

Fundamentals of Bayesian Epistemology 1

Introducing Credences

MICHAEL G. TITELBAUM

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

© Michael G. Titelbaum 2022

The moral rights of the author have been asserted

First Edition published in 2022

Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in
a retrieval system, or transmitted, in any form or by any means, without the
prior permission in writing of Oxford University Press, or as expressly permitted
by law, by licence or under terms agreed with the appropriate reprographics
rights organization. Enquiries concerning reproduction outside the scope of the
above should be sent to the Rights Department, Oxford University Press, at the
address above

You must not circulate this work in any other form
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2021949533

ISBN 978-0-19-870760-8 (hbk.)

ISBN 978-0-19-870761-5 (pbk.)

DOI: 10.1093/oso/9780198707608.001.0001

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY

Links to third party websites are provided by Oxford in good faith and
for information only. Oxford disclaims any responsibility for the materials
contained in any third party website referenced in this work.

Contents

<i>Quick Reference</i>	xv
<i>Preface</i>	xvii
0.1 What's in this book	xviii
0.2 How to read—and teach—this book	xix
0.3 Acknowledgments	xxi

VOLUME 1

I. OUR SUBJECT

1. Beliefs and Degrees of Belief	3
1.1 Binary beliefs	4
1.1.1 Classificatory, comparative, quantitative	4
1.1.2 Shortcomings of binary belief	5
1.2 From binary to graded	9
1.2.1 Comparative confidence	9
1.2.2 Bayesian epistemology	12
1.2.3 Relating beliefs and credences	14
1.3 The rest of this book	17
1.4 Exercises	19
1.5 Further reading	19

II. THE BAYESIAN FORMALISM

2. Probability Distributions	25
2.1 Propositions and propositional logic	25
2.1.1 Relations among propositions	28
2.1.2 State-descriptions	29
2.1.3 Predicate logic	30
2.2 The probability axioms	31
2.2.1 Consequences of the probability axioms	33
2.2.2 A Bayesian approach to the Lottery scenario	35
2.2.3 Doxastic possibilities	37
2.2.4 Probabilities are weird! The Conjunction Fallacy	38
2.3 Alternative representations of probability	39
2.3.1 Probabilities in Venn diagrams	39
2.3.2 Probability tables	41

2.3.3 Using probability tables	43
2.3.4 Odds	45
2.4 What the probability calculus adds	46
2.5 Exercises	48
2.6 Further reading	51
3. Conditional Credences	55
3.1 Conditional credences and the Ratio Formula	55
3.1.1 The Ratio Formula	56
3.1.2 Consequences of the Ratio Formula	59
3.1.3 Bayes's Theorem	61
3.2 Relevance and independence	64
3.2.1 Conditional independence and screening off	67
3.2.2 The Gambler's Fallacy	68
3.2.3 Probabilities are weird! Simpson's Paradox	70
3.2.4 Correlation and causation	72
3.3 Conditional credences and conditionals	75
3.4 Exercises	80
3.5 Further reading	85
4. Updating by Conditionalization	90
4.1 Conditionalization	90
4.1.1 Consequences of Conditionalization	94
4.1.2 Probabilities are weird! The Base Rate Fallacy	96
4.2 Evidence and certainty	99
4.2.1 Probabilities are weird! Total Evidence and the Monty Hall Problem	101
4.3 Priors and standards	105
4.3.1 Initial priors	105
4.3.2 Epistemic standards	107
4.3.3 Hypothetical priors	109
4.4 Exercises	114
4.5 Further reading	118
5. Further Rational Constraints	123
5.1 Subjective and Objective Bayesianism	124
5.1.1 Frequencies and propensities	124
5.1.2 Two distinctions in Bayesianism	129
5.2 Deference principles	133
5.2.1 The Principal Principle	133
5.2.2 Expert principles and Reflection	141
5.3 The Principle of Indifference	145
5.4 Credences for infinitely many possibilities	149
5.5 Jeffrey Conditionalization	156
5.6 Exercises	161
5.7 Further reading	164

<i>Glossary for Volume 1</i>	175
<i>Bibliography of Volume 1</i>	185
<i>Index of Names in Volume 1</i>	191

VOLUME 2

III. APPLICATIONS

6. Confirmation	195
6.1 Formal features of the confirmation relation	196
6.1.1 Confirmation is weird! The Paradox of the Ravens	196
6.1.2 Further adequacy conditions	200
6.2 Carnap's theory of confirmation	208
6.2.1 Confirmation as relevance	208
6.2.2 Finding the right function	210
6.3 Grue	215
6.4 Subjective Bayesian confirmation	221
6.4.1 Confirmation measures	225
6.4.2 Subjective Bayesian solutions to the Paradox of the Ravens	230
6.5 Exercises	235
6.6 Further reading	239
7. Decision Theory	246
7.1 Calculating expectations	247
7.1.1 The move to utility	250
7.2 Expected utility theory	251
7.2.1 Preference rankings and money pumps	251
7.2.2 Savage's expected utility	254
7.2.3 Jeffrey's theory	256
7.2.4 Risk aversion and Allais' Paradox	260
7.3 Causal Decision Theory	262
7.3.1 Newcomb's Problem	263
7.3.2 A causal approach	267
7.3.3 Responses and extensions	270
7.4 Exercises	271
7.5 Further reading	276

IV. ARGUMENTS FOR BAYESIANISM

8. Representation Theorems	285
8.1 Ramsey's four-step process	286
8.2 Savage's representation theorem	291
8.3 Representation theorems and probabilism	295
8.3.1 Objections to the argument	297
8.3.2 Reformulating the argument	300

X CONTENTS

8.4	Exercises	304
8.5	Further reading	306
9.	Dutch Book Arguments	310
9.1	Dutch Books	311
9.1.1	Dutch Books for probabilism	313
9.1.2	Further Dutch Books	315
9.2	The Dutch Book Argument	318
9.2.1	Dutch Books de pragmatized	321
9.3	Objections to Dutch Book Arguments	324
9.3.1	The Package Principle	326
9.3.2	Dutch Strategy objections	329
9.4	Exercises	333
9.5	Further reading	334
10.	Accuracy Arguments	338
10.1	Accuracy as calibration	339
10.2	The gradational accuracy argument for probabilism	343
10.2.1	The Brier score	343
10.2.2	Joyce's accuracy argument for probabilism	346
10.3	Objections to the accuracy argument for probabilism	351
10.3.1	The absolute-value score	351
10.3.2	Proper scoring rules	354
10.3.3	Are improper rules unacceptable?	359
10.4	Do we really need Finite Additivity?	361
10.5	An accuracy argument for Conditionalization	366
10.6	Exercises	368
10.7	Further reading	371

V. CHALLENGES AND OBJECTIONS

11.	Memory Loss and Self-locating Credences	381
11.1	Memory loss	382
11.1.1	The problem	382
11.1.2	A possible solution	385
11.1.3	Suppositional Consistency	389
11.2	Self-locating credences	392
11.2.1	The problem	393
11.2.2	The HTM approach	396
11.2.3	Going forward	401
11.3	Exercises	403
11.4	Further reading	407
12.	Old Evidence and Logical Omniscience	413
12.1	Old evidence	414
12.1.1	The problem	414

12.1.2 Solutions to the diachronic problem	417
12.1.3 Solutions to the synchronic problem	419
12.1.4 More radical solutions	424
12.2 Logical omniscience	428
12.2.1 Clutter avoidance and partial distributions	429
12.2.2 Logical confirmation and logical learning	432
12.2.3 Allowing logical uncertainty	433
12.2.4 Logical omniscience reconsidered	436
12.3 Exercises	439
12.4 Further reading	440
13. The Problem of the Priors and Alternatives to Bayesianism	445
13.1 The Problem of the Priors	446
13.1.1 Understanding the problem	448
13.1.2 Washing out of priors	453
13.2 Frequentism	457
13.2.1 Significance testing	459
13.2.2 Troubles with significance testing	461
13.3 Likelihoodism	467
13.3.1 Troubles with likelihoodism	470
13.4 Exercises	474
13.5 Further reading	476
14. Comparative Confidence, Ranged Credences, and Dempster-Shafer Theory	484
14.1 Comparative confidence	485
14.1.1 de Finetti's comparative conditions	486
14.1.2 The Scott Axiom	491
14.1.3 Extensions and challenges	496
14.2 Ranged credences	505
14.2.1 Ranged credences, representation, and evidence	510
14.2.2 Extensions and challenges	517
14.3 Dempster-Shafer theory	527
14.4 Exercises	533
14.5 Further reading	540
<i>Glossary for Volumes 1 & 2</i>	551
<i>Bibliography of Volumes 1 & 2</i>	573
<i>Index of Names in Volumes 1 & 2</i>	591

3

Conditional Credences

Chapter 2's discussion was confined to unconditional credence, an agent's outright degree of confidence that a particular proposition is true. This chapter takes up conditional credence, an agent's credence that one proposition is true on the supposition that another one is.

The main focus of this chapter is our fourth core normative Bayesian rule: the Ratio Formula. This rational constraint on conditional credences has a number of important consequences, including Bayes's Theorem (which gives Bayesianism its name).

Conditional credences are also central to the way Bayesians understand evidential relevance. I will define relevance as positive correlation, then explain how this notion has been used to investigate causal relations through the concept of screening off.

Having achieved a deeper understanding of the mathematics of conditional credences, I return at the end of the chapter to what exactly a conditional credence is. In particular, I discuss an argument by David Lewis that a conditional credence can't be understood as an unconditional credence in a conditional.

3.1 Conditional credences and the Ratio Formula

Arturo and Baxter know that two events will occur simultaneously in separate rooms: a fair coin will be flipped, and a clairvoyant will predict how it will land. Let H represent the proposition that the coin comes up heads, and C represent the proposition that the clairvoyant predicts heads. Suppose Arturo and Baxter each assign an unconditional credence of $1/2$ to H and an unconditional credence of $1/2$ to C .

Although Arturo and Baxter assign the same unconditional credences as each other to H and C , they still might take these propositions to be *related* in different ways. We could tease out those differences by saying to each agent, "I have no idea how the coin is going to come up or what the clairvoyant is going to say. But suppose for a moment the clairvoyant predicts heads. On this supposition, how confident are you that the coin will come up heads?" If

Arturo says 1/2 and Baxter says 99/100, that's a good indication that Baxter has more faith in the mystical than Arturo.

The quoted question in the previous paragraph elicits Arturo and Baxter's *conditional* credences, as opposed to the *unconditional* credences discussed in Chapter 2. An unconditional credence is a degree of belief assigned to a single proposition, indicating how confident the agent is that that proposition is true. A **conditional credence** is a degree of belief assigned to an ordered pair of propositions, indicating how confident the agent is that the first proposition is true on the supposition that the second is. We symbolize conditional credences as follows:

$$\text{cr}(H | C) = 1/2 \tag{3.1}$$

This equation says that a particular agent (in this case, Arturo) has a 1/2 credence that the coin comes up heads conditional on the supposition that the clairvoyant predicts heads. The vertical bar indicates a conditional credence; to the right of the bar is the proposition supposed; to the left of the bar is the proposition evaluated in light of that supposition. The proposition to the right of the bar is sometimes called the **condition**; I am not aware of any generally accepted name for the proposition on the left.

To be clear: A real agent never assigns any credences *ex nihilo*, without assuming at least some background information. An agent's unconditional credences in various propositions (such as *H*) are informed by her background information at that time. To assign a conditional credence, the agent combines her stock of background information with a *further* supposition that the condition is true. She then evaluates the other proposition in light of this combination.

A conditional credence is assigned to an *ordered* pair of propositions. It makes a difference which proposition is supposed and which is evaluated. Consider a case in which I'm going to roll a fair die and you have various credences involving the proposition *E* that it comes up even and the proposition 6 that it comes up six. Compare:

$$\text{cr}(6 | E) = 1/3 \tag{3.2}$$

$$\text{cr}(E | 6) = 1 \tag{3.3}$$

3.1.1 The Ratio Formula

Section 2.2 described Kolmogorov's probability axioms, which Bayesians take to represent rational constraints on an agent's unconditional credences.

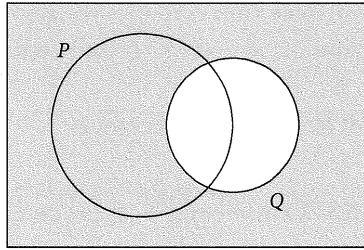


Figure 3.1 The region that dictates $\text{cr}(P | Q)$

Bayesians then add a constraint relating conditional to unconditional credences:

Ratio Formula: For any P and Q in \mathcal{L} , if $\text{cr}(Q) > 0$ then

$$\text{cr}(P | Q) = \frac{\text{cr}(P \& Q)}{\text{cr}(Q)}$$

Stated this way, the Ratio Formula remains silent on the value of $\text{cr}(P | Q)$ when $\text{cr}(Q) = 0$. There are various positions on how one should assign conditional credences when the condition has credence 0; we'll address some of them in our discussion of the infinite in Section 5.4.

Why should an agent's conditional credence equal the ratio of those unconditionals? Consider Figure 3.1. The rectangle represents all the possible worlds the agent entertains. The agent's unconditional credence in P is the fraction of that rectangle taken up by the P -circle. (The area of the rectangle is stipulated to be 1, so that fraction is the area of the P -circle divided by 1, which is just the area of the P -circle.) When we ask the agent to assign a credence conditional on the supposition that Q , she temporarily narrows her focus to just those possibilities that make Q true. In other words, she excludes from her attention the worlds I've shaded in the diagram, and considers only what's in the Q -circle. The agent's credence in P conditional on Q is the fraction of the Q -circle occupied by P -worlds. So it's the area of the PQ overlap divided by the area of the entire Q -circle, which is $\text{cr}(P \& Q)/\text{cr}(Q)$.

In the scenario in which I roll a fair die, your initial doxastic possibilities include all six outcomes of the die roll. I then ask for your credence that the die comes up six conditional on its coming up even—that is, $\text{cr}(6 | E)$. To assign this value, you exclude from consideration all the odd outcomes. You haven't actually *learned* that the die outcome is even; I've simply asked you to *suppose* that it comes up even and assign a confidence to other propositions in light of

that supposition. You distribute your credence equally over the outcomes that remain under consideration (2, 4, and 6), so your credence in six conditional on even is 1/3.

We get the same result from the Ratio Formula:

$$\text{cr}(6 | E) = \frac{\text{cr}(6 \& E)}{\text{cr}(E)} = \frac{1/6}{1/2} = \frac{1}{3} \quad (3.4)$$

The Ratio Formula allows us to calculate your conditional credences (confidences under a supposition) from your unconditional credences (confidences relative to no suppositions beyond your background information). Hopefully it's obvious why E gets an unconditional credence of 1/2 in this case; as for 6 & E , that's equivalent to just 6, so it gets an unconditional credence of 1/6.¹

Warning

Mathematicians often treat the Ratio Formula as a *definition* of conditional probability. From their point of view, a conditional probability has the value it does *in virtue of* two unconditional probabilities' standing in a certain ratio. But I do not want to reduce the possession of a conditional credence to the possession of two unconditional credences standing in a particular relation. I take a conditional credence to be a genuine mental state (an attitude toward an ordered pair of propositions) capable of being elicited in various ways, such as by asking an agent her confidence in a proposition given a supposition. So I will interpret the Ratio Formula as a rational constraint on how an agent's conditional credences should relate to her unconditional credences. As a normative *constraint* (rather than a *definition*), it can be violated—by assigning a conditional credence that doesn't equal the specified ratio.

The point of the previous warning is that the Ratio Formula is a rational constraint, and agents don't always meet all the rational constraints on their credences. Yet for agents who do satisfy the Ratio Formula, there can be no difference between their conditional credences without some difference in their unconditional credences as well. If we're both rational and I assign a different $\text{cr}(P | Q)$ value than you, we cannot assign the same values to both $\text{cr}(P \& Q)$ and $\text{cr}(Q)$. (A rational agent's conditional credences **supervene** on

her unconditional credences.) Fully specifying a rational agent's unconditional credence distribution suffices to specify her conditional credences as well.² For instance, we might specify Arturo's and Baxter's credence distributions using the following probability table:

C	H	cr_A	cr_B
T	T	1/4	99/200
T	F	1/4	1/200
F	T	1/4	1/200
F	F	1/4	99/200

Here cr_A represents Arturo's credences and cr_B represents Baxter's. Arturo's unconditional credence in C is identical to Baxter's—the values on the first two rows sum to 1/2 for each of them. Similarly, Arturo and Baxter have the same unconditional credence in H (the sum of the first and third rows). Yet Arturo and Baxter disagree in their confidence that the coin will come up heads (H) given that the clairvoyant predicts heads (C). Using the Ratio Formula, we calculate this conditional credence by dividing the value on the first row of the table by the sum of the values on the first two rows. This yields:

$$cr_A(H|C) = \frac{1/4}{1/2} = \frac{1}{2} \neq \frac{99}{100} = \frac{99/200}{100/200} = cr_B(H|C) \quad (3.5)$$

Baxter has high confidence in the clairvoyant's abilities. So on the supposition that the clairvoyant predicts heads, Baxter is almost certain that the flip comes up heads. Arturo, on the other hand, is skeptical, so supposing that the clairvoyant predicts heads leaves his opinions about the flip outcome unchanged.

3.1.2 Consequences of the Ratio Formula

Combining the Ratio Formula with the probability axioms yields further useful probability rules. First we have the

Law of Total Probability: For any proposition P and finite partition $\{Q_1, Q_2, \dots, Q_n\}$ in \mathcal{L} ,

$$cr(P) = cr(P|Q_1) \cdot cr(Q_1) + cr(P|Q_2) \cdot cr(Q_2) + \dots + cr(P|Q_n) \cdot cr(Q_n)$$

Suppose you're trying to predict whether I will bike to work tomorrow, but you're unsure if the weather will rain, hail, or be clear. The Law of Total Probability allows you to systematically work through the possibilities in that partition. You multiply your confidence that it will rain by your confidence that I'll bike should it rain. Then you multiply your confidence that it'll hail by your confidence in my biking given hail. Finally you multiply your unconditional credence that it'll be clear by your conditional credence that I'll bike given that it's clear. Adding these three products together yields your unconditional credence that I'll bike. (In the formula, the proposition that I'll bike plays the role of P and the three weather possibilities are Q_1 , Q_2 , and Q_3 .)

Next, the Ratio Formula connects conditional credences to Kolmogorov's axioms in a special way. Conditional credence is a two-place function, taking in an ordered pair of propositions and yielding a real number. Now suppose we designate some particular proposition R as our condition, and look at all of an agent's credences conditional on R . We now have a one-place function (because the second place has been filled by R) that we can think of as a distribution over the propositions in \mathcal{L} . Remarkably, if the agent's unconditional credences satisfy the probability axioms, then the Ratio Formula requires this conditional distribution $\text{cr}(\cdot | R)$ to satisfy those axioms as well. More formally, for any proposition R in \mathcal{L} such that $\text{cr}(R) > 0$, the following will all be true:

- For any proposition P in \mathcal{L} , $\text{cr}(P | R) \geq 0$.
- For any tautology \top in \mathcal{L} , $\text{cr}(\top | R) = 1$.
- For any mutually exclusive propositions P and Q in \mathcal{L} ,
 $\text{cr}(P \vee Q | R) = \text{cr}(P | R) + \text{cr}(Q | R)$.

(You'll prove these three facts in Exercise 3.4.)

Knowing that a conditional credence distribution is a probability distribution can be a handy shortcut. (It also has a significance for updating credences that we'll discuss in Chapter 4.) Because it's a probability distribution, a conditional credence distribution must satisfy all the consequences of the probability axioms we saw in Section 2.2.1. For example, if I tell you that $\text{cr}(P | R) = 0.7$, you can immediately tell that $\text{cr}(\sim P | R) = 0.3$, by the following conditional implementation of the Negation rule:

$$\text{cr}(\sim P | R) = 1 - \text{cr}(P | R) \tag{3.6}$$

Similarly, Entailment tells us that if $P \models Q$, then $\text{cr}(P | R) \leq \text{cr}(Q | R)$.

One special conditional distribution is worth investigating at this point: What happens when the condition R is a tautology? Imagine I ask you to report your unconditional credences in a bunch of propositions. Then I ask you to assign credences to those same propositions conditional on the further supposition of . . . nothing. I give you nothing more to suppose. Clearly you'll just report back to me the same credences. Bayesians represent vacuous information as a tautology, so this means that a rational agent's credences conditional on a tautology equal her unconditional credences. In other words, for any P in \mathcal{L} ,

$$\text{cr}(P|T) = \text{cr}(P) \quad (3.7)$$

This fact (whose proof I'll leave to the reader) will be important to our theory of updating later on.³

3.1.3 Bayes's Theorem

The most famous consequence of the Ratio Formula and Kolmogorov's axioms is

Bayes's Theorem: For any H and E in \mathcal{L} ,

$$\text{cr}(H|E) = \frac{\text{cr}(E|H) \cdot \text{cr}(H)}{\text{cr}(E)}$$

The first thing to say about Bayes's Theorem is *that it is a theorem*—it can be proven straightforwardly from the axioms and Ratio Formula. This is worth remembering, because there is a great deal of controversy about how Bayesians *apply* the theorem. (The significance they attach to this theorem is why Bayesians came to be called “Bayesians”.)

What philosophical significance could attach to an equation that is, in the end, just a truth of mathematics? The theorem was first articulated by the Reverend Thomas Bayes in the 1700s.⁴ Prior to Bayes, much of probability theory was concerned with problems of **direct inference**. Direct inference starts with the supposition of some probabilistic hypothesis, then asks how likely that hypothesis makes a particular experimental result. You probably learned to solve many direct inference problems in school, such as “Suppose I flip a fair coin 20 times; how likely am I to get exactly 19 heads?” Here the probabilistic

hypothesis H says that the coin is fair, while the experimental result E is that 20 flips yield exactly 19 heads. Your credence that the experimental result will occur on the supposition that the hypothesis is true— $\text{cr}(E | H)$ —is called the **likelihood**.⁵

Yet Bayes was also interested in **inverse inference**. Instead of making suppositions about hypotheses and determining probabilities of courses of evidence, his theorem allows us to calculate probabilities of hypotheses from suppositions about evidence. Instead of calculating the likelihood $\text{cr}(E | H)$, Bayes's Theorem shows us how to calculate $\text{cr}(H | E)$. A problem of inverse inference might ask, "Suppose a coin comes up heads on exactly 19 of 20 flips; how probable is it that the coin is fair?"

Assessing the significance of Bayes's Theorem, Hans Reichenbach wrote:

The method of indirect evidence, as this form of inquiry is called, consists of inferences that on closer analysis can be shown to follow the structure of the rule of Bayes. The physician's inferences, leading from the observed symptoms to the diagnosis of a specified disease, are of this type; so are the inferences of the historian determining the historical events that must be assumed for the explanation of recorded observations; and, likewise, the inferences of the detective concluding criminal actions from inconspicuous observable data. . . . Similarly, the general inductive inference from observational data to the validity of a given scientific theory must be regarded as an inference in terms of Bayes' rule. (Reichenbach 1935/1949, pp. 94–5)⁶

Here's an example of inverse inference: You're a biologist studying a particular species of fish, and you want to know whether the genetic allele coding for blue fins is dominant or recessive. Based on some other work you've done on fish, you're leaning toward recessive—initially you assign a 0.4 credence that the blue-fin allele is dominant. Given some background assumptions we won't worry about here,⁷ a direct inference from the theory of genetics tells you that if the allele is dominant, roughly three out of four species members will have blue fins; if the allele is recessive blue fins will appear on roughly 25% of the fish. But you're going to perform an inverse inference, making experimental observations to decide between genetic hypotheses. You will capture fish from the species at random and examine their fins. How significant will your first observation be to your credences in dominant versus recessive? When you contemplate various ways that observation might turn out, how should supposing one outcome or the other affect your credences about the allele? Before we do the calculation, try estimating how confident you should

be that the allele is dominant on the supposition that the first fish you observe has blue fins.

In this example our hypothesis H will be that the blue-fin allele is dominant. The evidence E to be supposed is that a randomly drawn fish has blue fins. We want to calculate the **posterior** value $\text{cr}(H|E)$. This value is called the “posterior” because it’s your credence in the hypothesis H *after* the evidence E has been supposed. In order to calculate this posterior, Bayes’s Theorem requires the values of $\text{cr}(E|H)$, $\text{cr}(H)$, and $\text{cr}(E)$.

$\text{cr}(E|H)$ is the likelihood of drawing a blue-finned fish on the hypothesis that the allele is dominant. On the supposition that the allele is dominant, 75% of the fish have blue fins, so your $\text{cr}(E|H)$ value should be 0.75. The other two values are known as **priors**; they are your unconditional credences in the hypothesis and the evidence *before* anything is supposed. We already said that your prior in the blue-fin dominant hypothesis H is 0.4. So $\text{cr}(H)$ is 0.4. But what about the prior in the evidence? How confident are you before observing any fish that the first one you draw will have blue fins?

Here we can apply the Law of Total Probability to the partition containing H and $\sim H$. This yields:

$$\text{cr}(E) = \text{cr}(E|H) \cdot \text{cr}(H) + \text{cr}(E|\sim H) \cdot \text{cr}(\sim H) \quad (3.8)$$

The values on the right-hand side are all either likelihoods, or priors related to the hypothesis. These values we can easily calculate. So

$$\text{cr}(E) = 0.75 \cdot 0.4 + 0.25 \cdot 0.6 = 0.45 \quad (3.9)$$

Plugging all these values into Bayes’s Theorem gives us

$$\text{cr}(H|E) = \frac{\text{cr}(E|H) \cdot \text{cr}(H)}{\text{cr}(E)} = \frac{0.75 \cdot 0.4}{0.45} = 2/3 \quad (3.10)$$

Observing a single fish has the potential to change your credences substantially. On the supposition that the fish you draw has blue fins, your credence that the blue-fin allele is dominant goes from its prior value of $2/5$ to a posterior of $2/3$.

Again, all of this is pure mathematics from a set of axioms that are rarely disputed. So why has Bayes’s Theorem been the focus of controversy? One issue is the role Bayesians give the theorem in *updating* attitudes over time; we’ll return to that application in Chapter 4. But the main idea Bayesians

take from Bayes—the idea that has proven controversial—is that probabilistic inverse inference is the key to induction. Bayesians think the primary way we ought to draw conclusions from data—how we ought to reason about scientific hypotheses, say, on the basis of experimental evidence—is by calculating posterior credences using Bayes’s Theorem. This view stands in direct conflict with other statistical methods, such as frequentism and likelihoodism. Advocates of those methods also have deep concerns about where Bayesians get the priors that Bayes’s Theorem requires. Once we’ve considerably deepened our understanding of Bayesian epistemology, we will discuss those concerns in Chapter 13, and assess frequentism and likelihoodism as alternatives to Bayesianism.

Before moving on, I’d like to highlight two useful alternative forms of Bayes’s Theorem. We’ve just seen that calculating the prior of the evidence— $\text{cr}(E)$ —can be easier if we break it up using the Law of Total Probability. Incorporating that maneuver into Bayes’s Theorem yields

$$\text{cr}(H | E) = \frac{\text{cr}(E | H) \cdot \text{cr}(H)}{\text{cr}(E | H) \cdot \text{cr}(H) + \text{cr}(E | \sim H) \cdot \text{cr}(\sim H)} \quad (3.11)$$

When a particular hypothesis H is under consideration, its negation $\sim H$ is known as the **catchall** hypothesis. So this form of Bayes’s Theorem calculates the posterior in the hypothesis from the priors and likelihoods of the hypothesis and its catchall.

In other situations we have multiple hypotheses under consideration instead of just one. Given a finite partition of n hypotheses $\{H_1, H_2, \dots, H_n\}$, the Law of Total Probability transforms the denominator of Bayes’s Theorem to yield

$$\text{cr}(H_i | E) = \frac{\text{cr}(E | H_i) \cdot \text{cr}(H_i)}{\text{cr}(E | H_1) \cdot \text{cr}(H_1) + \text{cr}(E | H_2) \cdot \text{cr}(H_2) + \dots + \text{cr}(E | H_n) \cdot \text{cr}(H_n)} \quad (3.12)$$

This version allows you to calculate the posterior of one particular hypothesis H_i in the partition from the priors and likelihoods of all the hypotheses.

3.2 Relevance and independence

Arturo doesn’t believe in hocus pocus; from his point of view, information about what a clairvoyant predicts is irrelevant to determining how a coin flip will come out. So supposing that a clairvoyant predicts heads makes no

difference to Arturo's confidence in a heads outcome. If C says the clairvoyant predicts heads, H says the coin lands heads, and cr_A is Arturo's credence distribution, we have

$$cr_A(H | C) = 1/2 = cr_A(H) \quad (3.13)$$

Generalizing this idea yields a key definition: Proposition P is **probabilistically independent** of proposition Q relative to distribution cr just in case

$$cr(P | Q) = cr(P) \quad (3.14)$$

In this case Bayesians also say that Q is **irrelevant** to P . When Q is irrelevant to P , supposing Q leaves an agent's credence in P unchanged.⁸

Notice that probabilistic independence is always relative to a distribution cr . The very same propositions P and Q might be independent relative to one distribution but dependent relative to another. (Relative to Arturo's credences the clairvoyant's prediction is irrelevant to the flip outcome, but relative to the credences of his friend Baxter—who believes in psychic powers—it is not.) In what follows I may omit reference to a particular distribution when context makes it clear, but you should keep the relativity of independence to a probability distribution in the back of your mind.

While Equation (3.14) will be our official *definition* of probabilistic independence, there are many equivalent tests for independence. Given the probability axioms and Ratio Formula, the following equations are all true just when Equation (3.14) is:

$$cr(P) = cr(P | \sim Q) \quad (3.15)$$

$$cr(P | Q) = cr(P | \sim Q) \quad (3.16)$$

$$cr(Q | P) = cr(Q) = cr(Q | \sim P) \quad (3.17)$$

$$cr(P \& Q) = cr(P) \cdot cr(Q) \quad (3.18)$$

The equivalence of Equations (3.14) and (3.15) tells us that Q is probabilistically independent of P just in case $\sim Q$ is. The equivalence of (3.14) and (3.17) shows us that independence is symmetric: if supposing Q makes no difference to an agent's credence in P , then supposing P won't change that agent's attitude toward Q . Finally, Equation (3.18) embodies a useful probability rule:

Multiplication: P and Q are probabilistically independent relative to cr if and only if $cr(P \& Q) = cr(P) \cdot cr(Q)$.

(Some authors define probabilistic independence using this biconditional, but we will define independence using Equation (3.14), then treat Multiplication as a consequence.)

We can calculate the probability of a conjunction by multiplying the probabilities of its conjuncts only when those conjuncts are *independent*. This trick will not work for any arbitrary propositions. The general formula for probability in a conjunction can be derived quickly from the Ratio Formula:

$$\text{cr}(P \& Q) = \text{cr}(P | Q) \cdot \text{cr}(Q) \quad (3.19)$$

When P and Q are probabilistically independent, the first term on the right-hand side equals $\text{cr}(P)$.

It's important not to get Multiplication and Finite Additivity confused. Finite Additivity says that the credence of a *disjunction* is the *sum* of the credences of its *mutually exclusive* disjuncts. Multiplication says that the credence of a *conjunction* is the *product* of the credences of its *independent* conjuncts. If I flip a fair coin twice in succession, heads on the first flip and heads on the second flip are independent, while heads on the first flip and tails on the first flip are mutually exclusive.

When two propositions fail to be probabilistically independent (relative to a particular distribution), we say those propositions are **relevant** to each other. Replace the “=” signs in Equations (3.14) through (3.18) with “>” signs and you have tests for Q 's being **positively relevant** to P . Once more the tests are equivalent—if any of the resulting inequalities is true, all of them are. Q is positively relevant to P when assuming Q makes you more confident in P . For example, since Baxter believes in mysticism, he takes the clairvoyant's predictions to be highly relevant to the outcome of the coin flip—supposing that the clairvoyant has predicted heads takes him from equanimity to near-certainty in a heads outcome. Baxter assigns

$$\text{cr}_B(H | C) = 99/100 > 1/2 = \text{cr}_B(H) \quad (3.20)$$

Like independence, positive relevance is symmetric. Supposing that the coin came up heads will make Baxter highly confident that the clairvoyant predicted it would.

Similarly, replacing the “=” signs with “<” signs above yields tests for **negative relevance**. For Baxter, the clairvoyant's predicting heads is negatively relevant to the coin's coming up tails. Like positive correlation, negative correlation is symmetric (supposing a tails outcome makes Baxter less

confident in a heads prediction). Note also that there are many synonyms in the statistics community for “relevance”. Instead of finding “positively/negatively relevant” terminology, you’ll sometimes find “positively/negatively dependent”, “positively/negatively correlated”, or even “correlated/anti-correlated”.

The strongest forms of positive and negative relevance are entailment and refutation. Suppose a hypothesis H has nonextreme prior credence. If a particular piece of evidence E entails the hypothesis, the probability axioms and Ratio Formula tell us that

$$\text{cr}(H | E) = 1 \quad (3.21)$$

Supposing E takes H from a middling credence to the highest credence allowed. Similarly, if E refutes H (what philosophers of science call **falsification**), then

$$\text{cr}(H | E) = 0 \quad (3.22)$$

Relevance will be most important to us because of its connection to confirmation, the Bayesian notion of evidential support. A piece of evidence confirms a hypothesis only if it’s relevant to that hypothesis. Put another way, learning a piece of evidence changes a rational agent’s credence in a hypothesis only if that evidence is relevant to the hypothesis. (Much more on this later.)

3.2.1 Conditional independence and screening off

The definition of probabilistic independence compares an agent’s conditional credence in a proposition to her unconditional credence in that proposition. But we can also compare conditional credences. When Baxter, who believes in the occult, hears a clairvoyant’s prediction about the outcome of a fair coin flip, he takes it to be highly correlated with the true flip outcome. But what if we ask Baxter to suppose that this particular clairvoyant is an impostor? Once he supposes the clairvoyant is an impostor, Baxter may take the clairvoyant’s predictions to be completely irrelevant to the flip outcome. Let C be the proposition that the clairvoyant predicts heads, H be the proposition that the coin comes up heads, and I be the proposition that the clairvoyant is an impostor. It’s possible for Baxter’s credences to satisfy both of the following equations at once:

$$\text{cr}(H | C) > \text{cr}(H) \tag{3.23}$$

$$\text{cr}(H | C \ \& \ I) = \text{cr}(H | I) \tag{3.24}$$

Inequality (3.23) tells us that unconditionally, Baxter takes C to be relevant to H . But conditional on the supposition of I , C becomes independent of H (Equation (3.24)); once Baxter has supposed I , adding C to his suppositions doesn't affect his confidence in H .

Generalizing this idea yields the following definition of **conditional independence**: P is probabilistically independent of Q conditional on R just in case

$$\text{cr}(P | Q \ \& \ R) = \text{cr}(P | R) \tag{3.25}$$

If this equality fails to hold, we say that Q is relevant to P conditional on R .

One more piece of terminology: We will say that R **screens off** P from Q when Q is unconditionally relevant to P , but irrelevant to P conditional on each of R and $\sim R$. That is, R screens off P from Q just in case all three of the following are satisfied:

$$\text{cr}(P | Q) \neq \text{cr}(P) \tag{3.26}$$

$$\text{cr}(P | Q \ \& \ R) = \text{cr}(P | R) \tag{3.27}$$

$$\text{cr}(P | Q \ \& \ \sim R) = \text{cr}(P | \sim R) \tag{3.28}$$

When these equations are met, P and Q are correlated but supposing either R or $\sim R$ makes that correlation disappear.⁹

Conditional independence and screening off are both best understood through real-world examples. We'll see a number of those in the next few sections.

3.2.2 The Gambler's Fallacy

People often act as if future chancy events will "compensate" for unexpected past results. When a good hitter strikes out many times in a row, someone will say he's "due" for a hit. If a fair coin comes up heads nineteen times in a row, many people become more confident that the next outcome will be tails.

This mistake is known as the **Gambler's Fallacy**.¹⁰ A person who makes the mistake is thinking along something like the following lines: In twenty flips of a fair coin, it's more probable to get nineteen heads and one tail than it is to

get twenty heads. So having seen nineteen heads, it's much more likely that the twentieth flip will come up tails.

This person is providing the right answer to the wrong question. If the question is "When a fair coin is flipped twenty times, is it more likely that you'll get a *total* of nineteen heads and one tail than it is that you'll get twenty heads?", the answer to that question is "yes"—in fact, it's twenty times as likely! But that's the wrong question to ask in this case. Instead of wondering what sorts of total outcomes are probable when one flips a fair coin twenty times, in this case it's more appropriate to ask: *given* that the coin has already come up heads nineteen times, how confident are we that the twentieth flip will be tails? This is a question about our conditional credence

$$\text{cr}(20\text{th flip tails} \mid \text{first 19 flips heads}) \quad (3.29)$$

How should we calculate this conditional credence? Ironically, it might be more reasonable to make a mistake in the *opposite* direction from the Gambler's Fallacy. If I see a coin come up heads nineteen times, shouldn't that make me suspect that it's biased toward heads? If anything, shouldn't supposing nineteen consecutive heads make me *less* confident that the twentieth flip will come up tails?

This line of reasoning would be appropriate to the present case if we hadn't stipulated that the coin is fair. For a rational agent, the outcome of the twentieth flip is probabilistically independent of the outcomes of the first nineteen flips conditional on the fact that the coin is fair. That is,

$$\begin{aligned} \text{cr}(20\text{th flip tails} \mid \text{first 19 flips heads \& fair coin}) = \\ \text{cr}(20\text{th flip tails} \mid \text{fair coin}) \end{aligned} \quad (3.30)$$

We can justify this equation as follows: Typically, information that a coin came up heads nineteen times in a row would alter your opinion about whether it's a fair coin. Changing your opinion about whether it's a fair coin would then affect your prediction for the twentieth flip. So typically, information about the first nineteen flips alters your credences about the twentieth flip *by way of* your opinion about whether the coin is fair. But if you've already established that the coin is fair, information about the first nineteen flips has no further significance for your prediction about the twentieth. So conditional on the coin's being fair, the first nineteen flips' outcomes are irrelevant to the outcome of the twentieth flip.

The left-hand side of Equation (3.30) captures the correct question to ask about the Gambler's Fallacy case. The right-hand side is easy to calculate; it's $1/2$. So after seeing a coin known to be fair come up heads nineteen times, we should be $1/2$ confident that the twentieth flip will be tails.¹¹

3.2.3 Probabilities are weird! Simpson's Paradox

Perhaps you're too much of a probabilistic sophisticate to ever commit the Gambler's Fallacy. Perhaps you successfully navigated Tversky and Kahneman's Conjunction Fallacy (Section 2.2.4) as well. But even probability experts sometimes have trouble with the counterintuitive relations that arise between conditional and unconditional dependence.

Here's an example of how odd things can get: In a famous case, the University of California, Berkeley's graduate departments were investigated for gender bias in admissions. The concern arose because in 1973 about 44% of overall male applicants were admitted to graduate school at Berkeley, while only 35% of female applicants were. Yet when the graduate departments (where admissions decisions are made) were studied one at a time, it turned out that individual departments either were admitting men and women at roughly equal rates, or in some cases were admitting a higher percentage of female applicants.

This is an example of **Simpson's Paradox**, in which probabilistic dependencies point in one direction conditional on each member of a partition, yet point the opposite way unconditionally. A Simpson's Paradox case involves a collection with a number of subgroups. Each of the subgroups displays the same correlation between two traits. Yet when we examine the collection as a whole, that correlation is reversed!¹²

To see how this can happen, consider another example: Over the course of the 2016–2017 NBA season, Houston Rockets player James Harden made a higher percentage of his two-point shots than Toronto Raptors player DeMar DeRozan did. Harden also made a higher percentage of his three-point shots than DeRozan. Yet when we look at all the shots attempted (both two- and three-pointers), DeRozan made a higher percentage than Harden overall.¹³

Here are the data for the two players:

	Two-pointers		Three-pointers		Combined	
DeRozan	688/1421	48.4%	33/124	26.6%	721/1545	46.7%
Harden	412/777	53.0%	262/756	34.7%	674/1533	44.0%

The second number in each box is the number of attempts; the first is the number of makes; the third is the percentage (makes divided by attempts). Looking at the table, you can see how Harden managed to shoot better than DeRozan from each distance yet have a worse shooting percentage overall. Since three-pointers are more difficult to make than two-pointers, each player made a much higher percentage of his two-point attempts than his three-point attempts. DeRozan shot over 90% of his shots from the easier-to-make two-point distance. Harden, on the other hand, shot almost half of his shots from downtown. Harden was taking many more low-percentage shots than DeRozan, so even though Harden was better at those shots, his overall percentage suffered.

Scrutiny revealed a similar effect in Berkeley's 1973 admissions data. Bickel, Hammel, and O'Connell (1975) concluded, "The proportion of women applicants tends to be high in departments that are hard to get into and low in those that are easy to get into." Although individual departments were admitting women at comparable rates to men, female applications were less successful overall because more were directed at departments with low admission rates.¹⁴

Simpson's Paradox can be thought of entirely in terms of numerical proportions, as we've just done with the basketball and admissions examples. But these examples can also be analyzed using conditional probabilities. Suppose, for instance, that you are going to select a Harden or DeRozan shot attempt at random from the 2016–2017 season, making your selection so that each of the 3,078 attempts they put up together is equally likely to be selected. How confident should you be that the selected attempt will be a make? How should that confidence change on the supposition that a DeRozan attempt is selected, or a two-point attempt?

Below is a probability table for your credences, assembled from the real-life statistics above. Here D says that it's a DeRozan attempt; 2 says it's a two-pointer; and M says it's a make. (Given the pool from which we're sampling, $\sim D$ means a Harden attempt and ~ 2 means it's a three-pointer.)

D	2	M	cr
T	T	T	688/3078
T	T	F	733/3078
T	F	T	33/3078
T	F	F	91/3078
F	T	T	412/3078
F	T	F	365/3078
F	F	T	262/3078
F	F	F	494/3078

A bit of calculation with this probability table reveals the following:

$$\text{cr}(M | D) > \text{cr}(M | \sim D) \quad (3.31)$$

$$\text{cr}(M | D \& 2) < \text{cr}(M | \sim D \& 2) \quad (3.32)$$

$$\text{cr}(M | D \& \sim 2) < \text{cr}(M | \sim D \& \sim 2) \quad (3.33)$$

If you're selecting an attempt from the total sample, DeRozan is more likely to make it than Harden. Put another way, DeRozan's taking the attempt is unconditionally positively relevant to its being made (Equation (3.31)). But DeRozan's shooting is negatively relevant to a make conditional on each of the two distances (Equations (3.32) and (3.33)). If you're selecting from only the two-pointers, or from only the three-pointers, the shot is more likely to be made if it's attempted by Harden.

3.2.4 Correlation and causation

You may have heard the expression "correlation is not causation." People typically use this expression to point out that just because two events have both occurred—and maybe occurred in close spatio-temporal proximity—that doesn't mean they had anything to do with each other. But "correlation" is a technical term in probability discussions. The propositions describing two events may both be true, or you might have high credence in both of them, yet they still might not be probabilistically correlated. For the propositions to be correlated, supposing one to be true must *change* the probability of the other. I'm confident that I'm under 6 feet tall and that my eyes are blue, but that doesn't mean I take those facts to be correlated.

Once we've understood probabilistic correlation correctly, does its presence always indicate a causal connection? When two propositions about empirical events are correlated, must the event described by one cause the event described by the other? Hans Reichenbach offered a classic counterexample to this proposal. He wrote:

Suppose two geysers which are not far apart spout irregularly, but throw up their columns of water always at the same time. The existence of a subterranean connection of the two geysers with a common reservoir of hot water is then practically certain. (1956, p. 158)

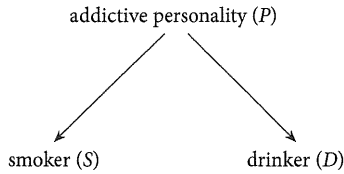


Figure 3.2 A causal fork

If you've noticed that two nearby geysers always spout simultaneously, seeing one spout will increase your confidence that the other is spouting as well. So your credences about the geysers are correlated. But you don't think one geyser's spouting *causes* the other to spout. Instead, you hypothesize an unobserved reservoir of hot water that is the **common cause** of both spouts.

Reichenbach proposed a famous principle about empirically correlated events:

Principle of the Common Cause: When event outcomes are probabilistically correlated, either one causes the other or they have a common cause.¹⁵

In addition to this principle, he offered a key mathematical insight about causation: a common cause screens its effects off from each other.

Let's work through an example of this insight concerning causation and screening off. Suppose the proposition that a particular individual is a drinker is positively relevant to the proposition that she's a smoker. According to the Principle of the Common Cause, there must be some causal link between these propositions. Perhaps drinking causes smoking—drinking creates situations in which one is more likely to smoke—or vice versa. Or they may be linked through a common cause: maybe being a smoker and being a drinker are both promoted by an addictive personality, which we can imagine results from a genetic endowment unaffected by one's behavior. (See Figure 3.2; the arrows indicate causal influence.)

Imagine the latter explanation is true, and moreover is the *only* true explanation of the correlation between drinking and smoking. That is, being a smoker and being a drinker are positively correlated only due to their both being caused by an addictive personality. Given this assumption, let's take a particular subject whose personality you're unsure about, and consider what happens to your credences when you make various suppositions about her.

If you begin by supposing that the subject drinks, this will make you more confident that she smokes—but *only because it makes you more confident*

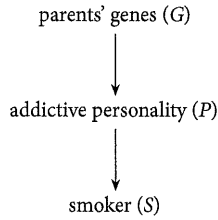


Figure 3.3 A causal chain

that the subject has an addictive personality. On the other hand, you might start by supposing that the subject has an addictive personality. That will certainly make you more confident that she's a smoker. But once you've made that adjustment, going on to suppose that she's a drinker won't affect your confidence in smoking. Information about drinking affects your smoking opinions only *by way of* helping you figure out whether she has an addictive personality, and the answer to the personality question was filled in by your initial supposition. Once an addictive personality is supposed, drinking has no further relevance to smoking. (Compare: Once a coin is supposed to be fair, the outcomes of its first nineteen flips have no relevance to the outcome of the twentieth.) Drinking becomes probabilistically independent of smoking conditional on suppositions about whether the subject has an addictive personality. That is,

$$\text{cr}(S | D) > \text{cr}(S) \quad (3.34)$$

$$\text{cr}(S | D \& P) = \text{cr}(S | P) \quad (3.35)$$

$$\text{cr}(S | D \& \sim P) = \text{cr}(S | \sim P) \quad (3.36)$$

Causal forks (as in Figure 3.2) give rise to screening off. P is a common cause of S and D , so P screens off S from D .

But that's not the only way screening off can occur. Consider Figure 3.3. Here we've focused on a different portion of the causal structure. Imagine that the subject's parents' genes causally influence whether she has an addictive personality, which in turn causally promotes smoking. Now her parents' genetics are probabilistically relevant to the subject's smoking, but that correlation is screened off by facts about her personality. Again, if you're uncertain whether the subject's personality is addictive, facts about her parents' genes will affect your opinion of whether she's a smoker. But once you've made a firm supposition about the subject's personality, suppositions about her parents' genetics have no further influence on your smoking opinions. In equation form:

$$\text{cr}(S | G) > \text{cr}(S) \quad (3.37)$$

$$\text{cr}(S | G \& P) = \text{cr}(S | P) \quad (3.38)$$

$$\text{cr}(S | G \& \sim P) = \text{cr}(S | \sim P) \quad (3.39)$$

P screens off S from G .¹⁶

Relevance, conditional relevance, and causation can interact in very complex ways.¹⁷ My goal here has been to introduce the main ideas and terminology employed in their analysis. The state of the art in this field has come a long way since Reichenbach; computational tools now available can look at statistical correlations among a large number of variables and hypothesize a causal structure lying beneath them. The resulting causal diagrams are known as **Bayes Nets**, and have practical applications from satellites to health care to car insurance to college admissions.

These causal methods all start from Reichenbach's insight that common causes screen off their effects. And what of his more metaphysically radical Principle of the Common Cause? It remains highly controversial.

3.3 Conditional credences and conditionals

I now want to circle back and get clearer on the nature of conditional credence. First, it's important to note that the conditional credences we've been discussing are indicative, not subjunctive. This distinction is familiar from the theory of conditional propositions. Compare:

If Shakespeare didn't write *Hamlet*, then someone else did.

If Shakespeare hadn't written *Hamlet*, then someone else would have.

The former conditional is indicative, while the latter is subjunctive. The traditional distinction between these two types of conditional begins with the assumption that a conditional is evaluated by considering possible worlds in which the antecedent is satisfied, then checking whether the consequent is true in those worlds as well. In evaluating an indicative conditional, the antecedent worlds are restricted to those among the agent's doxastic possibilities.¹⁸ Evaluating a subjunctive conditional, on the other hand, permits *counterfactual* reasoning involving antecedent worlds considered non-actual. So when you assess the subjunctive conditional above, you are allowed to consider worlds that make the antecedent true by making *Hamlet* never exist. But when evaluating the indicative conditional, you have to take into account

that *Hamlet* actually does exist, and entertain only worlds in which that's true. So you consider bizarre "author-conspiracy" worlds that, while far-fetched, satisfy the antecedent and are among your current doxastic possibilities. In the end, I'm guessing you take the indicative conditional to be true but the subjunctive to be false.

Now suppose I ask for your credence in the proposition that someone wrote *Hamlet*, conditional on the supposition that Shakespeare didn't. This value will be high, again because you take *Hamlet* to exist. In assigning this conditional credence, you aren't bringing into consideration possible worlds you'd otherwise ruled out (such as *Hamlet*-free worlds). Instead, you're focusing in on the narrow set of author-conspiracy worlds you currently entertain. As we saw in Figure 3.1, assigning a conditional credence strictly narrows the worlds under consideration; it doesn't expand your attention to worlds previously ruled out. Thus the conditional credences discussed in this book—and typically discussed in the Bayesian literature—are indicative rather than subjunctive.¹⁹

Are there more features of conditional propositions that can help us understand conditional credences? Might we understand conditional credences *in terms of* conditionals? Initiating his study of conditional degrees of belief, F.P. Ramsey warned against assimilating them to conditional propositions:

We are also able to define a very useful new idea—"the degree of belief in *p* given *q*". This does not mean the degree of belief in "If *p* then *q*", or that in "*p* entails *q*", or that which the subject would have in *p* if he knew *q*, or that which he ought to have. (1931, p. 82)²⁰

Yet many authors failed to heed Ramsey's warning. It's very tempting to equate conditional credences with some simple combination of conditional propositions and unconditional credences. For example, when I ask, "How confident are you in *P* given *Q*?", it's easy to hear that as "Given *Q*, how confident are you in *P*?" or just "If *Q* is true, how confident are you in *P*?" This simple slide might suggest that for any real number *r* and propositions *P* and *Q*,

$$\text{"cr}(P | Q) = r\text{" is equivalent to } "Q \rightarrow \text{cr}(P) = r\text{"} \quad (3.40)$$

Here I'm using the symbol " \rightarrow " to represent some kind of conditional. For the reasons discussed above, it should be an indicative conditional. But it need not be the material conditional symbolized by " \supset "; many authors think the material conditional's truth-function fails to accurately represent the meaning of natural-language indicative conditionals.

Endorsing the equivalence in (3.40) would require serious changes to the traditional logic of conditionals. We can demonstrate this in two ways. First, we usually take indicative conditionals to satisfy the disjunctive syllogism rule. (The material conditional certainly does!) This rule tells us that

$$"X \rightarrow Z" \text{ and } "Y \rightarrow Z" \text{ jointly entail } "(X \vee Y) \rightarrow Z" \quad (3.41)$$

for any propositions X , Y , and Z . Thus for any propositions A , B , and C and constant k we have

$$"A \rightarrow [\text{cr}(C) = k]" \text{ and } "B \rightarrow [\text{cr}(C) = k]" \text{ entail } "(A \vee B) \rightarrow [\text{cr}(C) = k]" \quad (3.42)$$

Combining (3.40) and (3.42) yields

$$"\text{cr}(C|A) = k" \text{ and } "\text{cr}(C|B) = k" \text{ entail } "\text{cr}(C|A \vee B) = k" \quad (3.43)$$

(3.43) may look appealing, as a sort of probabilistic analog of disjunctive syllogism. But it's false. Not only can one design a credence distribution satisfying the probability axioms and Ratio Formula such that $\text{cr}(C|A) = k$ and $\text{cr}(C|B) = k$ but $\text{cr}(C|A \vee B) \neq k$; one can even describe real-life examples in which it's rational for an agent to assign such a distribution. (See Exercise 3.14.) The failure of (3.43) is another case in which credences confound expectations developed by our experiences with classificatory states.

Second, we usually take indicative conditionals to satisfy *modus tollens*. But consider the following facts about me: Unconditionally, I am highly confident that I will be alive tomorrow. But conditional on the proposition that the sun just exploded, my confidence that I will be alive tomorrow is very low. Given these facts, *modus tollens*, and (3.40), I could run the following argument:

$$\text{cr}(\text{alive tomorrow} | \text{sun exploded}) \text{ is low.} \quad [\text{given}] \quad (3.44)$$

$$\text{If the sun exploded, cr}(\text{alive tomorrow}) \text{ is low.} \quad [(3.44), (3.40)] \quad (3.45)$$

$$\text{cr}(\text{alive tomorrow}) \text{ is high.} \quad [\text{given}] \quad (3.46)$$

$$\text{The sun did not explode.} \quad [\textit{modus tollens}] \quad (3.47)$$

While intriguing for its promise of astronomy by introspection, this argument is unsound. So I conclude that as long as indicative conditionals satisfy classical logical rules such as disjunctive syllogism and *modus tollens*, any analysis of conditional credences in terms of conditionals that uses (3.40) must be false.²¹

Perhaps we've mangled the transition from conditional credences to conditional propositions. Perhaps we should hear "How confident are you in P given Q ?" as "How confident are you in ' P , given Q '?", which is in turn "How confident are you in 'If Q , then P '?" Maybe a conditional credence is a credence in a conditional. Or perhaps more weakly: an agent assigns a particular conditional credence value whenever she unconditionally assigns that value to a conditional. In symbols, the proposal is that

$$\text{"cr}(P | Q) = r\text{" is equivalent to } \text{"cr}(Q \rightarrow P) = r\text{"} \quad (3.48)$$

for any real r , any propositions P and Q , any credence distribution cr , and some indicative conditional \rightarrow . If true, this equivalence would offer another possibility for analyzing conditional credences in terms of unconditional credences and conditionals.

We can quickly show that (3.48) fails if " \rightarrow " is read as the material conditional \supset . Under the material reading, the proposal entails that

$$\text{cr}(P | Q) = \text{cr}(Q \supset P) \quad (3.49)$$

Using the probability calculus and Ratio Formula, we can show that Equation (3.49) holds only when $\text{cr}(Q) = 1$ or $\text{cr}(Q \supset P) = 1$. (See Exercise 3.15.) This is a *triviality result*: It shows that Equation (3.49) can hold only in trivial cases, namely over the narrow range of conditionals for which the agent is either certain of the antecedent or certain of the conditional itself. Equation (3.49) does not express a truth that holds for *all* conditional credences in *all* propositions; nor does (3.48) when " \rightarrow " is read materially.

Perhaps the equivalence in (3.48) can be saved from this objection by construing its " \rightarrow " as something other than a material conditional. But David Lewis (1976) provided a clever objection that works whichever conditional \rightarrow we choose. Begin by selecting arbitrary propositions P and Q . We then derive the following from the proposal in Equation (3.48):

$$\text{cr}(Q \rightarrow P) = \text{cr}(P | Q) \quad [\text{from (3.48)}] \quad (3.50)$$

$$\text{cr}(Q \rightarrow P | P) = \text{cr}(P | Q \& P) \quad [\text{see below}] \quad (3.51)$$

$$\text{cr}(Q \rightarrow P | P) = 1 \quad [Q \& P \text{ entails } P] \quad (3.52)$$

$$\text{cr}(Q \rightarrow P | \sim P) = \text{cr}(P | Q \& \sim P) \quad [\text{see below}] \quad (3.53)$$

$$\text{cr}(Q \rightarrow P | \sim P) = 0 \quad [Q \& \sim P \text{ refutes } P] \quad (3.54)$$

$$\text{cr}(Q \rightarrow P) = \text{cr}(Q \rightarrow P | P) \cdot \text{cr}(P) + \text{cr}(Q \rightarrow P | \sim P) \cdot \text{cr}(\sim P) \quad [\text{Law of Total Prob.}] \quad (3.55)$$

$$\text{cr}(Q \rightarrow P) = 1 \cdot \text{cr}(P) + 0 \cdot \text{cr}(\sim P) \quad [(3.52), (3.54), (3.55)] \quad (3.56)$$

$$\text{cr}(Q \rightarrow P) = \text{cr}(P) \quad (3.57)$$

$$\text{cr}(P | Q) = \text{cr}(P) \quad [(3.50)] \quad (3.58)$$

Some of these lines require explanation. The idea of lines (3.51) and (3.53) is this: We've already seen that a credence distribution conditional on a particular proposition satisfies the probability axioms. This suggests that we should think of a distribution conditional on a proposition as being just like any other credence distribution. (We'll see more reason to think this in Chapter 4, note 3.) So a distribution conditional on a proposition should satisfy the proposal of (3.48) as well. If you conditionally suppose X , then under that supposition you should assign $Y \rightarrow Z$ the same credence you would assign Z were you to *further* suppose Y . In other words, anyone who maintains (3.48) should also maintain that for any X , Y , and Z ,

$$\text{cr}(Y \rightarrow Z | X) = \text{cr}(Z | Y \& X) \quad (3.59)$$

In line (3.51) the roles of X , Y , and Z are played by P , Q , and P ; in line (3.53) it's $\sim P$, Q , and P .

Lewis has given us another triviality result. Assuming the probability axioms and Ratio Formula, the proposal in (3.48) can hold only for propositions P and Q such that $\text{cr}(P | Q) = \text{cr}(P)$. In other words, it can hold only for propositions the agent takes to be independent.²² Or (taking things from the other end), the proposed equivalence can hold for all the conditionals in an agent's language only if the agent treats every pair of propositions in \mathcal{L} as independent!²³

So a rational agent's conditional credence will not in general equal her unconditional credence in a conditional. This is not to say that conditional credences have *nothing* to do with conditionals. A popular idea now usually called "Adams's Thesis" (Adams 1965) holds that an indicative conditional $Q \rightarrow P$ is *acceptable* to a degree equal to $\text{cr}(P | Q)$.²⁴ But we cannot maintain that an agent's conditional credence is always equal to her credence that some conditional is *true*.

This brings us back to a proposal I discussed in Chapter 1. One might try to relate degrees of belief to binary beliefs by suggesting that whenever an agent has an r -valued credence, she has a binary belief in a traditional proposition with r as part of its content. Working out this proposal for conditional

credences reveals how hopeless it is. Suppose an agent assigns $\text{cr}(P|Q) = r$. Would we suggest that the agent believes that if Q , then the probability of P is r ? This proposal mangles the logic of conditional credences. Perhaps the agent believes that the probability of “if Q , then P ” is r ? Lewis’s argument dooms this idea.

I said in Chapter 1 that the numerical value of an unconditional degree of belief is an attribute of the *attitude taken* toward a proposition, not a *constituent of* that proposition itself. As for conditional credences, $\text{cr}(P|Q) = r$ does not say that an agent takes some attitude toward a conditional proposition with a probability value in its consequent. Nor does it say that the agent takes some attitude toward a single, conditional proposition composed of P and Q . $\text{cr}(P|Q) = r$ says that the agent takes an r -valued attitude toward an *ordered pair* of propositions—neither of which need involve the number r .

3.4 Exercises

Unless otherwise noted, you should assume when completing these exercises that the credence distributions under discussion satisfy the probability axioms and Ratio Formula. You may also assume that whenever a conditional credence expression occurs, the condition has a nonzero unconditional credence so that the conditional credence is well defined.

Problem 3.1. ♪ A family has two children of different ages. Assume that each child has a probability of $1/2$ of being a girl, and that the probability that the elder is a girl is independent of the probability that the younger is.

- (a) Conditional on the older child’s being a girl, what’s the probability that the younger one is?
- (b) Conditional on at least one child’s being a girl, what’s the probability that they both are?

Problem 3.2. ♪ Flip and Flop are playing a game. They have a fair coin that they are going to keep flipping until one of two things happens: either the coin comes up heads twice in a row, or it comes up tails followed by heads. The first time one of these things happens, the game ends—if it ended with HH, Flip wins; if it ended with TH, Flop wins.

- (a) What’s the probability that Flip wins after the first two tosses of the coin? What’s the probability that Flop wins after the first two tosses of the coin?

- (b) Flip and Flop play their game until it ends (at which point one of them wins). What's the probability that Flop is the winner?²⁵

Problem 3.3. 🐞 One might think that real humans only ever assign credences that are rational numbers—and perhaps only rational numbers involving relatively small whole-number numerators and denominators. But we can construct simple conditions that *require* an irrational-valued credence distribution. For example, consider the scenario below.

You have a biased coin that you are going to flip twice in a row. Suppose your credence distribution satisfies all of the following conditions:

- (i) You are equally confident that the first flip will come up heads and that the second flip will come up heads.
- (ii) You treat the outcomes of the two flips as probabilistically independent.
- (iii) Given what you know about the bias, your confidence that the two flips will *both* come up heads equals your confidence in all of the other outcomes put together.

Assuming your credence distribution satisfies the three conditions above, how confident are you that the first flip will come up heads?²⁶

Problem 3.4. 🐞 Prove that credences conditional on a particular proposition form a probability distribution. That is, prove that for any proposition R in \mathcal{L} such that $\text{cr}(R) > 0$, the following three conditions hold:

- (a) For any proposition P in \mathcal{L} , $\text{cr}(P | R) \geq 0$.
- (b) For any tautology T in \mathcal{L} , $\text{cr}(T | R) = 1$.
- (c) For any mutually exclusive propositions P and Q in \mathcal{L} ,
 $\text{cr}(P \vee Q | R) = \text{cr}(P | R) + \text{cr}(Q | R)$.

Problem 3.5. 🐞 Pink gumballs always make my sister sick. (They remind her of Pepto Bismol.) Blue gumballs make her sick half of the time (they just look unnatural), while white gumballs make her sick only one-tenth of the time. Yesterday, my sister bought a single gumball randomly selected from a machine that's one-third pink gumballs, one-third blue, and one-third white. Applying the version of Bayes's Theorem in Equation (3.12), how confident should I be that my sister's gumball was pink conditional on the supposition that it made her sick?

Problem 3.6. 🐞

- (a) Prove Bayes's Theorem from the probability axioms and Ratio Formula. (Hint: Start by using the Ratio Formula to write down expressions involving $\text{cr}(H \& E)$ and $\text{cr}(E \& H)$.)

- (b) Exactly which unconditional credences must we assume to be positive in order for your proof to go through?
- (c) Where exactly does your proof rely on the probability axioms (and not just the Ratio Formula)?

Problem 3.7. ♪ Once more, consider the probabilistic credence distribution specified by this probability table (from Exercise 2.9):

P	Q	R	cr
T	T	T	0.1
T	T	F	0.2
T	F	T	0
T	F	F	0.3
F	T	T	0.1
F	T	F	0.2
F	F	T	0
F	F	F	0.1


Answer the following questions about this distribution:

- (a) What is $\text{cr}(P | Q)$?
- (b) Relative to this distribution, is Q positively relevant to P , negatively relevant to P , or probabilistically independent of P ?
- (c) What is $\text{cr}(P | R)$?
- (d) What is $\text{cr}(P | Q \& R)$?
- (e) Conditional on R , is Q positively relevant to P , negatively relevant to P , or probabilistically independent of P ?
- (f) Does R screen off P from Q ? Explain why or why not.

Problem 3.8. ♪ Prove that all the alternative statements of probabilistic independence in Equations (3.15) through (3.18) follow from our original independence definition. That is, prove that each equation (3.15) through (3.18) follows from Equation (3.14), the probability axioms, and the Ratio Formula. (Hint: Once you prove that a particular equation follows from Equation (3.14), you may use it in subsequent proofs.)

Problem 3.9. ♪ Show that probabilistic independence is not transitive. That is, provide a single probability distribution on which all of the following are true: X is independent of Y , and Y is independent of Z , but X is not independent


of Z . Show that your distribution satisfies all three conditions. (For an added chili pepper of difficulty, have your distribution assign every state-description a nonzero unconditional credence.)


Problem 3.10.  In the text we discussed what makes a *pair* of propositions probabilistically independent. If we have a larger collection of propositions, what does it take to make them all independent of each other? You might think all that's necessary is *pairwise independence*—for each pair within the set of propositions to be independent. But pairwise independence doesn't guarantee that each proposition will be independent of *combinations* of the others.

To demonstrate this fact, describe a real-world example (spelling out the propositions represented by X , Y , and Z) in which it would be rational for an agent to assign credences meeting all four of the following conditions:

- (i) $\text{cr}(X | Y) = \text{cr}(X)$
- (ii) $\text{cr}(X | Z) = \text{cr}(X)$
- (iii) $\text{cr}(Y | Z) = \text{cr}(Y)$
- (iv) $\text{cr}(X | Y \& Z) \neq \text{cr}(X)$

Show that your example satisfies all four conditions.



Problem 3.11.  Using the program PrSAT referenced in the Further Readings for Chapter 2, find a probability distribution satisfying all the conditions in Exercise 3.10, plus the following *additional* condition: Every state-description expressible in terms of X , Y , and Z must have a nonzero unconditional cr-value.


Problem 3.12. 

- (a) The 2016–2017 NBA season has just ended, and you're standing on a basketball court. Suddenly aliens appear, point to a spot on the court, and say that unless someone makes the next shot attempted from that spot, they will end your life. You're highly interested in self-preservation but terrible at basketball; luckily James Harden and DeMar DeRozan are standing right there. DeRozan says, "I had a better overall shooting percentage than Harden this year, so I should attempt the shot." Given the statistics on page 70, explain why DeRozan's argument is unconvincing.
- (b) Suppose the aliens pointed to a spot that would yield a three-point attempt. You're about to hand the ball to Harden, when DeRozan says,

“I know that’s a three-pointer, and Harden shot better from three-point range in general than I did this year. But from *that particular spot*, I had a better percentage than him.” Is DeRozan’s claim consistent with the statistics on page 70? (That is, could DeRozan’s claim be true while those statistics are also accurate?) Explain why or why not.


Problem 3.13. After laying down probabilistic conditions for a causal fork, Reichenbach demonstrated that a causal fork induces correlation. Consider the following four conditions:

- (i) $\text{cr}(A | C) > \text{cr}(A | \sim C)$
- (ii) $\text{cr}(B | C) > \text{cr}(B | \sim C)$
- (iii) $\text{cr}(A \& B | C) = \text{cr}(A | C) \cdot \text{cr}(B | C)$
- (iv) $\text{cr}(A \& B | \sim C) = \text{cr}(A | \sim C) \cdot \text{cr}(B | \sim C)$
- (a)  Assuming C is the common cause of A and B , explain what each of the four conditions means in terms of relevance, independence, conditional relevance, or conditional independence.
- (b)  Prove that if all four conditions hold, then $\text{cr}(A \& B) > \text{cr}(A) \cdot \text{cr}(B)$.

Problem 3.14.  In Section 3.3 I pointed out that the following statement (labeled Equation (3.43) there) does not hold for every constant k and propositions A , B , and C :

$$\text{“cr}(C | A) = k\text{” and “cr}(C | B) = k\text{” entail “cr}(C | A \vee B) = k\text{”}$$

- (a) Describe a real-world example (involving dice, or cards, or something more interesting) in which it’s rational for an agent to assign $\text{cr}(C | A) = k$ and $\text{cr}(C | B) = k$ but $\text{cr}(C | A \vee B) \neq k$. Show that your example meets this description.
- (b) Prove that if A and B are mutually exclusive, then whenever $\text{cr}(C | A) = k$ and $\text{cr}(C | B) = k$ it’s also the case that $\text{cr}(C | A \vee B) = k$.

Problem 3.15.  Fact: For any propositions P and Q , if $\text{cr}(Q) > 0$ then $\text{cr}(Q \supset P) \geq \text{cr}(P | Q)$.

- (a) Starting from a language \mathcal{L} with atomic propositions P and Q , build a probability table on its state-descriptions and use that table to prove the fact above.
- (b) Show that Equation (3.49) in Section 3.3 entails that either $\text{cr}(Q) = 1$ or $\text{cr}(Q \supset P) = 1$.

3.5 Further reading

INTRODUCTIONS AND OVERVIEWS

Alan Hájek (2011a). Conditional Probability. In: *Philosophy of Statistics*. Ed. by Prasanta S. Bandyopadhyay and Malcolm R. Forster. Vol. 7. Handbook of the Philosophy of Science. Amsterdam: Elsevier, pp. 99–136

Describes the Ratio Formula and its motivations. Then works through a number of philosophical applications of conditional probability, and a number of objections to the Ratio Formula. Also discusses conditional-probability-first formalizations (as described in note 3 of this chapter).

Todd A. Stephenson (2000). *An Introduction to Bayesian Network Theory and Usage*. Tech. rep. 03. IDIAP

Section 1 provides a nice, concise overview of what a Bayes Net is and how it interacts with conditional probabilities. (Note that the author uses A, B to express the *conjunction* of A and B .) Things get fairly technical after that as he covers algorithms for creating and using Bayes Nets. Sections 6 and 7, though, contain real-life examples of Bayes Nets for speech recognition, Microsoft Windows troubleshooting, and medical diagnosis.

Christopher R. Hitchcock (2021). Probabilistic Causation. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Spring 2021

While this entry is primarily about *analyses* of the concept of causation using probability theory, along the way Hitchcock includes impressive coverage of the Principle of the Common Cause, Simpson's Paradox, causal modeling with Bayes Nets, and related material.

CLASSIC TEXTS

Hans Reichenbach (1956). The Principle of Common Cause. In: *The Direction of Time*. Berkeley: University of California Press, pp. 157–60

Text in which Reichenbach introduces his account of common causes in terms of screening off. (Note that Reichenbach uses a period to express conjunction, and a comma rather than a vertical bar for conditional probabilities—what we would write as $\text{cr}(A | B)$ he writes as $P(B, A)$.)

David Lewis (1976). Probabilities of Conditionals and Conditional Probabilities. *The Philosophical Review* 85, pp. 297–315

Article in which Lewis presents his triviality argument concerning probabilities of conditionals.

EXTENDED DISCUSSION

Bas C. van Fraassen (1982). Rational Belief and the Common Cause Principle. In: *What? Where? When? Why?* Ed. by Robert McLaughlin. Dordrecht: Reidel, pp. 193–209

Frank Arntzenius (1993). The Common Cause Principle. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 2, pp. 227–37

Discuss the meaning and significance of Reichenbach's Principle of the Common Cause, then present possible counterexamples (including counterexamples from quantum mechanics).

Alan Hájek (2011b). Triviality Pursuit. *Topoi* 30, pp. 3–15

Explains the plausibility and significance of the claim that probabilities of conditionals are conditional probabilities. Then canvasses a variety of Lewis-style triviality arguments against that claim.

Notes

1. Here's a good way to double-check that $6 \ \& \ E$ is equivalent to 6 : Remember that equivalence is mutual entailment. Clearly $6 \ \& \ E$ entails 6 . Going in the other direction, 6 entails 6 , but 6 also entails E . So 6 entails $6 \ \& \ E$. When evaluating conditional credences using the Ratio Formula, we'll often find ourselves simplifying a conjunction down to just one or two of its conjuncts. For this to work, the conjunct that remains has to entail each of the conjuncts that was removed.
2. When I refer to an agent's "credence distribution" going forward, I will often be referring to both her unconditional and conditional credences. Strictly speaking this extends our definition of a "distribution", but since conditional credences supervene on unconditional for rational agents, not much damage will be done.
3. Some authors take advantage of Equation (3.7) to formalize probability theory in exactly the opposite order from the way I've been proceeding. They begin by introducing conditional credences and subject them to a number of constraints somewhat like

Kolmogorov's axioms. The desired rules for *unconditional* credences are then obtained by introducing the single constraint that for all P in \mathcal{L} , $\text{cr}(P) = \text{cr}(P|T)$. For more on this approach, its advocates, and its motivations, see Section 5.4.

4. Bayes never published the theorem; Richard Price found it in Bayes's notes and published it after Bayes's death in 1761. Pierre-Simon Laplace independently rediscovered the theorem later on and was responsible for much of its early popularization.
5. In everyday English "likely" is a synonym for "probable". Yet R.A. Fisher introduced the technical term "likelihood" to represent a particular *kind* of probability—the probability of some evidence given a hypothesis. This somewhat peculiar terminology has stuck.
6. Quoted in Galavotti (2005, p. 51).
7. For instance, we have to assume that the base rates of the alleles are equal in the population, none of the relevant phenotypes is fitter than any of the others, and the blue-finned fish don't assortatively mate with other blue-finned fish. (Thanks to Hayley Clatterback for discussion.)
8. Throughout this section and Section 3.2.1, I will assume that any proposition appearing in the condition of a conditional *cr*-expression has a nonzero *cr*-value. Defining probabilistic independence for propositions with probability 0 can get complicated. (See e.g. Fitelson and Hájek (2014))
9. One will sometimes see "screening off" defined without Equation (3.28) or its analogue. (That is, some authors define screening off in terms of R 's making the correlation between P and Q disappear, without worrying whether $\sim R$ has the same effect.) Equation (3.28) makes an important difference to our definition: in the Baxter example I does not screen off H from C according to our definition because when $\sim I$ is supposed, C becomes very relevant to H .

I have included Equation (3.28) in our definition because it connects our approach to the more general notion of screening off used in the statistics community. In statistics one often works with continuous random variables, and the idea is that random variable Z screens off X from Y if X and Y become independent conditional on each possible value of Z . Understanding proposition R as a dichotomous random variable (Chapter 2, note 7) converts this general definition of screening off into the particular definition I've given in the text.

Many authors also leave Equation (3.26) (or its analogue) implicit in their definitions of "screening off". But since examples of screening off always involve unconditional correlations that disappear under conditioning, I've made this feature explicit in my definition.

10. Not to be confused with the Rambler's Fallacy: I've said so many false things in a row, the next one must be true!
11. Twenty flips of a fair coin provide a good example of what statisticians call **IID trials**. "IID" stands for "independent, identically distributed." The flips are "independent" because each is probabilistically independent of all the others; information about the outcomes of other coin flips doesn't change the probability that a particular flip will come up heads. The flips are "identically distributed" because each flip has the same probability of producing heads (as contrasted with a case in which some of the flips are of a fair coin while others are flips of a biased coin).

12. This paradoxical phenomenon is named after E.H. Simpson because of a number of striking examples he gave in his (1951). Yet the phenomenon had been known to statisticians as early as Pearson, Lee, and Bramley-Moore (1899) and Yule (1903).
13. I learned about the Harden/DeRoza example from Reuben Stern, who in turn learned about it from a reddit post by a user named Jerome Williams. (I copied the specific data for the two players from stats.nba.com.) The UC Berkeley example was brought to the attention of philosophers by Cartwright (1979).
14. Notice that these findings only address one potential form of bias that might have been present in Berkeley's graduate admissions. For instance, they're consistent with the possibility that women were being actively discouraged from applying to the less selective departments.
15. I'm playing a bit fast and loose with the objects of discussion here. Throughout this chapter we're considering correlations in an agent's credence distribution. Reichenbach was concerned not with probabilistic correlations in an agent's credences but instead with correlations in objective frequencies or chance distributions (about which more in Chapter 5). But presumably if the Principle of the Common Cause holds for objective probability distributions, that provides an agent who views particular propositions as empirically correlated with some reason to suppose that the events described in those propositions either stand as cause to effect or share a common cause.
16. You might worry that Figure 3.3 presents a counterexample to Reichenbach's Principle of the Common Cause, because *G* and *S* are unconditionally correlated yet *G* doesn't cause *S* and they have no common cause. It's important to the principle that the causal relations need not be *direct*; for Reichenbach's purposes *G* counts as a cause of *S* even though it's not the immediate cause of *S*.
17. Just to indicate a few more complexities that may arise: While our discussion in the text concerns "direct" common causes, one can have an "indirect" common cause that doesn't screen off its effects from each other. For example, if we imagine merging Figures 3.2 and 3.3 to show how the subject's parents' genes are a common cause of both smoking and drinking by way of her addictive personality, it is possible to arrange the numbers so that her parents' genetics don't screen off smoker from drinker. Even more complications arise if some causal arrows do end-arounds past others—what if in addition to the causal structure just described, the parents' genetics tend to make *them* smokers, which in turn directly influences the subject's smoking behavior?
18. Here I assume that a rational agent will entertain an indicative conditional only if she takes its antecedent to be possible. For arguments in favor of this position, and citations to the relevant literature, see Moss (2018, Sect. 4.3) and Titelbaum (2013a, Sect. 5.3.2). The analogous assumption for conditional credences is that an agent assigns a conditional credence only when its condition is true in at least one of her doxastically possible worlds.
19. One *could* study a kind of attitude different from the conditional credences considered in this book—something like a subjunctive degree of belief. Joyce (1999) does exactly that, but is careful to distinguish his analysandum from standard conditional degrees of belief. (For instance, the arguments for the Ratio Formula given earlier in this chapter do not extend to Joycean subjunctive credences.) Schwarz (2018) then evaluates triviality arguments for subjunctive conditional credences much like the triviality arguments for indicatives I will go on to consider in this section.

20. I realize some of the “p”s and “q”s in this quote are flipped around from what one might expect, but that’s how it’s printed in my copy of Ramsey (1931). Context makes clear that the ordering is *not* why Ramsey rejects the proposed equivalents to “the degree of belief in p given q”; he’d still reject them were the order inverted.
21. A variety of recent positions in linguistics and the philosophy of language suggest that indicative conditionals with modal expressions in their consequents do not obey classical logical rules. Yalcin (2012), among others, classes probability locutions with these modals and so argues that, *inter alia*, indicative conditionals with probabilistic consequents do not keep *modus tollens* truth-preserving. (His argument could easily be extended to disjunctive syllogism as well.) Yet the alternative positive theory of indicative conditionals Yalcin offers does not analyze conditional credences in terms of conditionals either, so even if he’s correct, we would still need an independent understanding of what conditional credences are. (Thanks to Fabrizio Cariani for discussion of these points.)
22. A careful reader will note that the proof given fails when $cr(P|Q)$ takes an extreme value. If $cr(P|Q) = 0$, the condition in $cr(P|Q \& P)$ will have unconditional credence 0, while if $cr(P|Q) = 1$, the condition in $cr(P|Q \& \sim P)$ will have credence 0. So strictly speaking the trivality result is that (3.48) can hold only when the agent takes *P* and *Q* to be independent, or dependent in the strongest possible fashion. This is no more plausible than the less careful version I’ve given in the text. (Thanks to Glauber de Bona for being the careful reader who caught this for me!)
23. Fitelson (2015) proves a trivality result like Lewis’s using probability tables (instead of proceeding axiomatically). Moreover, he traces the trivality specifically to the combination of (3.48) with the assumption that the conditional \rightarrow satisfies what’s known as the “import-export” condition.
24. Interestingly, this idea is often traced back to a suggestion in Ramsey, known as “Ramsey’s test” (Ramsey 1929/1990, p. 155n).
25. Thanks to Irving Lubliner for inspiring this problem.
26. This problem was inspired by a problem of Branden Fitelson’s. Thanks to Catrin Campbell-Moore for devising this particularly elegant set of conditions.

Bibliography of Volumes 1 & 2

- Achinstein, Peter (1963). Variety and Analogy in Confirmation Theory. *Philosophy of Science* 3, pp. 207–21.
- Adams, Ernest (1962). On Rational Betting Systems. *Archiv für mathematische Logik und Grundlagenforschung* 6, pp. 7–29.
- Adams, Ernest (1965). The Logic of Conditionals. *Inquiry* 8, pp. 166–97.
- Alchourrón, Carlos E., Peter Gärdenfors, and David Makinson (1985). On the Logic of Theory Change: Partial Meet Contraction and Revision Functions. *The Journal of Symbolic Logic* 50, pp. 510–30.
- Allais, Maurice (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école Américaine. *Econometrica* 21, pp. 503–46.
- Amrhein, Valentin, Sander Greenland, and Blake McShane (2019). Scientists Rise up against Statistical Significance. *Nature* 567, pp. 305–7.
- Armendt, Brad (1980). Is There a Dutch Book Argument for Probability Kinematics? *Philosophy of Science* 47, pp. 583–8.
- Armendt, Brad (1992). Dutch Strategies for Diachronic Rules: When Believers See the Sure Loss Coming. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1, pp. 217–29.
- Arntzenius, Frank (1993). The Common Cause Principle. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 2, pp. 227–37.
- Arntzenius, Frank (2003). Some Problems for Conditionalization and Reflection. *The Journal of Philosophy* 100, pp. 356–70.
- Arrow, Kenneth J. (1951). *Social Choice and Individual Values*. New York: John Wiley and Sons.
- Bandyopadhyay, Prasanta S. and Malcolm R. Forster (2011). *Philosophy of Statistics*. Vol. 7. Handbook of the Philosophy of Science. Amsterdam: Elsevier.
- Barnett, Jeffrey A. (1981). Computational Methods for a Mathematical Theory of Evidence. In: *Proceedings of the 7th International Joint Conference on AI*, pp. 868–75.
- Bartha, Paul and Christopher R. Hitchcock (1999). No One Knows the Date or the Hour: An Unorthodox Application of Rev. Bayes's Theorem. *Philosophy of Science* 66, S339–53.
- Berger, James O. and Donald A. Berry (1988). Statistical Analysis and the Illusion of Objectivity. *American Scientist* 76, pp. 159–65.
- Bergmann, Merrie, James Moor, and Jack Nelson (2013). *The Logic Book*. 6th edition. New York: McGraw Hill.
- Berker, Selim (2013). Epistemic Teleology and the Separateness of Propositions. *Philosophical Review* 122, pp. 337–93.
- Bernoulli, Daniel (1738/1954). Exposition of a New Theory on the Measurement of Risk. *Econometrica* 22, pp. 23–36.
- Bernoulli, Jacob (1713). *Ars Conjectandi*. Basiliae.
- Bertrand, Joseph (1888/1972). *Calcul des probabilités*. 2nd edition. New York: Chelsea Publishing Company.

- Bickel, P.J., E.A. Hammel, and J.W. O'Connell (1975). Sex Bias in Graduate Admissions: Data from Berkeley. *Science* 187, pp. 398–404.
- Birnbaum, Allan (1962). On the Foundations of Statistical Inference. *Journal of the American Statistical Association* 57, pp. 269–306.
- Bolzano, Bernard (1837/1973). *Wissenschaftslehre*. Translated by Jan Berg under the title *Theory of Science*. Dordrecht: Reidel.
- Bovens, Luc and Stephan Hartmann (2003). *Bayesian Epistemology*. Oxford: Oxford University Press.
- Bradley, Darren (2010). Conditionalization and Belief *De Se*. *Dialectica* 64, pp. 247–50.
- Bradley, Darren (2011). Self-location Is No Problem for Conditionalization. *Synthese* 182, pp. 393–411.
- Bradley, Darren (2015). *A Critical Introduction to Formal Epistemology*. London: Bloomsbury.
- Bratman, Michael E. (1987). *Intention, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press.
- Brier, George (1950). Verification of Forecasts Expressed in Terms of Probability. *Monthly Weather Review* 78, pp. 1–3.
- Briggs, R.A. (2010). Putting a Value on Beauty. In: *Oxford Studies in Epistemology*. Ed. by Tamar Szabó Gendler and John Hawthorne. Vol. 3. Oxford University Press, pp. 3–34.
- Briggs, R.A. and Richard Pettigrew (2020). An Accuracy-Dominance Argument for Conditionalization. *Noûs* 54, pp. 162–81.
- Broome, John (1999). Normative Requirements. *Ratio* 12, pp. 398–419.
- Buchak, Lara (2013). *Risk and Rationality*. Oxford: Oxford University Press.
- Campbell-Moore, Catrin and Benjamin A. Levinstein (2021). Strict Propriety Is Weak. *Analysis* 81.1, pp. 8–13.
- Capotorti, Andrea and Barbara Vantaggi (2000). Axiomatic Characterization of Partial Ordinal Relations. *International Journal of Approximate Reasoning* 24, pp. 207–19.
- Carnap, Rudolf (1945). On Inductive Logic. *Philosophy of Science* 12, pp. 72–97.
- Carnap, Rudolf (1947). On the Application of Inductive Logic. *Philosophy and Phenomenological Research* 8, pp. 133–48.
- Carnap, Rudolf (1950). *Logical Foundations of Probability*. Chicago: University of Chicago Press.
- Carnap, Rudolf (1955/1989). Statistical and Inductive Probability. In: *Readings in the Philosophy of Science*. Ed. by Baruch A. Brody and Richard E. Grandy. 2nd edition. Hoboken: Prentice-Hall, pp. 279–87.
- Carnap, Rudolf (1962a). *Logical Foundations of Probability*. 2nd edition. Chicago: University of Chicago Press.
- Carnap, Rudolf (1962b). The Aim of Inductive Logic. In: *Logic, Methodology, and the Philosophy of Science*. Ed. by P. Suppes, E. Nagel, and A. Tarski. Stanford University: Stanford University Press, pp. 303–18.
- Carnap, Rudolf (1971). A Basic System of Inductive Logic, Part 1. In: *Studies in Inductive Logic and Probability*. Ed. by Rudolf Carnap and Richard C. Jeffrey. Vol. I. Berkeley: University of California Press, pp. 33–166.
- Carnap, Rudolf (1980). A Basic System of Inductive Logic, Part 2. In: *Studies in Inductive Logic and Probability*. Ed. by Richard C. Jeffrey. Vol. II. Berkeley: University of California Press, pp. 7–156.
- Carr, Jennifer (2017). Epistemic Utility Theory and the Aim of Belief. *Philosophy and Phenomenological Research* 95, pp. 511–34.

- Cartwright, Nancy (1979). Causal Laws and Effective Strategies. *Noûs* 13, pp. 419–37.
- Carver, Ronald P. (1978). The Case against Statistical Significance Testing. *Harvard Educational Review* 48, pp. 378–99.
- Chandler, Jake (2013). Contrastive Confirmation: Some Competing Accounts. *Synthese* 190, pp. 129–38.
- Cherniak, Christopher (1986). *Minimal Rationality*. Cambridge, MA: The MIT Press.
- Chihara, C. (1981). Quine and the Confirmational Paradoxes. In: *Midwest Studies in Philosophy 6: Foundations of Analytic Philosophy*. Ed. by P. French, H. Wettstein, and T. Uehling. Minneapolis: University of Minnesota Press, pp. 425–52.
- Christensen, David (1991). Clever Bookies and Coherent Beliefs. *The Philosophical Review* 100, pp. 229–47.
- Christensen, David (1999). Measuring Confirmation. *The Journal of Philosophy* 96, pp. 437–61.
- Christensen, David (2001). Preference-Based Arguments for Probabilism. *Philosophy of Science* 68, pp. 356–76.
- Christensen, David (2004). *Putting Logic in its Place*. Oxford: Oxford University Press.
- Colyvan, Mark (2004). The Philosophical Significance of Cox's Theorem. *International Journal of Approximate Reasoning* 37, pp. 71–85.
- Cox, Richard T. (1946). Probability, Frequency and Reasonable Expectation. *American Journal of Physics* 14, pp. 1–13.
- Cox, Richard T. (1961). *The Algebra of Probable Inference*. Baltimore, MD: The Johns Hopkins Press.
- Crupi, Vincenzo, Branden Fitelson, and Katya Tentori (2008). Probability, Confirmation, and the Conjunction Fallacy. *Thinking & Reasoning* 14, pp. 182–99.
- Crupi, Vincenzo, Katya Tentori, and Michel Gonzalez (2007). On Bayesian Measures of Evidential Support: Theoretical and Empirical Issues. *Philosophy of Science* 74, pp. 229–52.
- Davidson, Donald (1984). *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.
- Davidson, Donald, J.C.C. McKinsey, and Patrick Suppes (1955). Outlines of a Formal Theory of Value, I. *Philosophy of Science* 22, pp. 14–60.
- de Finetti, Bruno (1931/1989). Probabilism: A Critical Essay on the Theory of Probability and the Value of Science. *Erkenntnis* 31, pp. 169–223. Translation of B. de Finetti, *Probabilismo*, *Logos* 14, pp. 163–219.
- de Finetti, Bruno (1937/1964). Foresight: Its Logical Laws, its Subjective Sources. In: *Studies in Subjective Probability*. Ed. by Henry E. Kyburg Jr and H.E. Smokler. New York: Wiley, pp. 94–158. Originally published as “La prévision; ses lois logiques, ses sources subjectives” in *Annales de l'Institut Henri Poincaré* 7, pp. 1–68.
- de Finetti, Bruno (1949/1951). La ‘Logica del Plausibile’ Secondo la Concezione di Polya. *Atti della XLII Riunione, Società Italiana per il Progresso delle Scienze*. Presented in 1949, published in 1951, pp. 227–36.
- de Finetti, Bruno (1974). *Theory of Probability*. Vol. 1. New York: Wiley.
- de Finetti, Bruno (1995). *Filosofia della probabilità*. Ed. by Alberto Mura. Milan: Il Saggiatore.
- DeGroot, Morris H. (1970). *Optimal Statistical Decisions*. Hoboken, New Jersey: Wiley.
- Dempster, Arthur P. (1966). New Methods for Reasoning towards Posterior Distributions Based on Sample Data. *Annals of Mathematical Statistics* 37, pp. 355–74.
- Dickson, Michael and Davis Baird (2011). Significance Testing. In: *Philosophy of Statistics*. Ed. by Prasanta S. Bandyopadhyay and Malcolm R. Forster. Vol. 7. Handbook of the Philosophy of Science. Amsterdam: Elsevier.

- Domotor, Zoltan (1969). *Probabilistic Relational Structures and their Applications*. Technical Report 144. Stanford University: Institute for Mathematical Studies in the Social Sciences.
- Earman, John (1992). *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. Cambridge, MA: The MIT Press.
- Easwaran, Kenny (2013). Expected Accuracy Supports Conditionalization—and Conglomerability and Reflection. *Philosophy of Science* 80, pp. 119–42.
- Easwaran, Kenny (2014a). Decision Theory without Representation Theorems. *Philosophers' Imprint* 14, pp. 1–30.
- Easwaran, Kenny (2014b). Regularity and Hyperreal Credences. *Philosophical Review* 123, pp. 1–41.
- Eddington, A. (1939). *The Philosophy of Physical Science*. Cambridge: Cambridge University Press.
- Edwards, A.W.F. (1972). *Likelihood: An Account of the Statistical Concept of Likelihood and its Application to Scientific Inference*. Cambridge: Cambridge University Press.
- Eells, Ellery (1982). *Rational Decision and Causality*. Cambridge Studies in Philosophy. Cambridge: Cambridge University Press.
- Eells, Ellery (1985). Problems of Old Evidence. *Pacific Philosophical Quarterly* 66, pp. 283–302.
- Eells, Ellery and Branden Fitelson (2000). Measuring Confirmation and Evidence. *Journal of Philosophy* 97, pp. 663–72.
- Eells, Ellery and Branden Fitelson (2002). Symmetries and Asymmetries in Evidential Support. *Philosophical Studies* 107, pp. 129–42.
- Efron, B. (1986). Why Isn't Everyone a Bayesian? *The American Statistician* 40, pp. 1–5.
- Egan, Andy (2007). Some Counterexamples to Causal Decision Theory. *Philosophical Review* 116, pp. 93–114.
- Egan, Andy and Michael G. Titelbaum (2022). Self-locating Belief. Forthcoming in *The Stanford Encyclopedia of Philosophy*.
- Elga, Adam (2000). Self-locating Belief and the Sleeping Beauty Problem. *Analysis* 60, pp. 143–7.
- Elga, Adam (2007). Reflection and Disagreement. *Noûs* 41, pp. 478–502.
- Elga, Adam (2010). Subjective Probabilities Should Be Sharp. *Philosophers' Imprint* 10, pp. 1–11.
- Ellenberg, Jordan (2014). *How Not to Be Wrong: The Power of Mathematical Thinking*. New York: Penguin Press.
- Ellis, Robert Leslie (1849). On the Foundations of the Theory of Probabilities. *Transactions of the Cambridge Philosophical Society* VIII, pp. 1–6.
- Ellsberg, Daniel (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics* 75, pp. 643–69.
- Feldman, Richard (2007). Reasonable Religious Disagreements. In: *Philosophers without Gods: Meditations on Atheism and the Secular Life*. Ed. by Louise M. Antony. Oxford: Oxford University Press, pp. 194–214.
- Feller, William (1968). *An Introduction to Probability Theory and its Applications*. 3rd edition. New York: Wiley.
- Fermat, Pierre and Blaise Pascal (1654/1929). Fermat and Pascal on Probability. In: *A Source Book in Mathematics*. Ed. by D. Smith. Translated by Vera Sanford. New York: McGraw-Hill, pp. 546–65.
- Fine, Terrence L. (1973). *Theories of Probability: An Examination of Foundations*. New York, London: Academic Press.

- Fishburn, Peter C. (1981). Subjective Expected Utility: A Review of Normative Theories. *Theory and Decision* 13, pp. 129–99.
- Fishburn, Peter C. (1986). The Axioms of Subjective Probability. *Statistical Science* 1, pp. 335–45.
- Fisher, Ronald A. (1956). *Statistical Methods and Scientific Inference*. Edinburgh: Oliver and Boyd.
- Fitelson, Branden (2006). Logical Foundations of Evidential Support. *Philosophy of Science* 73, pp. 500–12.
- Fitelson, Branden (2007). Likelihoodism, Bayesianism, and Relational Confirmation. *Synthese* 156, pp. 473–89.
- Fitelson, Branden (2008). A Decision Procedure for Probability Calculus with Applications. *The Review of Symbolic Logic* 1, pp. 111–25.
- Fitelson, Branden (2012). Evidence of Evidence Is Not (Necessarily) Evidence. *Analysis* 72, pp. 85–8.
- Fitelson, Branden (2015). The Strongest Possible Lewisian Triviality Result. *Thought* 4, pp. 69–74.
- Fitelson, Branden and Alan Hájek (2014). Declarations of Independence. *Synthese* 194, pp. 3979–95.
- Fitelson, Branden and James Hawthorne (2010a). How Bayesian Confirmation Theory Handles the Paradox of the Ravens. *The Place of Probability in Science*. Ed. by Ellery Eells and J. Fetzer. *Boston Studies in the Philosophy of Science* 284, pp. 247–75.
- Fitelson, Branden and James Hawthorne (2010b). The Wason Task(s) and the Paradox of Confirmation. *Philosophical Perspectives* 24, pp. 207–41.
- Fitelson, Branden and David McCarthy (ms). Accuracy and Comparative Likelihood. Unpublished manuscript.
- Fitelson, Branden and Daniel Osherson (2015). Remarks on “Random Sequences”. *Australasian Journal of Logic* 12, pp. 11–16.
- Foley, Richard (1987). *The Theory of Epistemic Rationality*. Cambridge, MA: Harvard University Press.
- Foley, Richard (1993). *Working without a Net*. Oxford: Oxford University Press.
- Foley, Richard (2009). Beliefs, Degrees of Belief, and the Lockean Thesis. In: *Degrees of Belief*. Ed. by Franz Huber and Christoph Schmidt-Petri. Vol. 342. Synthese Library. Springer, pp. 37–48.
- Frigerio, Roberta et al. (2005). Education and Occupations Preceding Parkinson’s Disease. *Neurology* 65, pp. 1575–83.
- Gaifman, H. and M. Snir (1982). Probabilities over Rich Languages. *Journal of Symbolic Logic* 47, pp. 495–548.
- Gaifman, Haim (2004). Reasoning with Limited Resources and Assigning Probabilities to Arithmetical Statements. *Synthese* 140, pp. 97–119.
- Galavotti, Maria Carla (2005). *Philosophical Introduction to Probability*. CSLI Lecture Notes 167. Stanford, CA: CSLI Publications.
- Gallo, Valentina et al. (2018). Exploring Causality of the Association between Smoking and Parkinson’s Disease. *International Journal of Epidemiology* 48, pp. 912–25.
- Gandenberger, Greg (2014). *Titelbaum’s Counterexample to the Law of Likelihood*. Blog post available at URL: <https://gandenberger.org/2014/05/26/titelbaum-counterexample/>.
- Gandenberger, Greg (2015). A New Proof of the Likelihood Principle. *British Journal for the Philosophy of Science* 66, pp. 475–503.
- Gandenberger, Greg (ms). New Responses to Three Purported Counterexamples to the Likelihood Principle. Unpublished manuscript.

- Garber, Daniel (1983). Old Evidence and Logical Omniscience in Bayesian Confirmation Theory. In: *Testing Scientific Theories*. Ed. by John Earman. Vol. 10. Minnesota Studies in the Philosophy of Science. Minneapolis: University of Minnesota Press, pp. 99–132.
- Gibbard, A. and W. Harper (1978/1981). Counterfactuals and Two Kinds of Expected Utility. In: *IFS: Conditionals, Belief, Decision, Chance, and Time*. Ed. by W. Harper, Robert C. Stalnaker, and G. Pearce. Dordrecht: Reidel, pp. 153–90.
- Gillies, Donald (2000). Varieties of Propensity. *British Journal for the Philosophy of Science* 51, pp. 807–35.
- Glass, David H. and Mark McCartney (2015). A New Argument for the Likelihood Ratio Measure of Confirmation. *Acta Analytica* 30, pp. 59–65.
- Glymour, Clark (1980). *Theory and Evidence*. Princeton, NJ: Princeton University Press.
- Good, I.J. (1952). Rational Decisions. *Journal of the Royal Statistical Society, Series B* 14, pp. 107–14.
- Good, I.J. (1962). Subjective Probability as the Measure of a Non-measurable Set. In: *Logic, Methodology, and the Philosophy of Science*. Ed. by P. Suppes, E. Nagel, and A. Tarski. Stanford, CA: Stanford University Press, pp. 319–29.
- Good, I.J. (1967). The White Shoe Is a Red Herring. *British Journal for the Philosophy of Science* 17, p. 322.
- Good, I.J. (1968). The White Shoe qua Herring Is Pink. *British Journal for the Philosophy of Science* 19, pp. 156–7.
- Good, I.J. (1971). Letter to the Editor. *The American Statistician* 25, pp. 62–3.
- Goodman, Nelson (1946). A Query on Confirmation. *The Journal of Philosophy* 43, pp. 383–5.
- Goodman, Nelson (1955). *Fact, Fiction, and Forecast*. Cambridge, MA: Harvard University Press.
- Greaves, Hilary (2013). Epistemic Decision Theory. *Mind* 122, pp. 915–52.
- Greaves, Hilary and David Wallace (2006). Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility. *Mind* 115, pp. 607–32.
- Hacking, Ian (1965). *The Logic of Statistical Inference*. Cambridge: Cambridge University Press.
- Hacking, Ian (1967). Slightly More Realistic Personal Probability. *Philosophy of Science* 34, pp. 311–25.
- Hacking, Ian (1971). The Leibniz-Carnap Program for Inductive Logic. *The Journal of Philosophy* 68, pp. 597–610.
- Hacking, Ian (2001). *An Introduction to Probability and Inductive Logic*. Cambridge: Cambridge University Press.
- Hájek, Alan (1996). ‘Mises Redux’—Redux: Fifteen Arguments against Finite Frequentism. *Erkenntnis* 45, pp. 209–27.
- Hájek, Alan (2003). What Conditional Probability Could Not Be. *Synthese* 137, pp. 273–323.
- Hájek, Alan (2009a). Arguments for—or against—Probabilism? In: *Degrees of Belief*. Ed. by Franz Huber and Christoph Schmidt-Petri. Vol. 342. Synthese Library. Springer, pp. 229–51.
- Hájek, Alan (2009b). Fifteen Arguments against Hypothetical Frequentism. *Erkenntnis* 70, pp. 211–35.
- Hájek, Alan (2011a). Conditional Probability. In: *Philosophy of Statistics*. Ed. by Prasanta S. Bandyopadhyay and Malcolm R. Forster. Vol. 7. Handbook of the Philosophy of Science. Amsterdam: Elsevier, pp. 99–136.
- Hájek, Alan (2011b). Triviality Pursuit. *Topoi* 30, pp. 3–15.

- Hájek, Alan (2019). Interpretations of Probability. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Fall 2019. URL: <http://plato.stanford.edu/archives/fall2019/entries/probability-interpret/>.
- Hájek, Alan and James M. Joyce (2008). Confirmation. In: *The Routledge Companion to Philosophy of Science*. Ed. by Stathis Psillos and Martin Curd. New York: Routledge, pp. 115–28.
- Hall, Ned (2004). Two Mistakes about Credence and Chance. *Australasian Journal of Philosophy* 82, pp. 93–111.
- Halpern, Joseph Y. (1999). Cox's Theorem Revisited. *Journal of Artificial Intelligence Research* 11, pp. 429–35.
- Halpern, Joseph Y. (2003). *Reasoning about Uncertainty*. Cambridge, MA: MIT Press.
- Halpern, Joseph Y. (2004). Sleeping Beauty Reconsidered: Conditioning and Reflection in Asynchronous Systems. *Proceedings of the Twentieth Conference on Uncertainty in AI*, pp. 226–34.
- Halpern, Joseph Y. (2005). Sleeping Beauty Reconsidered: Conditioning and Reflection in Asynchronous Systems. In: *Oxford Studies in Epistemology*. Ed. by Tamar Szabó Gendler and John Hawthorne. Vol. 1. Oxford: Oxford University Press, pp. 111–42.
- Harman, Gilbert (1986). *Change in View*. Boston: The MIT Press.
- Harrison-Trainor, Matthew, Wesley H. Holliday, and Thomas F. Icard III (2016). A Note on Cancellation Axioms for Comparative Probability. *Theory and Decision* 80, pp. 159–66.
- Hart, Casey and Michael G. Titelbaum (2015). Intuitive Dilation? *Thought* 4, pp. 252–62.
- Haverkamp, Nick and Moritz Schulz (2012). A Note on Comparative Probability. *Erkenntnis* 76, pp. 395–402.
- Hawthorne, James (2014). Inductive Logic. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2014. URL: <http://plato.stanford.edu/archives/win2014/entries/logic-inductive/>.
- Hawthorne, James (2016). A Logic of Comparative Support: Qualitative Conditional Probability Relations Representable by Popper Functions. In: *The Oxford Handbook of Probability and Philosophy*. Ed. by Alan Hájek and Christopher R. Hitchcock. Oxford: Oxford University Press, pp. 277–95.
- Hawthorne, James and Branden Fitelson (2004). Re-solving Irrelevant Conjunction with Probabilistic Independence. *Philosophy of Science* 71, pp. 505–14.
- Hedden, Brian (2013). Incoherence without Exploitability. *Noûs* 47, pp. 482–95.
- Hempel, Carl G. (1945a). Studies in the Logic of Confirmation (I). *Mind* 54, pp. 1–26.
- Hempel, Carl G. (1945b). Studies in the Logic of Confirmation (II). *Mind* 54, pp. 97–121.
- Hesse, Mary (1963). *Models and Analogies in Science*. London: Sheed & Ward.
- Heukelom, Floris (2015). A History of the Allais Paradox. *The British Journal for the History of Science* 48, pp. 147–69.
- Hintikka, Jaakko (1975). Impossible Possible Worlds Vindicated. *Journal of Philosophical Logic* 4, pp. 475–84.
- Hitchcock, Christopher R. (2004). Beauty and the Bets. *Synthese* 139, pp. 405–20.
- Hitchcock, Christopher R. (2021). Probabilistic Causation. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Spring 2021. URL: <https://plato.stanford.edu/archives/spr2021/entries/causation-probabilistic/>.
- Holton, Richard (2014). Intention as a Model for Belief. In: *Rational and Social Agency: The Philosophy of Michael Bratman*. Ed. by Manuel Vargas and Gideon Yaffe. Oxford: Oxford University Press, pp. 12–37.
- Hooker, C.A. (1968). Goodman, 'Grue' and Hempel. *Philosophy of Science* 35, pp. 232–47.

- Hosiasson-Lindenbaum, Janina (1940). On Confirmation. *Journal of Symbolic Logic* 5, pp. 133–48.
- Howson, Colin (1992). Dutch Book Arguments and Consistency. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 2, pp. 161–8.
- Howson, Colin (2014). Finite Additivity, Another Lottery Paradox and Conditionalisation. *Synthese* 191, pp. 989–1012.
- Howson, Colin and Peter Urbach (2006). *Scientific Reasoning: The Bayesian Approach*. 3rd edition. Chicago: Open Court.
- Huber, Franz (2016). Formal Representations of Belief. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Spring 2016. URL: <https://plato.stanford.edu/archives/spr2016/entries/formal-belief/>.
- Hume, David (1739–40/1978). *A Treatise of Human Nature*. Ed. by L.A. Selby-Bigge and Peter H. Niddich. 2nd edition. Oxford: Oxford University Press.
- Humphreys, Paul (1985). Why Propensities Cannot Be Probabilities. *Philosophical Review* 94, pp. 557–70.
- Icard III, Thomas F. (2016). Pragmatic Considerations on Comparative Probability. *Philosophy of Science* 83, pp. 348–70.
- Jackson, Elizabeth G. (2020). The Relationship between Belief and Credence. *Philosophy Compass* 15, pp. 1–13.
- Jaynes, E.T. (1957a). Information Theory and Statistical Mechanics I. *Physical Review* 106, pp. 62–30.
- Jaynes, E.T. (1957b). Information Theory and Statistical Mechanics II. *Physical Review* 108, pp. 171–90.
- Jaynes, E.T. (1968). Prior Probabilities. *IEEE Transactions on Systems Science and Cybernetics* SEC-4, pp. 227–41.
- Jeffrey, Richard C. (1965). *The Logic of Decision*. 1st edition. McGraw-Hill Series in Probability and Statistics. New York: McGraw-Hill.
- Jeffrey, Richard C. (1970). Dracula Meets Wolfman: Acceptance vs. Partial Belief. In: *Induction, Acceptance, and Rational Belief*. Ed. by M. Swain. Dordrecht: Reidel, pp. 157–85.
- Jeffrey, Richard C. (1983). *The Logic of Decision*. 2nd edition. Chicago: University of Chicago Press.
- Jeffrey, Richard C. (1987). Indefinite Probability Judgment: A Reply to Levi. *Philosophy of Science* 54, pp. 586–91.
- Jeffrey, Richard C. (1993). Causality and the Logic of Decision. *Philosophical Topics* 21, pp. 139–51.
- Jeffrey, Richard C. (2004). *Subjective Probability: The Real Thing*. Cambridge: Cambridge University Press.
- Johnson, W.E. (1932). Probability: The Deductive and Inductive Problems. *Mind* 41, pp. 409–23.
- Joyce, James M. (1998). A Nonpragmatic Vindication of Probabilism. *Philosophy of Science* 65, pp. 575–603.
- Joyce, James M. (1999). *The Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press.
- Joyce, James M. (2005). How Probabilities Reflect Evidence. *Philosophical Perspectives* 19, pp. 153–78.
- Joyce, James M. (2009). Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief. In: *Degrees of Belief*. Ed. by Franz Huber and Christoph Schmidt-Petri. Vol. 342. Synthese Library. Springer, pp. 263–97.

- Joyce, James M. (2010). A Defense of Imprecise Credences in Inference and Decision Making. *Philosophical Perspectives* 24, pp. 281–323.
- Kahneman, Daniel and Amos Tversky (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica* XLVII, pp. 263–91.
- Kaplan, Mark (1996). *Decision Theory as Philosophy*. Cambridge: Cambridge University Press.
- Kelley, Mikayla (ms). On Accuracy and Coherence with Infinite Opinion Sets. Unpublished manuscript.
- Kemeny, John G. (1955). Fair Bets and Inductive Probabilities. *The Journal of Symbolic Logic* 20, pp. 263–73.
- Kemeny, John G. and Paul Oppenheim (1952). Degree of Factual Support. *Philosophy of Science* 19, pp. 307–24.
- Keynes, John Maynard (1921). *Treatise on Probability*. London: Macmillan and Co., Limited.
- Keynes, John Maynard (1923). *A Tract on Monetary Reform*. Macmillan and Co., Limited.
- Keynes, John Maynard (1937). The General Theory of Employment. *The Quarterly Journal of Economics* 51, pp. 209–23.
- Kierland, Brian and Bradley Monton (2005). Minimizing Inaccuracy for Self-locating Beliefs. *Philosophy and Phenomenological Research* 70, pp. 384–95.
- Kim, Jaegwon (1988). What Is “Naturalized Epistemology”? *Philosophical Perspectives* 2, pp. 381–405.
- Kim, Namjoong (2009). Sleeping Beauty and Shifted Jeffrey Conditionalization. *Synthese* 168, pp. 295–312.
- Knight, Frank Hyneman (1921). *Risk, Uncertainty and Profit*. Boston and New York: Houghton Mifflin.
- Kolmogorov, A.N. (1933/1950). *Foundations of the Theory of Probability*. Translation edited by Nathan Morrison. New York: Chelsea Publishing Company.
- Konek, Jason and Benjamin A. Levinstein (2019). The Foundations of Epistemic Decision Theory. *Mind* 128, pp. 69–107.
- Kornblith, Hilary (1993). Epistemic Normativity. *Synthese* 94, pp. 357–76.
- Kraft, Charles H., John W. Pratt, and A. Seidenberg (1959). Intuitive Probability on Finite Sets. *The Annals of Mathematical Statistics* 30, pp. 408–19.
- Krantz, D.H. et al. (1971). *Foundations of Measurement, Vol 1: Additive and Polynomial Representations*. Cambridge: Academic Press.
- Kuhn, Thomas S. (1957). *The Copernican Revolution: Planetary Astronomy in the Development of Western Thought*. New York: MJF Books.
- Kyburg Jr, Henry E. (1961). *Probability and the Logic of Rational Belief*. Middletown: Wesleyan University Press.
- Kyburg Jr, Henry E. (1970). Conjunctivitis. In: *Induction, Acceptance, and Rational Belief*. Ed. by M. Swain. Boston: Reidel, pp. 55–82.
- Laddaga, R. (1977). Lehrer and the Consensus Proposal. *Synthese* 36, pp. 473–7.
- Lance, Mark Norris (1995). Subjective Probability and Acceptance. *Philosophical Studies* 77, pp. 147–79.
- Lange, Alexandra (2019). Can Data Be Human? The Work of Giorgia Lupi. *The New Yorker*. Published May 25, 2019.
- Lange, Marc (2000). Is Jeffrey Conditionalization Defective by Virtue of Being Non-commutative? Remarks on the Sameness of Sensory Experience. *Synthese* 123, pp. 393–403.
- Laplace, Pierre-Simon (1814/1995). *Philosophical Essay on Probabilities*. Translated from the French by Andrew Dale. New York: Springer.

- Lehman, R. Sherman (1955). On Confirmation and Rational Betting. *Journal of Symbolic Logic* 20, pp. 251–62.
- Lehrer, K. and Carl Wagner (1983). Probability Amalgamation and the Independence Issue: A Reply to Laddaga. *Synthese* 55, pp. 339–46.
- Leitgeb, Hannes and Richard Pettigrew (2010a). An Objective Justification of Bayesianism I: Measuring Inaccuracy. *Philosophy of Science* 77, pp. 201–35.
- Leitgeb, Hannes and Richard Pettigrew (2010b). An Objective Justification of Bayesianism II: The Consequences of Minimizing Inaccuracy. *Philosophy of Science* 77, pp. 236–72.
- Lele, Subhash R. (2004). Evidence Functions and the Optimality of the Law of Likelihood. In: *The Nature of Scientific Evidence: Statistical, Philosophical, and Empirical Considerations*. Chicago: University of Chicago Press, pp. 191–216.
- Levi, Isaac (1974). On Indeterminate Probabilities. *The Journal of Philosophy* 71, pp. 391–418.
- Levi, Isaac (1980). *The Enterprise of Knowledge*. Boston: The MIT Press.
- Levi, Isaac (1987). The Demons of Decision. *The Monist* 70, pp. 193–211.
- Lewis, C.I. (1946). *An Analysis of Knowledge and Valuation*. La Salle, IL: Open Court.
- Lewis, David (1971). Immodest Inductive Methods. *Philosophy of Science* 38, pp. 54–63.
- Lewis, David (1976). Probabilities of Conditionals and Conditional Probabilities. *The Philosophical Review* 85, pp. 297–315.
- Lewis, David (1979). Attitudes *de dicto* and *de se*. *The Philosophical Review* 88, pp. 513–43.
- Lewis, David (1980). A Subjectivist's Guide to Objective Chance. In: *Studies in Inductive Logic and Probability*. Ed. by Richard C. Jeffrey. Vol. 2. Berkeley: University of California Press, pp. 263–94.
- Lewis, David (1981a). Causal Decision Theory. *Australasian Journal of Philosophy* 59, pp. 5–30.
- Lewis, David (1981b). 'Why Ain'cha Rich?' *Noûs* 15, pp. 377–80.
- Lewis, David (1994). Humean Supervenience Debugged. *Mind* 103, pp. 473–90.
- Lewis, David (1996). Desire as Belief II. *Mind* 105, pp. 303–13.
- Lewis, David (2001). Sleeping Beauty: Reply to Elga. *Analysis* 61, pp. 171–6.
- Lichtenstein, S., B. Fischhoff, and L. Phillips (1982). Calibration of Probabilities: The State of the Art to 1980. In: *Judgment under Uncertainty: Heuristics and Biases*. Ed. by Daniel Kahneman, P. Slovic, and Amos Tversky. Cambridge: Cambridge University Press, pp. 306–34.
- Lindley, Dennis V. (1982). Scoring Rules and the Inevitability of Probability. *International Statistical Review* 50, pp. 1–26.
- Lindley, Dennis V. (1985). *Making Decisions*. 2nd edition. London: Wiley.
- Liu, Liping and Ronald R. Yager (2008). Classic Works of the Dempster-Shafer Theory of Belief Functions: An Introduction. In: *Classic Works of the Dempster-Shafer Theory of Belief Functions*. Ed. by Ronald R. Yager and Liping Liu. Vol. 219. Studies in Fuzziness and Soft Computing. Berlin: Springer, pp. 1–34.
- Locke, John (1689/1975). *An Essay Concerning Human Understanding*. Ed. by Peter H. Nidditch. Oxford: Oxford University Press.
- MacFarlane, John (2005). Making Sense of Relative Truth. *Proceedings of the Aristotelian Society* 105, pp. 321–39.
- Maher, Patrick (1993). *Betting on Theories*. Cambridge Studies in Probability, Induction, and Decision Theory. Cambridge: Cambridge University Press.
- Maher, Patrick (1996). Subjective and Objective Confirmation. *Philosophy of Science* 63, pp. 149–74.
- Maher, Patrick (2002). Joyce's Argument for Probabilism. *Philosophy of Science* 96, pp. 73–81.

- Maher, Patrick (2010). Explication of Inductive Probability. *Journal of Philosophical Logic* 39, pp. 593–616.
- Makinson, David C. (1965). The Paradox of the Preface. *Analysis* 25, pp. 205–7.
- Makinson, David C. (2011). Conditional Probability in the Light of Qualitative Belief Change. *Journal of Philosophical Logic* 40, pp. 121–53.
- Mayo, Deborah (2018). *Statistical Inference as Severe Testing: How to Get beyond the Statistics Wars*. Cambridge: Cambridge University Press.
- Mazurkiewicz, Stefan (1932). Zur Axiomatik der Wahrscheinlichkeitsrechnung. *Comptes rendues des séances de la Société des Sciences et des Lettres de Varsovie* 25, pp. 1–4.
- Meacham, Christopher J. G. (2008). Sleeping Beauty and the Dynamics of De Se Beliefs. *Philosophical Studies* 138, pp. 245–70.
- Meacham, Christopher J. G. (2010a). Unravelling the Tangled Web: Continuity, Internalism, Uniqueness, and Self-locating Belief. In: *Oxford Studies in Epistemology*. Ed. by Tamar Szabó Gendler and John Hawthorne. Vol. 3. Oxford: Oxford University Press, pp. 86–125.
- Meacham, Christopher J.G. (2010b). Two Mistakes Regarding the Principal Principle. *British Journal for the Philosophy of Science* 61, pp. 407–31.
- Meacham, Christopher J. G. (2016). Ur-Priors, Conditionalization, and Ur-Prior Conditionalization. *Ergo* 3, pp. 444–92.
- Meacham, Christopher J.G. and Jonathan Weisberg (2011). Representation Theorems and the Foundations of Decision Theory. *Australasian Journal of Philosophy* 89, pp. 641–63.
- Mellor, D.H. (2013). Review of *Probability in the Philosophy of Religion*. *Analysis* 73, pp. 548–54.
- Moore, David S., George P. McCabe, and Bruce A. Craig (2009). *Introduction to the Practice of Statistics*. 6th edition. New York: W.H. Freeman and Company.
- Moore, G.E. (1939). Proof of an External World. *Proceedings of the British Academy* 25, pp. 273–300.
- Moss, Sarah (2012). Updating as Communication. *Philosophy and Phenomenological Research* 85, pp. 225–48.
- Moss, Sarah (2015). Credal Dilemmas. *Noûs* 49, pp. 665–83.
- Moss, Sarah (2018). *Probabilistic Knowledge*. Oxford: Oxford University Press.
- Murphy, Allan H. (1973). A New Vector Partition of the Probability Score. *Journal of Applied Meteorology* 12, pp. 595–600.
- Murphy, Allan H. and Robert L. Winkler (1977). Reliability of Subjective Probability Forecasts of Precipitation and Temperature. *Journal of the Royal Statistical Society, Series C* 26, pp. 41–7.
- Neal, Radford M. (2006). *Puzzles of Anthropic Reasoning Resolved Using Full Non-indexical Conditioning*. Tech. rep. 0607. Department of Statistics, University of Toronto.
- Neyman, J. and Egon Pearson (1967). *Joint Statistical Papers*. Cambridge: Cambridge University Press.
- Nicod, Jean (1930). *Foundations of Geometry and Induction*. Translated by Philip Wiener. New York: Harcourt, Brace and Company.
- Nozick, Robert (1969). Newcomb's Problem and Two Principles of Choice. In: *Essays in Honor of Carl G. Hempel*. Synthese Library. Dordrecht: Reidel, pp. 114–15.
- Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science* 349. DOI: 10.1126/science.aac4716.
- Papineau, David (2012). *Philosophical Devices: Proofs, Probabilities, Possibilities, and Sets*. Oxford: Oxford University Press.
- Paris, J.B. (1994). *The Uncertain Reasoner's Companion*. Cambridge: Cambridge University Press.

- Pascal, Blaise (1670/1910). *Pensées*. Translated by W.F. Trotter. London: Dent.
- Pearson, K., A. Lee, and L. Bramley-Moore (1899). Genetic (Reproductive) Selection: Inheritance of Fertility in Man. *Philosophical Transactions of the Royal Society A* 73, pp. 534–39.
- Pedersen, Arthur Paul and Gregory Wheeler (2014). Demystifying Dilation. *Erkenntnis* 79, pp. 1305–42.
- Peirce, Charles Sanders (1910/1932). Notes on the Doctrine of Chances. In: *Collected Papers of Charles Sanders Peirce*. Ed. by Charles Hartshorne and Paul Weiss. Cambridge, MA: Harvard University Press, pp. 404–14.
- Peterson, Martin (2009). *An Introduction to Decision Theory*. Cambridge Introductions to Philosophy. Cambridge: Cambridge University Press.
- Pettigrew, Richard (2013a). A New Epistemic Utility Argument for the Principal Principle. *Episteme* 10, pp. 19–35.
- Pettigrew, Richard (2013b). Epistemic Utility and Norms for Credences. *Philosophy Compass* 8, pp. 897–908.
- Pettigrew, Richard (2014). Accuracy, Risk, and the Principle of Indifference. *Philosophy and Phenomenological Research* 92(1), pp. 35–59.
- Pettigrew, Richard (2016). *Accuracy and the Laws of Credence*. Oxford: Oxford University Press.
- Pettigrew, Richard (2021). On the Expected Utility Objection to the Dutch Book Argument for Probabilism. *Noûs* 55, pp. 23–38.
- Pettigrew, Richard and Michael G. Titelbaum (2014). Deference Done Right. *Philosophers' Imprint* 14, pp. 1–19.
- Pettigrew, Richard and Jonathan Weisberg, eds (2019). *The Open Handbook of Formal Epistemology*. Published open access online by PhilPapers. URL: <https://philpapers.org/rec/PETTOH-2>.
- Piccione, Michele and Ariel Rubinstein (1997). On the Interpretation of Decision Problems with Imperfect Recall. *Games and Economic Behavior* 20, pp. 3–24.
- Pollock, John L. (2001). Defeasible Reasoning with Variable Degrees of Justification. *Artificial Intelligence* 133, pp. 233–82.
- Popper, Karl R. (1935/1959). *The Logic of Scientific Discovery*. London: Hutchinson & Co.
- Popper, Karl R. (1938). A Set of Independent Axioms for Probability. *Mind* 47, pp. 275–9.
- Popper, Karl R. (1954). Degree of Confirmation. *British Journal for the Philosophy of Science* 5, pp. 143–9.
- Popper, Karl R. (1955). Two Autonomous Axiom Systems for the Calculus of Probabilities. *British Journal for the Philosophy of Science* 6, pp. 51–7.
- Popper, Karl R. (1957). The Propensity Interpretation of the Calculus of Probability and the Quantum Theory. *The Colston Papers* 9. Ed. by S. Körner, pp. 65–70.
- Predd, J. et al. (2009). Probabilistic Coherence and Proper Scoring Rules. *IEEE Transactions on Information Theory* 55, pp. 4786–92.
- Pryor, James (2004). What's Wrong with Moore's Argument? *Philosophical Issues* 14, pp. 349–78.
- Quinn, Warren S. (1990). The Puzzle of the Self-Torturer. *Philosophical Studies* 59, pp. 79–90.
- Ramsey, Frank P. (1929/1990). General Propositions and Causality. In: *Philosophical Papers*. Ed. by D.H. Mellor. Cambridge: Cambridge University Press, pp. 145–63.
- Ramsey, Frank P. (1931). Truth and Probability. In: *The Foundations of Mathematics and other Logic Essays*. Ed. by R.B. Braithwaite. New York: Harcourt, Brace and Company, pp. 156–98.
- Rees, Martin (2000). *Just Six Numbers: The Deep Forces that Shape the Universe*. New York: Basic Books.

- Reichenbach, Hans (1935/1949). *The Theory of Probability*. English expanded version of the German original. Berkeley: University of California Press.
- Reichenbach, Hans (1938). *Experience and Prediction*. Chicago: University of Chicago Press.
- Reichenbach, Hans (1956). The Principle of Common Cause. In: *The Direction of Time*. Berkeley: University of California Press, pp. 157–60.
- Reiss, Julian and Jan Sprenger (2017). Scientific Objectivity. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2017. Metaphysics Research Lab, Stanford University. URL: <https://plato.stanford.edu/archives/win2017/entries/scientific-objectivity/>.
- Renyi, Alfred (1970). *Foundations of Probability*. San Francisco: Holden-Day.
- Resnik, Michael D. (1987). *Choices: An Introduction to Decision Theory*. Minneapolis: University of Minnesota Press.
- Roche, William (2014). Evidence of Evidence Is Evidence under Screening-off. *Episteme* 11, pp. 119–24.
- Roeper, P. and H. Leblanc (1999). *Probability Theory and Probability Logic*. Toronto: University of Toronto Press.
- Rosenkrantz, Roger (1981). *Foundations and Applications of Inductive Probability*. Atascadero, CA: Ridgeview Press.
- Rosenthal, Jeffrey S. (2009). A Mathematical Analysis of the Sleeping Beauty Problem. *The Mathematical Intelligencer* 31, pp. 32–37.
- Royall, Richard M. (1997). *Statistical Evidence: A Likelihood Paradigm*. New York: Chapman & Hall/CRC.
- Salmon, Wesley (1966). *The Foundations of Scientific Inference*. Pittsburgh: University of Pittsburgh Press.
- Salmon, Wesley (1975). Confirmation and Relevance. In: *Induction, Probability, and Confirmation*. Ed. by Grover Maxwell and Robert M. Jr. Anderson. Vol. VI. Minnesota Studies in the Philosophy of Science. Minneapolis: University of Minnesota Press, pp. 3–36.
- Savage, Leonard J. (1954). *The Foundations of Statistics*. New York: Wiley.
- Savage, Leonard J. (1967). Difficulties in the Theory of Personal Probability. *Philosophy of Science* 34, pp. 305–10.
- Schervish, M.J., T. Seidenfeld, and J.B. Kadane (2004). Stopping to Reflect. *Journal of Philosophy* 101, pp. 315–22.
- Schervish, M.J., T. Seidenfeld, and J.B. Kadane (2009). Proper Scoring Rules, Dominated Forecasts, and Coherence. *Decision Analysis* 6, pp. 202–21.
- Schick, Frederic (1986). Dutch Bookies and Money Pumps. *The Journal of Philosophy* 83, pp. 112–19.
- Schoenfeld, Miriam (2017). The Accuracy and Rationality of Imprecise Credences. *Noûs* 51, pp. 667–85.
- Schwarz, Wolfgang (2010). *Lewis on Updating and Self-location*. Blog post available at URL: <https://www.umsu.de/wo/2010/563>.
- Schwarz, Wolfgang (2018). Subjunctive Conditional Probability. *Journal of Philosophical Logic* 47, pp. 47–66.
- Scott, Dana (1964). Measurement Structures and Linear Inequalities. *Journal of Mathematical Psychology* 1, pp. 233–47.
- Seelig, C. (1956). *Albert Einstein: A Documentary Biography*. London: Staples Press.
- Seidenfeld, Teddy (1986). Entropy and Uncertainty. *Philosophy of Science* 53, pp. 467–91.
- Seidenfeld, Teddy, M.J. Schervish, and J.B. Kadane (2017). Non-Conglomerability for Countably Additive Measures That Are Not κ -Additive. *The Review of Symbolic Logic* 10, pp. 284–300.

- Seidenfeld, Teddy and Larry Wasserman (1993). Dilation for Sets of Probabilities. *The Annals of Statistics* 21, pp. 1139–54.
- Selvin, Steve (1975). A Problem in Probability. *The American Statistician* 29. Published among the Letters to the Editor, p. 67.
- Shafer, Glenn (1976). *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton University Press.
- Shafer, Glenn (1981). Constructive Probability. *Synthese* 48, pp. 1–60.
- Shapiro, Amram, Louise Firth Campbell, and Rosalind Wright (2014). *The Book of Odds*. New York: Harper Collins.
- Shimony, Abner (1955). Coherence and the Axioms of Confirmation. *Journal of Symbolic Logic* 20, pp. 1–28.
- Shimony, Abner (1988). An Adamite Derivation of the Calculus of Probability. In: *Probability and Causality*. Ed. by J.H. Fetzer. Dordrecht: Reidel, pp. 151–61.
- Shogenji, Tomoji (2003). A Condition for Transitivity in Probabilistic Support. *British Journal for the Philosophy of Science* 54, pp. 613–16.
- Shogenji, Tomoji (2012). The Degree of Epistemic Justification and the Conjunction Fallacy. *Synthese* 184, pp. 29–48.
- Shope, R.K. (1978). The Conditional Fallacy in Contemporary Philosophy. *Journal of Philosophy* 75, pp. 397–413.
- Shortliffe, E. and B. Buchanan (1975). A Model of Inexact Reasoning in Medicine. *Mathematical Biosciences* 23, pp. 351–79.
- Simpson, E.H. (1951). The Interpretation of Interaction in Contingency Tables. *Journal of the Royal Statistical Society, Series B* 13, pp. 238–41.
- Skipper, Mattias and Jens Christian Bjerring (2022). Bayesianism for Nonideal Agents. *Erkenntnis* 87, pp. 93–115.
- Skyrms, Brian (1980a). *Causal Necessity: A Pragmatic Investigation of the Necessity of Laws*. New Haven CT: Yale University Press.
- Skyrms, Brian (1980b). Higher Order Degrees of Belief. In: *Prospects for Pragmatism*. Ed. by D.H. Mellor. Cambridge: Cambridge University Press, pp. 109–37.
- Skyrms, Brian (1983). Three Ways to Give a Probability Function a Memory. In: *Testing Scientific Theories*. Ed. by John Earman. Vol. 10. Minnesota Studies in the Philosophy of Science. Minneapolis: University of Minnesota Press, pp. 157–61.
- Skyrms, Brian (1987a). Coherence. In: *Scientific Inquiry in Philosophical Perspective*. Ed. by N. Rescher. Pittsburgh: University of Pittsburgh Press, pp. 225–42.
- Skyrms, Brian (1987b). Dynamic Coherence and Probability Kinematics. *Philosophy of Science* 54, pp. 1–20.
- Skyrms, Brian (2000). *Choice & Chance: An Introduction to Inductive Logic*. 4th edition. Stamford, CT: Wadsworth.
- Smith, C.A.B. (1961). Consistency in Statistical Inference and Decision. *Journal of the Royal Statistical Society, Series B* 23, pp. 1–25.
- Smithies, Declan (2015). Ideal Rationality and Logical Omniscience. *Synthese* 192, pp. 2769–93.
- Sober, Elliott (2008). *Evidence and Evolution*. Cambridge: Cambridge University Press.
- Spohn, Wolfgang (2012). *The Laws of Belief: Ranking Theory and its Philosophical Applications*. Oxford: Oxford University Press.
- Staffel, Julia (2019). *Unsettled Thoughts: Reasoning and Uncertainty in Epistemology*. Oxford: Oxford University Press.
- Stalnaker, Robert C. (1972/1981). Letter to David Lewis. In: *Ifs: Conditionals, Belief, Decision, Chance, and Time*. Ed. by W. Harper, Robert C. Stalnaker, and G. Pearce. Dordrecht: Reidel, pp. 151–2.

- Stalnaker, Robert C. (2008). *Our Knowledge of the Internal World*. Oxford: Oxford University Press.
- Stalnaker, Robert C. (2011). Responses to Stoljar, Weatherson and Boghossian. *Philosophical Studies* 155, pp. 467–79.
- Stefánsson, H. Orri (2017). What Is “Real” in Probabilism? *Australasian Journal of Philosophy* 95, pp. 573–87.
- Stefánsson, H. Orri (2018). On the Ratio Challenge for Comparativism. *Australasian Journal of Philosophy* 96, pp. 380–90.
- Stephenson, Todd A. (2000). *An Introduction to Bayesian Network Theory and Usage*. Tech. rep. 03. IDIAP.
- Sturgeon, Scott (2010). Confidence and Coarse-grained Attitudes. In: *Oxford Studies in Epistemology*. Ed. by Tamar Szabó Gendler and John Hawthorne. Vol. 3. Oxford: Oxford University Press, pp. 126–49.
- Suppes, Patrick (1955). The Role of Subjective Probability and Utility in Decision-making. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, pp. 61–73.
- Suppes, Patrick (1969). *Studies in the Methodology and Foundations of Science*. Berlin: Springer.
- Suppes, Patrick (1974). *Probabilistic Metaphysics*. Uppsala: University of Uppsala Press.
- Suppes, Patrick (2002). *Representation and Invariance of Scientific Structures*. Stanford, CA: CSLI Publications.
- Tal, Eyal and Juan Comesaña (2017). Is Evidence of Evidence Evidence? *Noûs* 51, pp. 95–112.
- Talbott, William J. (1991). Two Principles of Bayesian Epistemology. *Philosophical Studies* 62, pp. 135–50.
- Talbott, William J. (2005). Review of “Putting Logic in its Place: Formal Constraints on Rational Belief”. *Notre Dame Philosophical Reviews*. URL: <http://ndpr.nd.edu/>.
- Talbott, William J. (2016). Bayesian Epistemology. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2016. Metaphysics Research Lab, Stanford University. URL: <https://plato.stanford.edu/archives/win2016/entries/epistemology-bayesian/>.
- Teller, Paul (1973). Conditionalization and Observation. *Synthese* 26, pp. 218–58.
- Tennant, Neil (2017). Logicism and Neologicism. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2017. Metaphysics Research Lab, Stanford University. URL: <https://plato.stanford.edu/archives/win2017/entries/logicism/>.
- Tentori, Katya, Vincenzo Crupi, and Selena Russo (2013). On the Determinants of the Conjunction Fallacy: Probability versus Inductive Confirmation. *Journal of Experimental Psychology: General* 142, pp. 235–55.
- Titelbaum, Michael G. (2008). The Relevance of Self-locating Beliefs. *Philosophical Review* 117, pp. 555–605.
- Titelbaum, Michael G. (2010). Not Enough There There: Evidence, Reasons, and Language Independence. *Philosophical Perspectives* 24, pp. 477–528.
- Titelbaum, Michael G. (2013a). *Quitting Certainties: A Bayesian Framework Modeling Degrees of Belief*. Oxford: Oxford University Press.
- Titelbaum, Michael G. (2013b). Ten Reasons to Care about the Sleeping Beauty Problem. *Philosophy Compass* 8, pp. 1003–17.
- Titelbaum, Michael G. (2015a). Continuing On. *Canadian Journal of Philosophy* 45, pp. 670–91.
- Titelbaum, Michael G. (2015b). Rationality’s Fixed Point (Or: In Defense of Right Reason). In: *Oxford Studies in Epistemology*. Ed. by Tamar Szabó Gendler and John Hawthorne. Vol. 5. Oxford University Press, pp. 253–94.

- Titelbaum, Michael G. (2016). Self-locating Credences. In: *The Oxford Handbook of Probability and Philosophy*. Ed. by Alan Hájek and Christopher R. Hitchcock. Oxford: Oxford University Press, pp. 666–80.
- Titelbaum, Michael G. (2019). Precise Credences. In: *The Open Handbook of Formal Epistemology*. Ed. by Richard Pettigrew and Jonathan Weisberg. The PhilPapers Foundation, pp. 1–56.
- Tversky, Amos and Daniel Kahneman (1974). Judgment under Uncertainty: Heuristics and Biases. *Science* 185, pp. 1124–31.
- Tversky, Amos and Daniel Kahneman (1983). Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment. *Psychological Review* 90, pp. 293–315.
- Tversky, Amos and Daniel Kahneman (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty* 5, pp. 297–323.
- Tversky, Amos and Derek J. Koehler (1994). Support Theory: A Nonextensional Representation of Subjective Probability. *Psychological Review* 101, pp. 547–67.
- van Enk, Steven J. (2015). Betting, Risk, and the Law of Likelihood. *Ergo* 2, pp. 105–21.
- van Fraassen, Bas C. (1980). Rational Belief and Probability Kinematics. *Philosophy of Science* 47, pp. 165–87.
- van Fraassen, Bas C. (1981). A Problem for Relative Information Minimizers. *British Journal for the Philosophy of Science* 32, pp. 375–79.
- van Fraassen, Bas C. (1982). Rational Belief and the Common Cause Principle. In: *What? Where? When? Why?* Ed. by Robert McLaughlin. Dordrecht: Reidel, pp. 193–209.
- van Fraassen, Bas C. (1983). Calibration: A Frequency Justification for Personal Probability. In: *Physics Philosophy and Psychoanalysis*. Ed. by R. Cohen and L. Laudan. Dordrecht: Reidel, pp. 295–319.
- van Fraassen, Bas C. (1984). Belief and the Will. *The Journal of Philosophy* 81, pp. 235–56.
- van Fraassen, Bas C. (1989). *Laws and Symmetry*. Oxford: Clarendon Press.
- van Fraassen, Bas C. (1995). Belief and the Problem of Ulysses and the Sirens. *Philosophical Studies* 77, pp. 7–37.
- van Fraassen, Bas C. (1999). Conditionalization: A New Argument For. *Topoi* 18, pp. 93–6.
- van Fraassen, Bas C. (2005). Conditionalizing on Violated Bell's Inequalities. *Analysis* 65, pp. 27–32.
- Van Horn, Kevin S. (2003). Constructing a Logic of Plausible Inference: A Guide to Cox's Theorem. *International Journal of Approximate Reasoning* 34, pp. 3–24.
- Venn, John (1866). *The Logic of Chance*. London, Cambridge: Macmillan.
- Vineberg, Susan (2011). Dutch Book Arguments. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2011. URL: <https://plato.stanford.edu/archives/su2011/entries/dutch-book/>.
- von Mises, Richard (1928/1957). *Probability, Statistics and Truth*. (English edition of the original German *Wahrscheinlichkeit, Statistik und Wahrheit*.) New York: Dover.
- von Neumann, J. and O. Morgenstern (1947). *Theory of Games and Economic Behavior*. 2nd edition. Princeton, NJ: Princeton University Press.
- Vranas, Peter B.M. (2004). Hempel's Raven Paradox: A Lacuna in the Standard Bayesian Solution. *British Journal for the Philosophy of Science* 55, pp. 545–60.
- Wainer, Howard (2011). *Uneducated Guesses: Using Evidence to Uncover Misguided Education Policies*. Princeton, NJ: Princeton University Press.
- Walley, Peter (1991). *Statistical Reasoning with Imprecise Probabilities*. London: Chapman and Hall.
- Wasserstein, Ronald L. and Nicole A. Lazar (2016). The ASA's Statement on p-values: Context, Process, and Purpose. *The American Statistician*. DOI: 10.1080/00031305.2016.1154108.

- Weatherson, Brian (2002). Keynes, Uncertainty and Interest Rates. *Cambridge Journal of Economics* 26, pp. 47–62.
- Weatherson, Brian (2011). Stalnaker on Sleeping Beauty. *Philosophical Studies* 155, pp. 445–56.
- Weatherson, Brian (2015). For Bayesians, Rational Modesty Requires Imprecision. *Ergo* 2, pp. 529–45.
- Weatherson, Brian and Andy Egan (2011). Epistemic Modals and Epistemic Modality. In: *Epistemic Modality*. Ed. by Andy Egan and Brian Weatherson. Oxford: Oxford University Press, pp. 1–18.
- Weintraub, Ruth (2001). The Lottery: A Paradox Regained and Resolved. *Synthese* 129, pp. 439–49.
- Weirich, Paul (2012). Causal Decision Theory. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2012. URL: <https://plato.stanford.edu/archives/win2012/entries/decision-causal/>.
- Weisberg, Jonathan (2007). Conditionalization, Reflection, and Self-Knowledge. *Philosophical Studies* 135, pp. 179–97.
- Weisberg, Jonathan (2009). Varieties of Bayesianism. In: *Handbook of the History of Logic*. Ed. by Dov M. Gabbay, Stephan Hartmann, and John Woods. Vol. 10: Inductive Logic. Oxford: Elsevier.
- White, Roger (2005). Epistemic Permissiveness. *Philosophical Perspectives* 19, pp. 445–59.
- White, Roger (2006). Problems for Dogmatism. *Philosophical Studies* 131, pp. 525–57.
- White, Roger (2010). Evidential Symmetry and Mushy Credence. In: *Oxford Studies in Epistemology*. Ed. by Tamar Szabó Gendler and John Hawthorne. Vol. 3. Oxford: Oxford University Press, pp. 161–86.
- Williams, J. Robert G. (ms). A Non-Pragmatic Dominance Argument for Conditionalization. Unpublished manuscript.
- Williams, J. Robert G. (2016). Probability and Non-Classical Logic. In: *Oxford Handbook of Probability and Philosophy*. Ed. by Alan Hájek and Christopher R. Hitchcock. Oxford: Oxford University Press.
- Williamson, Timothy (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.
- Williamson, Timothy (2007). How Probable Is an Infinite Sequence of Heads? *Analysis* 67, pp. 173–80.
- Wittgenstein, Ludwig (1921/1961). *Tractatus Logico-Philosophicus*. Translated by D.F. Pears and B.F. McGuinness. London: Routledge.
- Wong, S.M. et al. (1991). Axiomatization of Qualitative Belief Structure. *IEEE Transactions on Systems, Man and Cybernetics* 21, pp. 726–34.
- Woolston, Chris (2015). Psychology Journal Bans P Values. *Nature* 519.9. URL: <https://doi.org/10.1038/519009f>.
- Wroński, Leszek and Godziszewski, Michał Tomasz (2017). *The Stubborn Non-probabilist—‘Negation Incoherence’ and a New Way to Block the Dutch Book Argument*. International Workshop on Logic, Rationality and Interaction. Berlin, Heidelberg: Springer, pp. 256–67.
- Yalcin, Seth (2012). A Counterexample to Modus Tollens. *Journal of Philosophical Logic* 41, pp. 1001–24.
- Yule, G.U. (1903). Notes on the Theory of Association of Attributes in Statistics. *Biometrika* 2, pp. 121–34.
- Zynda, Lyle (1995). Old Evidence and New Theories. *Philosophical Studies* 77, pp. 67–95.
- Zynda, Lyle (2000). Representation Theorems and Realism about Degrees of Belief. *Philosophy of Science* 67, pp. 45–69.