Confirmation, Decision, and Evidential Probability

PEDEN, WILLIAM,JOHN

CONFIRMATION, DECISION, AND EVIDENTIAL PROBABILITY

William Peden

Henry Kyburg's theory of Evidential Probability offers a neglected tool for approaching problems in confirmation theory and decision theory. I use Evidential Probability to examine some persistent problems within these areas of the philosophy of science. Formal tools in general and probability theory in particular have great promise for conceptual analysis in confirmation theory and decision theory, but they face many challenges.

In each chapter, I apply Evidential Probability to a specific issue in confirmation theory or decision theory. In Chapter 1, I challenge the notion that Bayesian probability offers the best basis for a probabilistic theory of evidence. In Chapter 2, I criticise the conventional measures of quantities of evidence that use the degree of imprecision of imprecise probabilities. In Chapter 3, I develop an alternative to orthodox utility-maximizing decision theory using Kyburg's system. In Chapter 4, I confront the orthodox notion that Nelson Goodman's New Riddle of Induction makes purely formal theories of induction untenable. Finally, in Chapter 5, I defend probabilistic theories of inductive reasoning against John D. Norton's recent collection of criticisms.

My aim is the development of fresh perspectives on classic problems and contemporary debates. I both defend and exemplify a formal approach to the philosophy of science. I argue that Evidential Probability has great potential for clarifying our concepts of evidence and rationality.

CONFIRMATION, DECISION, AND EVIDENTIAL PROBABILITY

A thesis submitted for the degree of

Doctor of Philosophy

by

William Peden

Department of Philosophy

Durham University

2017

I confirm that no part of the material contained in this thesis has previously been submitted for any degree in this or any other university. All the material is the author's own work, except for quotations and paraphrases which have been suitably indicated.

William Peden

ACKNOWLEDGEMENTS

A canny set of eyes is crucial to developing and revising philosophical writing. I have been fortunate to have three pairs of eyes assisting me; their owners have also provided instruction, advice, encouragement, and example into the bargain. Julian Reiss and Nancy Cartwright have been tremendous supervisors; I knew in 2013 that there was nowhere else in the world that could offer a better-suited team for me and I am very grateful to them both. It was in 2009 that Nancy became the first philosopher to suggest that I undertake a PhD and it is wonderful that she has become such an important part of this journey. Furthermore, I must give special thanks to Wendy Parker, who has been a brilliant additional supervisor.

Any PhD must build on past knowledge and on enthusiasm that was nurtured in the past. I have had a number of inspiring and helpful instructors over the years, but special thanks must go to Mr. Johnson, Mr. Vaughn-Sharpe, and Ms. Cunningham, who taught me the essentials of writing a discursive essay and how to teach a class of teenagers; Dr. Richmond and Dr. Levy, the salient kindlers of my interest in philosophy at Edinburgh; and Professor Milne, who has never officially instructed me and yet who has always been a rich source of information, advice, and assistance.

I was privileged to present work based on this thesis at several places in Durham: the Philosophy Departmental Seminar; the Global Policy Institute; the Institute for Hazard, Risk, and Resilience; and the postgraduate philosophy society Eidos. Furthermore, at the Centre for Humanities Engaging Science and Society (CHESS), I have had multiple opportunities to present and discuss my work, which has been greatly improved as a result. I have also been assisted by frequent mentoring from Matthew, Sara, and Geoffrey in the Durham Philosophy

Department. Finally, I have benefited from discussions at the Durham University Mathematics Department and the Durham University Economics and Finance Department. Outwith Durham, I am grateful to the Non-Monotonic Logic Group at Bochum, who organised a very useful conference on argument strength and gave me the opportunity to give a talk based on Chapter 3 of this thesis. Chapter 4 was modified based on the comments of an anonymous referee at *Erkenntnis*.

My family has been a tireless source of emotional and intellectual support. When my life has been dramatic and painful during the past few years, I have always known that they were there to offer reassurance, advice, and attentive ears. Beyond this assistance, I am indebted to my parents and grandmother for their exemplification of the virtues of patience, kindness, perseverance, indefatigability, and forgiveness, which has guided me even when their direct support was unavailable.

Friends, both old and new, have helpful me a lot during the writing of this thesis. Perhaps it is atypical for a doctoral study to be a time of increased sociability, but I have been lucky enough to acquire a large number of new friends in the past four years, as well as rekindling several very special old friendships. Any attempt at a complete list would be doomed to fail, but particularly special thanks must go to Ross, Steph, Alex, David, Rune, Jen, Lenka, Jovita, Liz, Simon, Sarah, Anthony, Anna, Alasdair, Fergus, Findlay, Darren, Jamie, Yang, Erin, Katherine, and Sasha, who have all provided close friendship at important challenging moments, even if I was being more than a little thrawn. Additionally, Lars Christensen and Ryan Murphy have provided an especially helpful combination of friendship, encouragement, advice, and professional support. This list is far from exhaustive.

Contents

INTRODUCTION

Science has a tremendous importance for epistemologists. In turn, *confirmation theory* and *normative decision theory* are important areas of the epistemology of science. Confirmation theorists are interested in modelling the standards for evidence to either support or undermine a hypothesis. Normative decision theorists aim to model ideally rational choice; in the context of the philosophy of science, their discipline is especially relevant for modelling rational decision-making regarding (1) the theories to select for acceptance or for testing, (2) the methods with which to develop them or test them, and (3) other choices made by scientists. These formal models are not historical or sociological descriptions of actual scientific practice, nor are they descriptions of "the" scientific view of confirmation and decision, since different scientists have different ideas about these concepts; instead, they are models of ideal reasoning.

There are many reasons why models can be useful for confirmation theory and decision theory. I shall list several:

Confirmation Theory: Unclear Cases

Sometimes it is intuitive that a body of evidence confirms or disconfirms a hypothesis. It might also be clear that this relation is weak or strong. In other cases, these claims are far more controversial. It would be useful to have standards of reasoning that conform to our firm judgements of the evidential relations (at least insofar as these judgements withstand scrutiny) but also provide useful guidance when our intuitions are weak or conflicting. This programme in confirmation theory is comparable to how some philosophers have viewed formal deductive logic[1]. For instance, philosophers disagree about whether empirical evidence can confirm hypotheses like Newton's laws (interpreted as universal generalizations over an infinite number of objects) as Rudolf Carnap[2] denies this claim, whereas Colin Howson and Peter Urbach affirm it[3]. This dispute could be resolved by establishing a common confirmation theory.

Confirmation Theory: Conflicting General Claims

Plausible general claims about confirmation can be incompatible. For example, consider the claims:

(a) If a hypothesis H is more probable after learning some evidence E, then E confirms H.

(b) A set of inconsistent hypotheses cannot each be confirmed by internally consistent evidence.

(c) The hypotheses 'Average British male height is 182.5 cm' and 'Average British male height is 183.5 cm' can both become more probable upon learning that the sample mean of height for a sample of British males is 183 cm.

At least one of these claims must be false. Some confirmation theories will endorse

---

[1] Kyburg and Teng (2001) p. 35-39.

[2] Carnap (1962) p. 570-571.

[3] Howson and Urbach (1993) p. 392.

one or two of these theses, but no logically consistent confirmation theory can allow all three. Therefore, the choice of confirmation theory can help adjudicate between apparently correct but logically inconsistent general claims about confirmation.

Normative Decision-Theory

Normative decision theorists aim to formalise standards of rational choice for an agent facing uncertainty. For this inquiry, confirmation theory has a fundamental importance, because such an investigation involves epistemic concepts such as evidence, uncertainty, rational belief revision, and personal probability. Confirmation theorists can use formal methods to clarify, systematise, and analyse these epistemic concepts in decision theory.

Today, the predominant confirmation theory is Bayesianism. In this theory, one models confirmation using the probabilistic relations between statements. In general, if the hypothesis is *more* probable given the evidence than if the evidence were false, then the hypothesis is confirmed by the evidence; if the hypothesis is *less* probable given the evidence than if the evidence is false, then the hypothesis is disconfirmed by the evidence. Subjective Bayesianism probabilities are typically (a) real numbers[4] and (b) almost entirely determined via subjective judgements. Insights from Bayesianism are used in decision theory, in statistics, critical thinking courses, and in other disciplines[5].

Confirmation theorists such as Bayesians often formulate definitions of two distinct

[4] They are points along the number line. Bayesian probabilities are typically fractions with values from 0 to 1. In Chapter 4, I discuss Imprecise Bayesians, who do not require that probabilities are real numbers.

[5] Talbott (2016).

but related concepts. Firstly, there is the qualitative concept of some evidence confirming, disconfirming, or being neutral towards a hypothesis. Confirmation, in this qualitative sense, is binary: either the evidence confirms the hypothesis (perhaps relative to some background knowledge) or it does not. For example, as I shall discuss in Chapter 1 Subsection 2.2, many Bayesians define this concept as favourable probabilistic relevance[6]. Secondly, there is the concept of the *degree* to which some evidence confirms or disconfirms a hypothesis. There are many differing analyses of this concept. Some major rival definitions of the *degree of confirmation* of a hypothesis H by some evidence E relative to background knowledge K include:

The Difference Measure: $(H, E \mid K) = P(H \mid E \wedge K) - P(H \mid K)$.

The Log-Ratio Measure: $\log(\frac{P(H \mid E \wedge K)}{P(H \mid K)})$.

The Log-Likelihood Measure: $\log(\frac{P(E \mid H \wedge K)}{P(E \mid \neg H \wedge K)})$.

The Normalized Difference Measure: $P(H \mid E \wedge K) - P(H \mid \neg E \wedge K)$.

Bayesian epistemologists who have endorsed a degree of confirmation in recent decades have tended to adopt one of these measures, a measure that is ordinally equivalent to one of these measures, or a measure that is similar except without an explicit reference to background knowledge[7]. On all of these definitions, E has a positive degree of confirmation with respect to H relative to K if and only if E is positively probabilistically relevant to H relative to K. In other words, for all of the above measures, the degree of confirmation is

[6] Fitelson (2008) p. 620.

[7] Fitelson (2001) p. 124.

greater than zero if and only if P(H | E ^ K) > P(H | K)[8]. Consequently, a Bayesian who adopted a *qualitative* definition of confirmation in terms of positive *quantitative* degree of confirmation on one of the above measures would be effectively adopting the standard definition. In this thesis, my discussions of confirmation will be almost entirely focused on qualitative definitions of confirmation, rather than definitions of degrees of confirmation.

There have also been definitions of concepts close to degrees of confirmation that do not require favourable probabilistic relevance. For instance, Popper does not believe that scientific hypotheses can be more probable given some evidence than in the absence of that evidence, because he believes that P(H | E) = 0, for any scientific hypothesis H. Nonetheless, he offers a definition of degrees of confirmation for H given E relative to K:

Popperian Definition: $\left(\frac{P(E \mid H \wedge K) - P(E \mid K)}{P(E \mid H \wedge K) + P(E \mid K)}\right) \; (1 + P(H \mid K)\, P(H \mid E \wedge K))$[9]

This definition is in the context of Popper's probability system, in which conditional probabilities like P(E | H ^ K) are defined even when P(H) = 0. He also offers an almost identical definition without the reference to background knowledge K[10].

A Bayesian could also adopt such a definition. If, like Popper, they defined conditional probabilities such as P(E | H ^ K) was defined even when P(H) = 0, then it would

[8] Eells and Fitelson (2000) p. 663-664.

[9] Popper (1980) p. 401.

[10] Popper (1980) p. 400. Popper also notes a simpler approach that also satisfies his criteria in a footnote on page 400.

be possible for the degree of confirmation to be increased by additions to E without increasing the probability of H. On the other hand, if conditional probability is defined as in Chapter 1 Subsection 2.1.1, then this will not be possible.

Although probabilistic analyses of confirmation are predominant, they are nonetheless controversial. Peter Achinstein argues that the notion of evidential support cannot be defined using probabilities[11]. Susan Haack argues that Bayesian probabilities do not correspond to our concepts of evidential support[12]. John D. Norton has recently provided a useful compilation of the most powerful outstanding criticisms of Bayesian confirmation theory[13].

Most probabilist confirmation theories are Bayesian, but it is possible to be a non-Bayesian probabilist. John Maynard Keynes developed one of the most influential probabilistic confirmation theories, but he was not a Bayesian[14]. While Keynes and the Bayesians agree that confirmation is ultimately a probabilistic notion, they disagree on many other important issues. For instance, Bayesian probabilities are real-valued, whereas Keynesian probabilities are not always real-valued[15]. In general, all Bayesian confirmation theorists are probabilists, but not all probabilists are Bayesians.

Henry E. Kyburg was another non-Bayesian probabilist. His theory of Evidential Probability differs from Bayesianism in a number of important respects. One important

[11] Achinstein (1994).

[12] Haack (2003) p. 75-76.

[13] Norton (2011).

[14] Keynes (1921).

[15] Keynes (1921) p. 27-28.

difference is that, in Evidential Probability, the probability values are expressed using two real numbers, rather than just one. Another difference is that a subjective choice of values determines the values of Subjective Bayesian probabilities, whereas evidential probabilities are always derived from information about relative frequencies.

In this thesis, I apply Kyburg's theory of Evidential Probability to examine several different topics in confirmation theory and decision theory. I argue that probabilist confirmation theorists can avoid some classic objections and paradoxes by using the resources from Evidential Probability. One common thread is that Kyburg's system helps confirmation theorists to focus on some neglected aspects of evidential support and these overlooked dimensions of reasoning are useful for addressing these objections and paradoxes.

In Chapter 1, I critically examine the project of analysing the concept of 'relevant evidence' using probabilistic theories of confirmation. I begin by providing some reasons why evidential relevance is important, before describing the prevalent Bayesian probabilistic analysis of this concept. Their analysis faces a severe challenge from Karl Popper's 'Paradox of Ideal Evidence', which indicates that the Bayesian definition of evidence is too narrow. I discuss some probabilist responses to Popper's paradox, which are variously too broad or too narrow. To set up my alternative, I explain Kyburg's Evidential Probabilist theory. I use this probability theory as the basis of a revised definition of evidential relevance. I argue that this definition avoids the Paradox of Ideal Evidence and provides an analysis of evidential relevance that is neither too broad nor too narrow.

In Chapter 2, I look at proposals that have been made for using imprecise probabilities to measure the quantity of relevant evidence. Firstly, I introduce the concept of the quantity of relevant evidence, which Keynes called the “weight of argument”. Secondly, I discuss a measure developed by Peter Walley in *Statistical Reasoning with Imprecise Probabilities* (1991). I argue that it performs very poorly as a measure of the quantity of relevant evidence. Thirdly, I discuss a measure proposed by Kyburg, which performs somewhat better. However, it is nonetheless very problematic, and I briefly suggest an alternative research programme for measuring the weight of argument using Evidential Probability.

In Chapter 3, I discuss the Ellsberg Paradox. This is a problem in normative decision theory, in which most people’s choices in a type of decision-problem are irrational, according to the standard normative decision theory, yet their reasoning does not seem mistaken. I describe the standard approach to normative decision theory and explain the challenge that the Ellsberg Paradox presents to it. I consider some prominent responses to this paradox and argue that neither response is entirely satisfactory. I discuss Kyburg’s response and conclude that it is interesting but unnecessarily radical, given the other tools that his theory offers. Finally, I use these tools from Evidential Probability develop an answer by developing a novel decision theory that avoids the Ellsberg Paradox and yet retains much of standard decision theory.

One of the most historically important paradoxes in confirmation theory has been Nelson Goodman’s New Riddle of Induction. Most philosophers of science regard this paradox as a fatal problem for confirmation theories that use only formal relations to analyse evidential support. In Chapter 4, I challenge this consensus by arguing that Goodman’s

Riddle does not pose a problem for formalism as such. I begin by discussing the historical impact of the Riddle on confirmation theory. Since Goodman's principal target was Carl Hempel's formalist theory of confirmation, I discuss whether Hempel's theory is directly challenged by Goodman's Riddle. I argue that the New Riddle of Induction *as such* is not a problem for Hempel's theory, but only because of the severe limitations of that theory. A formalist should want to prove that a richer confirmation theory can also avoid Goodman's paradox. To answer the Riddle, I consider the concept of the reliability of evidence. I argue that a sophisticated formalist can answer the New Riddle of Induction by formalising and utilising this aspect of scientific reasoning. I describe how this can be done in Kyburg's model of scientific knowledge. Interestingly, by weakening a common idealization in confirmation theory (that the evidence is always known with complete certainty) a formalist can answer their greatest single challenge.

As I noted earlier, the probabilistic approach to studying confirmation is popular but controversial, and Norton is a prominent contemporary critic. In Chapter 5 I argue that, even if Norton's arguments against Bayesianism are sound, a probabilist can avoid them by using Evidential Probability. I focus my discussion on inductive reasoning Firstly, I distinguish inductive reasoning from other parts of the scientific method. Secondly, I present an Evidential Probabilist theory of confirmation and illustrate it using a variety of forms of induction. Thirdly, I discuss Norton's criticisms of Bayesianism and argue that none of them presents a problem for an Evidential Probabilist. I conclude that probabilists who are troubled by Norton's criticisms of Bayesianism can view Kyburg's system as offering an alternative theory of induction.

Bayesianism has reached tremendous prominence in philosophy. As a result, it has suffered the criticisms that come with such prominence. By providing numerous new inquiries using Evidential Probability, my thesis provides a rare discussion of a neglected cousin of Bayesianism. Like a cousin, it shares some family resemblances and yet it has some very distinctive features. At the same time, I provide a discussion of important topics regarding confirmation and decision that have too often been neglected. Consequently, my thesis is a novel contribution to the disciplines of confirmation theory and normative decision theory.

CHAPTER 1: RELEVANCE AND THE PARADOX OF IDEAL EVIDENCE

For many philosophers modelling scientific reasoning, reducing methodological concepts to probabilistic relations has a great appeal. Probability is a concept that has undergone voluminous mathematical and philosophical investigation. Evidence is one concept for which there is a standard probabilistic analysis[16]. This concept plays an important role in the philosophy of science: methodologists claim that one should (ideally) consider all the *relevant* evidence in scientific reasoning; theories of rational choice involve identifying and appropriately responding to *relevant* evidence; and as I discuss in Chapter 2, a significant number of philosophers have argued that there is an importance to the concept of the quantity of *relevant* evidence.

The standard probabilistic theory of evidential relevance comes from Bayesianism: a statement is evidentially relevant to another statement if and only if it is probabilistically relevant within Bayesian probability theory. However, Popper's 'Paradox of Ideal Evidence' challenges this analysis. Popper presents a scenario in which the Bayesian definition misses apparently relevant evidence.

In this chapter, I argue that an alternative definition that uses Kyburg's system of Evidential Probability can avoid Popper's paradox. In Section 1, I explain evidential relevance and its importance. In Section 2, I examine the standard probabilistic approach to analysing evidential relevance. In Section 3, I present Popper's paradox. In Section 4, I critically examine the existing responses to the paradox. I describe Kyburg's theory of

---

[16] Fitelson (2008) p. 620.

Evidential Probability in Section 5. Finally, in Section 6, I use this theory to answer the Paradox of Ideal Evidence.

SECTION 1: EVIDENTIAL RELEVANCE

No statement is ‘evidence’ by itself; a statement can only be evidence *regarding* some other statement. I shall use ‘evidential relevance’ to refer to this relation between the evidence and the other statement. All logically contingent statements are evidentially relevant to at least one other statement (for instance, their negation) and in this weak sense they are ‘evidence’, but ‘evidential relevance’ refers to the relationship between a particular hypothesis and a particular evidence-statement. An evidence-statement (often simply called “evidence”) often provides a description of some non-linguistic evidence, which could consist of observations, measurements, testimony, and so on.

There is an important distinction between relevance and confirmation. If a statement confirms a hypothesis, then it is evidentially relevant to that hypothesis. However, statements can be evidentially relevant to a hypothesis without confirming it. Most obviously, they can disconfirm the hypothesis. For instance, a statement reporting the discovery of a quadruple star system is relevant to the hypothesis that ‘Most multiple star systems are triples’, but the evidential relation is disconfirmation.

Evidential relevance plays an important role in every aspect of science, politics, and our personal lives. For instance, imagine that a research institute has commissioned a team of economists to study whether the inclusion of Turkey into the European Union will increase

UK GDP. Additionally, this team of economists has a mandate to acquire some new data on the issue. In accordance with their mandate, the economists acquire statistical data on the issue and perform statistical tests. How can the researchers tell if they have achieved their mandate of acquiring relevant evidence? It would be helpful if the philosophy of science could provide a clear and principled answer to this question, as well as more complex applied questions about evidence.

At a more abstract level, there are methodological principles that refer to the notion of evidence. For example, Carnap's "Requirement of Total Evidence" is a popular methodological thesis: one should try to take into account *all* of the available relevant evidence when using a confirmation theory to evaluate the plausibility of a hypothesis[17]. Some philosophers, like Alan Hájek, have criticised Carnap's principle by claiming that what constitutes "relevant evidence" is obscure[18]. For example, suppose you are wondering if there is a black hole in a particular region of space. You learn that there was a minute shortening in a super-fine rod. You also have strong reasons to believe that the known background causes cannot explain this phenomenon. In this context, a report of the measurement of the rod seems to confirm that the black hole exists. Therefore, the report is relevant evidence. However, you also acquired a large quantity of other knowledge: that this measurement occurred at a particular time in your field of reference, the implicit knowledge that it coincided with the Red Spot storm on Jupiter, and an indefinitely expandable set of second-order statements like 'I learned that this event occurred', 'I learned that I learned that this event occurred' and so on. How should you identify the relevant evidence from among this

[17] Carnap (1947) p. 138-139.

[18] Hájek (2012).

vast motley of information? If it is impossible to demarcate the "total relevant evidence" from the total available knowledge, then the satisfaction conditions of the Requirement of Total Evidence are opaque. Therefore, the clarification of evidential relevance has a methodological importance.

Evidential relevance is also a significant part of normative decision theory, which is the study of ideally rational choice. Ken Gemes notes that philosophers could clarify decision-theoretic imperatives like "Do not expend resources on irrelevant Evidence" by analysing the concept of evidential relevance[19]. Perhaps no general analysis of this concept is possible, but *if* such a characterisation can be achieved, then it would enrich normative decision theory. Gemes's point is especially significant for the development of artificial intelligence: while human judgements of evidential relevance are often intuitive and subconscious, such a reliance on tacit reasoning is not possible in the design of artificial intelligence. If confirmation theorists can provide a general and formal method for identifying whether information bears on a hypothesis, then this would be a useful contribution to artificial intelligence research.

Additionally, a sound understanding of evidential relevance is a requirement for the application of a theory of rationality. As Hempel argues, the assessment of the rationality of a particular agent's decision at a time *t* requires (1) the identification of the information that was available to that agent at *t* and (2) an assessment of its relevance to the agent's beliefs concerning the context of the decision[20]. If an agent makes a choice that they would regret

[19] Gemes (2007) p. 161.

[20] Hempel (1965) p. 464.

had they taken into account some information that was clearly relevant and easily available, then their reasoning is intuitively irrational. Similarly, it might be possible to rationalise someone's apparently unreasonable choice by considering the entirety of their relevant evidence. Such considerations can be important for historical judgements: it was arguably rational for the Ancient Greeks to sacrifice sheep to the Olympian gods in order to acquire more sheep, because this was likely to be a successful strategy given their total relevant evidence.

Additionally, confirmation theorists and formal epistemologists in general have made a number of very broad generalisations about evidence relevance. For instance, Hempel claims that all relevant evidence either confirms or disconfirms a hypothesis[21]. In contrast, John Maynard Keynes[22] and Janina Hosiasson[23] contend that evidence can be relevant without confirming or disconfirming a hypothesis[24]. Different definitions of evidential relevance imply contrary positions within this debate.

Keynes also argues that the *quantity* of relevant evidence is a pertinent issue within epistemology and normative decision theory. I shall discuss Keynes's concept in the next chapter. Here, the important points are that (1) an analysis of the concept of the *quantity* of relevant evidence can be no clearer than the analysis of relevance itself and (2) one could develop a method of identifying changes in the quantity of relevant evidence via a formal means of certifying a statement as 'relevant'.

[21] E.g. Hempel (1945) p. 3.

[22] Keynes (1921) Chapter VI.

[23] Hosiasson (1931).

[24] E.g. Hempel (1945) p. 3.

This is just a sample of cases in which evidential relevance is a significant part of an interesting controversy or can have useful applications. Clearly, there is much to gain from the study of evidential relevance. I shall now turn to the Bayesian analysis of this important concept.

SECTION 2: THE BAYESIAN ANALYSIS OF EVIDENTIAL RELEVANCE

In this section, I first describe Bayesian confirmation theory, before explaining the Bayesian definition of evidential relevance.

2.1 Bayesianism

2.1.1 Bayesian Probabilities

In contemporary philosophy, Bayesianism is perhaps the most popular theory of probability. I shall use ‘Bayesian’ to describe several different philosophies of probability. All share a common formalization and a common type of interpretation of probability. Bayesians use a probability function P that assigns unconditional probabilities to a domain of statements. The function P must be consistent with the axioms of the probability calculus, which I shall describe in Subsection 2.1.2. Consider a very simple domain that contains three statements H, E, and K. The statement H is a hypothesis; the statement E is some putative evidence for this hypothesis; and the statement K is a conjunction of background knowledge. Bayesians can assign values for the conjunctions of each statement in the domain and use the probability calculus to derive the *marginal probabilities* P(H), P(E), and P(K). In turn, Bayesians can apply the probability calculus to deriving the *joint probabilities* P(H ^ E), P(H ^ K), P(E ^ K), and P(H ^ E ^ K) from the marginal probabilities. Finally, once these values are all determined, it is possible to calculate the *conditional probabilities* like P(H | E ^ K), where these conditional probabilities are defined. The symbol '|' indicates that the probability is the probability of the statement on the left given the truth of the statement on the right. A conditional probability such as P(Φ | Ψ), where Φ and Ψ are statements in the domain, exists when P(Ψ) > 0.

Conditional probabilities play a crucial role in standard Bayesian analyses of relevance. They can be defined by Bayes's Theorem:

Key

P(H): The "prior probability" of H.

P(E): The "expectedness" of E.

P(E | H): The "likelihood" of E given H.

$$P(H \mid E) = \frac{P(H \wedge E)}{P(E)} = \frac{P(H)P(E \mid H)}{P(E)}$$

Bayesians can adopt a variety of axioms of probability that may or may not include the equation above: for example, Colin Howson and Peter Urbach take the definition above as an axiom[25] whereas others take conditional probability to be defined by such an equation,

[25] Howson and Urbach (1993) p. 22.

and it is also possible to regard conditional probability as primitive[26]. The terms in the key above are the prevailing labels within Bayesianism. None of them is entirely intuitive. The "prior probability" is not necessarily determined *a priori*, nor must it be determined prior to the values for P(H), P(H | E), or P(E | H) when calculating a conditional probability. Similarly, the "expectedness" of E could just as accurately be called the "prior probability" of E, since it is just the marginal probability of E. Finally, the "likelihood" has no straightforward relation to the term in ordinary language, because in their non-technical usage, "likelihood" and "probability" are synonymous. The likelihood is simply the conditional probability for E given H. Nonetheless, these are common usages in the literature and I shall stick to them.

Bayesians interpret these probabilities as *epistemic* probabilities. Such probabilities are given interpretations that refer to epistemic concepts like belief, knowledge, rationality, evidence, ignorance, and so on. Epistemic probabilities are also used in descriptive contexts such as psychology and economics. These probabilities are distinct from probabilities that represent non-normative phenomena, such as relative frequencies or propensities.

2.1.2 The Bayesian Formal Framework

In addition to the formal features discussed above, all Bayesians (excluding the Imprecise Bayesians, whom I shall describe in Chapter 2) use a probabilistic framework with the following features in their confirmation theories:

---

[26] For example, Hájek (2003) proposes using conditional probability as primitive.

(1) **The Domain of Probability Functions**: The domain of a Bayesian probability function is a set Ω of statements (or propositions or sentences etc.) that is closed under the connectives of negation and disjunction. This strong form of deductive closure means that if Ω contains the statement 'Fa' and the statement 'Fb', then Ω contains the negated statements '¬Fa' and '¬Fb', as well the disjunction 'Fa v Fb', and all permutations of these operations. Due to the interdefinability of the connectives in propositional logic, this closure under negation and disjunction also entails closure under all other propositional logic truth functions of the standard propositional calculus.

(2) **The Co-Domain of Probability Functions**: The co-domain of Bayesian probability functions is the set of real numbers. (The co-domain of a function is the set of its possible outputs.) For instance, if we compute the joint probability of the statements Φ and Ψ, where Φ and Ψ are in Ω, then we shall have the value $P(\Phi \wedge \Psi) = r$ for some r, where $r \in \mathbb{R}$ and $\mathbb{R}$ is the set of real numbers[27].

(3) **Additivity**: Standard Bayesian probability functions satisfy the axioms of additive probability and all the theorems of these axioms. The axioms are:

(i) $1 \geq P(\Phi) \geq 0$ for all Φ in the domain of P.

(ii) $P(T) = 1$ for any tautology T.

[27] A non-Bayesian probability function P_{nb} can have a different co-domain. The function might only provide numerical inequalities, like $P(\Phi \wedge \Psi) > 0.5$. Alternatively, it might only provide information about a comparative ordering, like $P(\Phi \wedge \Psi) > P(\chi \wedge \xi)$.

(iii) If $P(\Phi \wedge \Psi)$ are mutually exclusive, then $P(\Phi \vee \Psi) = P(\Phi) + P(\Psi)$[28].

From these axioms, one can derive the many useful theorems of the probability calculus.

2.1.3 The Interpretation of Probability

In the sense I am using the term 'Bayesian', all Bayesians endorse the following theses about the interpretation of this formal framework:

(1) **Epistemic Probability**: The probabilities assigned using the function P have epistemic interpretations. There are many Bayesian interpretations, but all of them agree that there is some epistemic significance to probability. The principal Bayesian interpretations are:

(i) *Subjective Bayesians*: This is currently the most popular Bayesian interpretation. They interpret the probabilities as a rational person's degrees of belief in the statements of the domain, where being "rational" involves satisfying a number of constraints. Standard Subjectivist constraints include the axioms of additive probability and conditionalization. Some Subjective Bayesians propose additional constraints. Apart from satisfying these constraints, the degrees of belief are an arbitrary matter for the person who has them. (A subjectivist could also interpret them as the rationally constrained degrees of belief of a

[28] Adapted from Howson and Urbach (1993) p. 21. The principal differences are that Howson and Urbach take the definition of conditional probability above as an axiom rather than a definition and only require that $P(\Phi) \geq 0$; they infer the upper bound for $P(\Phi)$ from an axiom that $P(\Phi) \geq 0$ and Axioms (ii) and (iii) via four proofs in Howson and Urbach (1993) p. 24-25. For reasons of brevity, I simply place the upper bound requirement in Axiom (i), rather than reproduce all four proofs.

computer or a group of people.) Howson and Urbach[29] develop a prominent Subjective Bayesian probability theory that has very few constraints.

(ii) *Objective Bayesianism*: This view is less common than the Subjective interpretation. Objective Bayesians, like Subjective Bayesians, interpret probabilities as rational degrees of belief. However, Objective Bayesians put much stronger constraints on the values that the degrees of belief might take. Objective Bayesians regard degree of belief as either entirely constrained or almost entirely constrained. There is a spectrum of between these views: an extreme Objectivist would contend that the values of degrees of belief are entirely constrained, whereas an extreme Subjectivist would contend that the values are entirely arbitrary. Edwin T. Jaynes[30] and Jon Williamson[31] are prominent Objectivists.

(iii) *Logical Bayesianism*: This logical interpretation is now rare, but it was an important part of the development of Bayesian epistemology. Logical Bayesians regard probability as a logical relation between two statements (or propositions or sentences or sets etc.) that is akin to deductive entailment. (Some logicists, like Keynes, were not Bayesians.) Logical Bayesians combine the logical interpretation of probability with the Bayesian formalism. The paradigmatic Logical Bayesian is Carnap[32] and a more recent example is Patrick Maher[33]. This approach is not always distinguished from Objective Bayesianism[34], but others

---

[29] Howson and Urbach (1993).

[30] Jaynes (2003).

[31] Williamson (2010).

[32] Carnap (1962) is his most extensive development of this interpretation.

[33] Maher (2010) is the apex of Maher's own Logical Bayesian project.

[34] Franklin (2001) uses 'the logical interpretation of probability' and 'objective Bayesianism' as synonyms.

rigorously distinguish logical interpretations of probability from Objective Bayesian interpretations[35].

Bayesians do not need to interpret *all* probability statements as epistemic. They can be pluralists about probability. Pluralists believe that there are multiple legitimate interpretations. Carnap was an early and influential pluralist Bayesian[36].

Most Bayesians also believe that an epistemic probability distribution should be updated by conditionalization:

(2) **Conditionalization**: When a statement E is ‘learned’, in the sense that its probability switches to 1, then one updates every statement by conditionalization to form a new probability distribution P´. Conditionalization of a statement, like H, occurs when the new probability P´ for H is changed such that such that H acquires the probability given by P(H | E), so that P'(H) = P(H | E).

Not all Bayesians adopt conditionalization as a universal norm. For instance, some Bayesians like Richard C. Jeffrey[37] regard it as a special case of Jeffrey Conditionalization. Additionally, some Objective Bayesians like Williamson reject conditionalization under some circumstances[38]. However, in conventional versions of Bayesianism, conditionalization is the

35 Rowbottom (2008).

36 Carnap (1945).

37 Jeffrey (1965).

38 Williamson (2011).

only updating method. Conditionalization is an important part of Bayesian epistemology: if it is removed from the system, then some important standard arguments for Bayesianism (such as the argument that Bayesians will converge towards long-run agreement given some quite general assumptions) are no longer sound.

2.2 Bayesianism and Evidential Relevance

The formalism above provides the basic terms for the standard probabilistic definition of evidential relevance. This analysis was originally developed by Keynes in a non-Bayesian context[39]. It has subsequently become the conventional approach within Bayesianism[40].

Bayesians define evidential relevance as a three-place relation between the hypothesis H, the evidence E, and the available background knowledge K. The definition includes K because differences in the background knowledge can affect the evidential relations. For example, suppose that you are making a random selection from a deck of 52 cards. You are wondering if the selected card is a face card. If your background knowledge contains the information that the deck is a normal deck, then the information that it is a Club is irrelevant to the hypothesis that it is a face card, because the proportion of face Clubs in a normal deck is equal to the proportion of face cards in the deck as a whole. In contrast, imagine if your background knowledge contained the information that the deck is normal *except* that 12 out of the 13 Clubs in the deck are face cards. Relative to this background knowledge, learning

[39] Keynes (1921) p. 55.

[40] Howson and Urbach (1993) p. 117.

that the card is a Club is intuitively relevant to the hypothesis. Consequently, evidence can be relevant or irrelevant depending on the background knowledge.

Most Bayesians adopt the following definition of relevance:

Probabilistic Definition of Evidential Relevance: E is relevant to H relative to K if and only if (1) $P(H \mid E \wedge K) \neq P(H \mid K)$ and (2) $P(E \wedge K) \neq 0$.

(The second clause is needed because $P(H \mid E \wedge K)$ is undefined, under the standard definition, when $P(E \wedge K) = 0$.)

Thus, on the standard Bayesian analysis, E is evidentially relevant to H given K when the probability of H given E and K differs from the probability given K alone. Broadly, they identify evidential relevance with probabilistic relevance, while evidential irrelevance occurs when there is probabilistic irrelevance.

For example, imagine that you have a coin whose bias or fairness is unknown. You assign a prior probability of 0.5 to the hypothesis that the head will land on heads if you toss the coin. Subsequently, a friend tells you that she tossed the coin 20 times, and that it landed heads on 18 out of the 20 tosses. Assume that the conditional probability of the coin landing heads given your friend’s report and your background knowledge is greater than 0.5. According to the standard approach, your friend’s report is evidentially relevant, because it is probabilistically relevant.

Bayesian confirmation theorists also define confirmation and disconfirmation in terms of probabilistic relevance. Confirmation is identified with positive probabilistic relevance. Disconfirmation is identified with negative probabilistic relevance. Thus, in Bayesian confirmation theory, relevant evidence can either confirm the hypothesis or it can disconfirm the hypothesis[41]. Implicitly, Bayesians side with Hempel on the impossibility of neutral evidence (evidence that neither confirms nor disconfirms a hypothesis) in contrast to philosophers like Keynes and Hosiasson.

Bayesianism seems to offer an analysis of evidential relevance that is tremendously general in scope, because it is not restricted to one part of science or one type of evidential reasoning. It also offers the basis for analysing some other interesting concepts. For example, Keynes argued that the quantity of relevant evidence was important for confirmation theory and decision theory[42]. Bayesians can define changes in this quantity in the following way: E adds to the quantity of evidence for H if and only if $P(H \mid E \wedge K) \neq P(H \mid K)$, because this entails that E is relevant to H.

Despite these advantages, the Bayesian analysis of evidential relevance faces a number of problems. I shall now turn to one of the most persistent challenges to this analysis.

[41] Howson and Urbach (1993) p. 117.

[42] Keynes (1921) Chapter VI.

SECTION 3: THE PARADOX OF IDEAL EVIDENCE

Popper's Paradox of Ideal Evidence (PIE) is a criticism of the standard analysis of evidential relevance. He develops the following scenario: suppose that you have a coin, *Z*, when you have no knowledge of the bias or fairness of *Z*. Assume that you assign a probability of 0.5 to the hypothesis H, which is the conjecture that "the *n*th unobserved toss of *Z* will be heads". You subsequently learn E, which is a statistical report that is "ideally favourable" to this assignment of 0.5, such as a report stating that 1,500 out of the 3,000 tosses landed heads[43]. Suppose that the conditional probability of H given your newly acquired evidence is equal to the prior probability, such that P(H | E ^ K) = 0.5. According to the standard probabilist definition of evidential relevance, the report is irrelevant to H, but this is counterintuitive. The standard probabilistic definition seems to be too narrow[44].

Popper assumes that the Bayesian *must* assign P(H) = 0.5. However, there are some theories of epistemic probability, like Subjective Bayesianism, in which it is consistent with Popper's scenario that P(H | K) has any value from 0 to 1. However, the important point for Popper's argument is that P(H | K) = 0.5 is *possible* in any Bayesian probability theory, even though it is not always *mandatory*. Furthermore, Popper could adapt his scenario to any other prior probability via modifying the sample mean: if P(H | K) = *r*, then a sample mean of *r* will be "ideally favourable" to the assignment P(H | K) = *r* under the conditions that he postulates.

I shall also illustrate Popper's basic point using a non-numismatic example: recall the

[43] Popper's argument can also be made in terms of margins of error: if the sample mean is within a margin of error of 0.5, such that P(H | E ^ K) = P(H | K), then there is also an apparent counterexample to the standard probabilist definition of relevance.

[44] Popper (1980) p. 407-408.

team of economists who have been commissioned to study the effects of Turkey joining the EU on the UK's GDP. Imagine that, after a survey of the existing evidence, the team concludes that the available evidence is equivocal to the hypothesis that Turkish membership will increase the GDP of the UK: the probability of the hypothesis given the available background knowledge is 50%. The second part of the team's mandate requires them to acquire some *new* evidence. Imagine that the team gathers a large database of new statistics on UK-Turkish trade, migration figures, capital flows, and other economic information about the direct and indirect economic relations between the two countries. The economists conduct statistical tests using this data and estimate the effects of Turkish membership on UK GDP. They conclude that the total evidence is still equivocal: there is a 50% probability that Turkish membership will increase UK GDP. The standard probabilistic definition of evidential relevance provides the absurd judgement that all of their new data is entirely irrelevant to the hypothesis.

The PIE was subsequently adapted as an objection to probabilistic analyses of the strength of evidence by Achinstein[45]. Kyburg also describes the paradox as a challenge to probabilistic theories of evidence[46]. The PIE is also an example of the criticism that is sometimes made of Bayesianism: it does not supply the means of distinguishing the state of lacking any evidence from the epistemic state of having acquired evidence that is equivocal[47].

[45] Achinstein (1978) p. 29-30.

[46] Kyburg (1970) p. 168.

[47] Reiss (2014) p. 289 and Norton (2011) p. 408-415.

Other philosophers, such as James M. Joyce, use the PIE to motivate a distinction between (1) the quantity of evidence and (2) the balance of the evidence for or against a hypothesis[48].

Popper was making two criticisms with his scenario. He directs both at philosophers such as Keynes and Carnap, who adopted epistemic interpretations of probability. The first criticism is the objection that I have described: the probabilist analyses of evidence says that intuitively relevant evidence is irrelevant. The second objection is a criticism of how Keynes and Carnap represent an agent's beliefs: the degree of belief in the outcome of the *n*th coin toss will not change after learning about the coin tosses[49]. Initially, your degree of belief that the *n*th toss will be heads is 0.5. After acquiring the report, your degree of belief will still be 0.5. Yet Popper suggests that something about your belief in the conjecture has altered, but the probabilistic framework does not formalise this change. He concludes that epistemic probabilities do not adequately represent how beliefs respond to evidence.

The bulk of the literature on the PIE concerns this second aspect of the paradox. The conventional response is (a) to accept that there is no change in a Bayesian reasoner's degree of belief *in the hypothesis*, but also (b) to note that there are *other* statements for which the degree of belief has changed. For instance, even if the probability of the *n*th toss landing heads is unchanged by learning sample report of coin tosses, this information might reduce the expected standard deviation[50] for tossing the coin. Consequently, there is an increase in

[48] Joyce (2005) p. 176.

[49] Popper (1980) p. 408.

[50] The standard deviation of a set of trials (such as coin tosses) is a measure of the dispersion of those results. If all of the results are at the mean, then their standard deviation is a minimum. The standard deviation increases as more of the trials differ from the mean value. Formally, the standard deviation of a sample is $\sqrt{\sum(x-\overline{x})^2 \div (n-1)}$, where x is the value of a member of the sample, $\overline{x}$ is the mean value and n is the sample size.

the probability that all of 10 tosses of the coin will land heads on 5 occasions. Therefore, the Bayesian probabilities can represent a change in one's beliefs after learning the report, even though there is no change in one's belief regarding the *n*th toss in particular. This response was first formulated by Jeffrey[51] and it has been further developed by Bayesians like Howson and Urbach[52].

However, even if this response to Popper is adequate, there is still the problem that he presents for the Bayesian definition of evidential relevance. In this chapter, I shall focus on this first aspect of the PIE. I shall ignore the second aspect. (I shall return to it in Chapter 3 Subsection 6.3.) There have been some responses to this second problem that Popper raised, and I shall now consider them.

SECTION 4: RESPONSES TO THE PARADOX

One obvious response to the PIE (which Popper would recommend) is to abandon probabilistic analyses of evidential relevance. C. A. Hooker and D. C. Stove are examples of philosophers who adopt this position, despite being broadly sympathetic to Bayesianism[53]. However, there are costs to this strategy. Probability is a well-explored concept that is amenable to formal modelling, and accordingly it is a promising basis for an *analysans* of relevance. For that reason, it is worth attempting to maintain a form of probabilistic

---

[51] Jeffrey (1960).

[52] Howson and Urbach (1993) p. 401-403.

[53] Hooker and Stove (1968) p. 310.

definition, in spite of the PIE. I shall focus in this section on responses that retain a probabilistic definition.

One might question the relevance of the coin-tossing report in the PIE, especially if one has a strong intuition that data cannot be both neutral *and* relevant regarding a hypothesis. However, it is plausible that one could “combine evidence” by concatenating the coin-tossing report with similarly neutral data from a physical model of the *Z*’s dynamics. Additionally, imagine that a casino manager asks you to obtain relevant evidence that the *n*th toss will land heads and you know that 0.5 is the accepted value for P(H). Suppose that you have the spare time to carry out 3,000 tosses of *Z* and you report that the coin had landed heads on 1,500 occasions. Intuitively, you have satisfied the request. Popper does not seem to be making any tendentious conceptual claims about relevance in the PIE.

Given that the PIE indicates that the standard probabilistic definition is too narrow, a probabilist response should expand the definition to include the type of scenario that Popper describes, while also avoiding excessive breadth. In this section, I shall critically discuss several probabilist attempts to improve the standard analysis, as well as listing some more radical alternatives. All of them are consistent with a Bayesian interpretation of the probabilities involved. However, I shall argue that none of them is satisfactory.

4.1 Keynes's Strict Definition

Keynes anticipates Popper's problem and proposes a "stricter and more complicated definition" of relevance to address such scenarios[54]. This strict definition contains the standard probabilist definition as a special case. Keynes regards the standard definition as adequate for most situations, whereas his Strict Definition is more robust.

Keynes's Strict Definition of Evidential Relevance: E is relevant to H relative to K if and only if (1) $P(E \wedge K) \neq 0$ and (2) either:

(i) $P(H \mid E \wedge K) \neq P(H \mid K)$

- or -

(ii) $(E \wedge K)$ implies a statement J such that $P(H \mid J \wedge K) \neq P(H \mid K)$.

The intuition behind Keynes's strict definition is that if E has a part that is relevant to H given K, then E is relevant to H given K. For instance, consider a detective who acquires two testimonies regarding a fatal collision between a car and a pedestrian. One witness says that the car was driven erratically prior to the crash, whereas the other says that the car was being driven normally. The detective has no reason to think that the testimony of either witness is more reliable than the other. Assume that the probability (relative to the detective's total knowledge) that the car was being driven erratically is unchanged by learning both of the testimonies. The conjunction of the testimonies seems relevant, but it is irrelevant

[54] Keynes (1921) p. 55.

according to the standard probabilistic definition. In contrast, if we use Keynes's strict definition, we can say that the conjunction of testimonies is relevant, because each individual testimony is entailed by their conjunction, and these individual testimonies are probabilistically relevant to the hypothesis given the detective's background knowledge.

In the PIE, the statistical report E that '1,500 out of 3,000 tosses of this coin landed heads' implies the statement J, that '1,500 out of 1,500 tosses of this coin landed heads', since J describes a subset of the tosses recorded in E. Assume that J can combine with the background knowledge (while ignoring the 1,500 tosses of *Z* that landed tails) such that it is more likely that the *n*th toss will land heads given J and K than given K alone. It follows that J will be relevant to H, because learning J would increase the probability of H given K, and so E is relevant to H on Keynes's strict definition of relevance, because (E ^ K) implies J.

However, Carnap proved a trivialization result for Keynes's strict definition: it entails that every statement is relevant to every other statement, except in special circumstances[55]. If P(H | K) is neither 0 nor 1 and P(E | K) ≠ 1, then E is relevant to H on Keynes's definition. For any statement E, it will be the case that (E ^ K) implies (E v H). Provided that H and E do not have extreme probabilities of 1 or 0 relative to K, the probability of H given (E v H) will be different from the probability of H given K alone.

To see Carnap's point, consider three probabilistically and evidentially independent statements H, E, and K. Assume the following probability distribution:

(1) $P(H \wedge E) = P(H)P(E) = 0.5625$

[55] Carnap (1962) p. 420.

(2) $P(H \wedge \neg E) = P(H)P(\neg E) = 0.1875$

(3) $P(\neg H \wedge \neg E) = P(\neg H)P(\neg E) = 0.0625$

(4) $P(\neg H \wedge E) = P(\neg H)P(E) = 0.1875$

(5) $P(K \wedge H \wedge E) = P(K)P(H)P(E) = 0.5625$

We can derive the marginal probabilities P(H) and P(E) from the joint probabilities above. Since P(H) = P(H ^ E) + P(H ^ ¬E), it follows from (1) and (2) that:

(6) $P(H) = 0.5625 + 0.1875 = 0.75$

Similarly, from (1) and (4):

(7) $P(E) = 0.5625 + 0.1875 = 0.75$

From (5), (6), and (7):

(8) $P(K) = \frac{0.5625}{P(H)P(E)} = 1$

Finally, I shall calculate the conditional probabilities. From (6), (7), and (8):

(9) $P(H \mid K) = \frac{P(H \wedge K)}{P(K)} = 0.75$

(10) $P(E \mid K) = \frac{P(E \wedge K)}{P(K)} = 0.75$

From (1), (7), and (8):

(11) $P(H \mid E \wedge K) = \frac{P(H \wedge E \wedge K)}{P(E \wedge K)} = \frac{P(H \wedge E)}{P(E)} = 0.75$

By symmetry of probabilistic irrelevance, from (10) and (11):

(12) $P(E \mid H \wedge K) = 0.75$

E implies (E v H) by disjunction introduction. Using the above details, axiom (iii) in Subsection 2.1.2 implies that:

(13) $P(E\ v\ H) = P(E) + P(H) - P(H \wedge E)$

Using (6), (7), and (13):

(14) $P(E\ v\ H) = 0.75 + 0.75 - 0.5625 = 0.9375$

From (8) and (14):

(15) $P((E\ v\ H) \wedge K) = 0.9375$

H implies (E v H). Hence, from (8) and (15):

(16) $P((E\ v\ H) \wedge K \mid H) = 1$

Applying Bayes' Theorem and the values from (6), (15), and (16):

(17) $P(H \mid (E \; v \; H) \wedge K) = \frac{P((E \; v \; H) \wedge K) \mid H)P(H)}{P((E \; v \; H) \wedge K)} = 0.8$

Substituting J for (E v H), it follows from (9) and (17) that:

(1) $P(H \mid J \wedge K) \; = \; 0.8 \; \neq \; P(H \mid K)$

Therefore, J is probabilistically relevant to H given K. Since (E ^ K) implies J, it follows from Keynes's strict definition that E is relevant to H. Yet E, H, and K can be almost any statements. Carnap's trivialization proof demonstrates that Keynes's strict definition is far too broad.

There are a number of ways of modifying Keynes's strict definition to address the PIE and Carnap's trivialization proof. One approach is to adopt a different interpretation of 'implies'. Keynes understands 'implies' to mean logical entailment in classical logic:

Classical Logical Entailment: φ logically entails ψ in classical logic if and only if there is no logically possible world in which φ is true and ψ is false.

However, there are a tremendous number of alternative interpretations of entailment in the philosophy of logic. It would be interesting to explore the consequences of adopting different interpretations of 'implies' for Keynes's analysis of relevance, but it would also be an extensive project, with no guarantee of success.

One possible option would be to restrict disjunction introduction, at least within the context of Keynes's strict definition. This rule allows the inference of (φ v ψ) from φ or from ψ, regardless of what propositions that φ or ψ might be. Some philosophers have argued that this principle is the source of many problems, including problems in confirmation theory[56]. However, merely removing this rule is insufficient to avoid Carnap-style trivialization proofs, because these can be formulated using inference rules other than disjunction introduction. For instance, one can use the rule that E implies (¬E → H). Consider the following probability distribution:

(1) $P(H \wedge E) = P(H)P(E) = 1/4$

(2) $P(H \wedge \neg E) = P(H)P(\neg E) = 1/4$

(3) $P(\neg H \wedge E) = P(\neg H)P(E) = 1/4$

(4) $P(\neg H \wedge \neg E) = P(\neg H)P(\neg E) = 1/4$

(5) $P(K \wedge H \wedge E) = P(K)P(H)P(E) = 1/4$

(6) $P(K \wedge H) = P(H) = 1/2$

(7) $P(K \wedge E) = P(E) = 1/2$

(8) $P(K) = 1$

In accordance with Carnap's claim, P(H | K) is neither 0 nor 1 and P(E | K) ≠ 1. Additionally, note that H, E, and K are probabilistically independent.

Suppose that a conditional is true if its antecedent is false. Thus, E implies the conditional (¬E → H). Let 'J' refer to this conditional. J is equivalent to ¬(¬H ^ ¬E) and therefore its probability can be quickly derived from (4):

[56] Weingartner (1994) p. 93.

(9) $P(J \mid K) = 1 - P(\neg H \wedge \neg E \mid K) = 3/4$

J is also equivalent to ((H ^ E) v (¬H ^ E) v (H ^ ¬E))[57]. Therefore:

(10) $P(J \mid H \wedge E \wedge K)) = 1$

(11) $P(J \mid H \wedge \neg E \wedge K)) = 1$

From (1), (9), (10), and (11), the following conditional probabilities can be derived:

(12) $P(H \wedge E \mid J \wedge K) \ = \frac{P(J \mid H \wedge E \wedge K)P(H \wedge E)}{P(J \wedge K)} = \frac{(1)(1/4)}{(3/4)} = 1/3$

(13) $P(H \wedge \neg E \mid J \wedge K) \ = \frac{P(J \mid H \wedge \neg E)P(H \wedge \neg E)}{P(J)} = \frac{(1)(1/4)}{(3/4)} = 1/3$

H is equivalent to ((H ^ E) v (H ^ ¬E)). Therefore, from (10) and (11):

(12) $P(J \mid H \wedge K) = 1$

From (6), (9), and (12):

[57] Note that this is a logical equivalence, not an entailment, so disjunction introduction is not being 'smuggled' in here. The same applies to the equivalence of H and ((H ^ E) v (H ^ ¬E)) later in the proof.

(14) $P(H \mid J \wedge K) = \frac{P(J \wedge K \mid H)P(H)}{P(J \wedge K)} = \frac{(1)(1/2)}{(3/4)} = 2/3$

Finally, from (6) and (8):

(14) $P(H \mid K) = \frac{P(H \wedge K)}{P(K)} = \frac{1/2}{1} = 1/2$

Since P(H | J ^ K) > P(H | K), it follows that J is probabilistically relevant to H relative to K. As E implies J, this proof provides an equivalent result to Carnap's trivialization proof for Keynes's strict definition of relevance, except using (¬E → H) rather than (H v E) as J.

However, one could go further than simply forbidding the use of disjunction introduction: that E implies (¬E → H) is also questionable. Nonetheless, my proof above demonstrates that it is not sufficient to eliminate just *one* rule from Classical Logic in order to avoid Carnap's trivialization proof. More generally, one could investigate the consequences of using non-classical systems in Keynes's definition[58]. The consequences of such a view for formulating a definition of evidential relevance would be an exciting topic for further investigation, but an extended inquiry into the broader consequences of such a maneuverer would be needed before one could be confident that it would not create new paradoxes. Therefore, I shall neither endorse nor reject the views of Weingartner and similar positions, but I shall note that they reveal the range of tactics that a defender of Keynes's strict definition might use. Keynes's strict definition *does* have some strong *prima facie*

[58] For example, Relevance Logicians argue that both the disjunction introduction rule *and* the particular conditional introduction rule that I used in the proof are objectionable, so their systems would be a natural place for a defender of Keynes's strict definition to start.

plausibility, so one cannot rule out that the basic notion is untenable; what one can rule in is that Keynes's strict definition leads to triviality under Classical Logical Entailment or some interesting modifications of this form of entailment.

Another possible alternative, material entailment, will not help Keynes's strict definition.

Material Entailment: φ materially entails ψ if and only if ψ is true or φ is false.

Reinterpreting 'implies' as material entailment will not narrow the scope of Keynes's definition, because all logical entailments are also material entailments. For instance, 'Either William Shakespeare or Christopher Marlowe wrote *Romeo and Juliet*, and Christopher Marlowe did not write *Romeo and Juliet*' logically implies 'William Shakespeare wrote *Romeo and Juliet*', but the first statement also materially entails the second. In contrast, not all material entailments are logical entailments. 'Bertrand Russell met Ludwig Wittgenstein' materially entails 'Bertrand Russell met John Maynard Keynes', because it is not the case that the first statement is true and the second false. However, the first statement does not logically entail the second, since it would be possible that Russell would have met Wittgenstein but not met Keynes. In general, logical entailments are a subset of material entailments. In particular, E always materially implies (E v H), since the latter is a logical implication of the former, and consequently it is also a material implication. It is not possible to obviate Carnap's trivialization proof by reinterpreting 'implies' as material implication.

I shall now examine two modifications of Keynes's strict definition that probabilists have made in response to the PIE and other paradoxes of evidential relevance such as

Carnap's trivialization proof. I shall argue that the first is too broad and the second is too narrow.

4.2 Rod O'Donnell's Response

Rod O'Donnell aims avoid the PIE by adding the condition that data is relevant when it implies a statement J and the negation ¬J is relevant on Keynes's strict definition:

O'Donnell's Strict Definition of Evidential Relevance: E is relevant to H relative to K if and only if $P(E \wedge K) \neq 0$ and at least one the following are true:

(i) $P(H \mid E \wedge K) \neq P(H \mid K)$.

(ii) $(E \wedge K)$ implies a statement J such that $P(H \mid J \wedge K) \neq P(H \mid K)$ and $P(J \mid K) \neq 0$.

(iii) $(E \wedge K)$ implies a statement J such that its negation ¬J is relevant according to the first two conditions[59].

Returning to the PIE, O'Donnell claims that the report of coin tosses E and your background knowledge K jointly imply J_1, where J_1 is 'The expected relative frequency of heads in tosses of this coin is within a margin of error of 0.5.'[60] The negation $\neg J_1$ denies this claim. In O'Donnell's scenario, the following conditional probabilities hold:

[59] O'Donnell (1992) p. 49-50.

[60] O'Donnell (1992) p. 49-50.

(a) $P(H \mid E \wedge K) = P(H \mid K) = 0.5$.

(The conditional probability of H given ($E \wedge K$) is equal to the probability given K, since E states that the coin tosses have a relative frequency equal to the prior probability.)

(b) $P(H \mid J_1 \wedge K) = P(H \mid K) = 0.5$.

(The conditional probability of H given ($J_1 \wedge K$) is equal to the probability given K, since J_1 states that the expected relative frequency of head tosses is within the margin of error of the prior probability.)

(c) $P(H \mid \neg J_1 \wedge K) = P(H \mid K) \neq 0.5$.

(The conditional probability of H given ($E \wedge K$) is *not* equal to the probability given K, since $\neg J_1$ states that the relative frequency of the coin tosses is not within the margin of error of 0.5.)

Since $\neg J_1$ is relevant to H given K, it follows from O'Donnell's definition that J_1 is relevant to H given K. Since J_1 is relevant to H given K and ($E \wedge K$) implies J_1, it follows that E is relevant to H given K. In this way, O'Donnell's definition provides the intuitively correct judgement that the report of coin tosses is relevant in Popper's scenario.

The basic notion that a statement is relevant if its negation is relevant is plausible. O'Donnell locates some historical precedent for this idea in discussions of evidential relevance by Keynes[61] and Stove[62], whose uses of this condition is motivated independently of the PIE. Thus, if it were satisfactory, his definition would have the advantage of being a non-*ad hoc* response to the PIE.

However, O'Donnell's discussion of the PIE has two important problems. Firstly, O'Donnell seems to misdiagnose Keynes's strict definition. E implies that there were 1,500 coin tosses that landed heads, and this implied statement is straightforwardly relevant to H given K according to Keynes's strict definition. O'Donnell's might be intrinsically intuitive, but it is superfluous for determining that E is evidentially relevant.

Secondly, O'Donnell's definition is no narrower than Keynes's strict definition: anything that is relevant on Keynes's definition will be relevant on O'Donnell's strict definition, because O'Donnell does not add any necessary conditions to Keynes's definition. The only modification that O'Donnell makes is an additional sufficient condition for relevance. Consequently, his definition is still subject to Carnap's trivialization proof, and hence unacceptably broad as an analysis of evidential relevance. O'Donnell does not discuss Carnap's trivialization proof, so naturally he does not suggest a response. Like Keynes, he has not managed to navigate between the Scylla of Popper's PIE and the Charybdis of Carnap's trivialization proof.

---

[61] Keynes (1921) p. 121.

[62] Stove (1986) p. 82.

4.3 Gemes's Analysis of Relevance

Gemes aims to retain the basic intuition in Keynes's strict definition: if a part of E is relevant to H, then E is relevant to H, but he modifies the analysis of propositional content (being a "part") to avoid Carnap's trivialization proof and other counterintuitive features of the analysis of content used by Keynes[63].

In Keynes's strict definition, J is a part of the content of E if and only if E implies J. I shall use the term 'Implication Analysis' to name this analysis of logical content. It can lead to very strange results. For instance:

Key

H: The orbits of the planets are elliptical.

E: The Sun is at a focal point of the planets' orbits.

J: The orbits of the planets are not elliptical.

Consider the conjunction:

(1) H ^ E.

By conjunction elimination, (1) implies:

[63] Gemes (2007) p. 163-164.

(2) E.

(2) implies, by disjunction introduction:

(3) E v J.

However, by the same rule, (3) can also be derived from:

(4) J.

J intuitively confirms (E v J), but it would be very peculiar to say that 'J confirms a part of (H ^ E)', because J is inconsistent with H and J is irrelevant to E. Yet, according to the Implication Analysis, (3) is a part of (1). Assuming that (4) confirms (3), we must say that (4) confirms a part of (1) if we are using the Implication Analysis of propositional content.

Unlike philosophers such as Weingartner, Gemes does not object to disjunction introduction: he is arguing that while (ϕ v ψ) is a *consequence* of any proposition ϕ, it is not thereby a part of the *propositional content* of ψ[64]. Gemes develops an alternative analysis that has potential benefits for defenders of classical logic: by rejecting the Implication Analysis, they can retain the classic analysis of logical consequence, while nonetheless avoiding any paradoxical results of using that analysis for logical content. However, even if Gemes's analysis of logical content is satisfactory, there might still be reasons (such as those raised by Weingartner) for rejecting features of the classic analysis of logical consequence, such as

[64] Gemes (2007) p. 164.

paradoxes where disjunction introduction or *ex falso quodlibet* seem to be responsible.

Using his alternative analysis of propositional content that avoids these problems, Gemes develops a revised version of Keynes's strict definition of relevance:

Gemes's Definition of Relevance: E is relevant to H given K if and only if

(1) $P(E \wedge K) \neq 0$ and (2) there is a part J of the propositional content of E such that

$P(H \mid J \wedge K) \neq P(H \mid K)$[65].

(Gemes does not have an explicit reference to background knowledge K in his definition. I have added it to preserve consistency with other definitions.)

I shall not discuss Gemes's account of propositional content or how it avoids both the PIE and Carnap's trivialization proof, because his account is highly complex and the details are unnecessary for my criticism. Gemes's definition faces a problem that I shall call the 'Problem of Corroborating Evidence'[66]. For example:

[65] Gemes (2007) p. 165.

[66] I am using 'corroborate' in a non-technical sense, rather than in Popper's sense of the term.

Key

H: All swans are white.

E: Bob is a swan and Bob is not white.

K: Our background knowledge, which includes the information that Arnold is a swan and Arnold is not white.

Intuitively, E is relevant to H given K, because a counterexample is relevant to a universal generalisation, even if you know that there is another counterexample. However, obviously $P(H \mid K) = 0$ and $P(H \mid E \wedge K) = 0$. Thus, according to Gemes's definition, E is not directly evidentially relevant to H given K. Furthermore, there can be no J such that (a) J is part of the logical content of E and (b) J is relevant to H given K, because J cannot raise H's probability given K and zero is the lowest possible value of the function P. In general, the probability calculus implies that additions to the conditions in an extreme-valued probability do not alter the probability.

For example, suppose that H is inconsistent with the background knowledge and $P(J \wedge K) > 0$. Consider the Bayesian probability distribution in which:

(1) $P(H \wedge J \wedge K) = 0$

(2) $P(H \wedge K) = 0$

(3) $P(J \wedge K) = 0.5$

Applying Bayes' Theorem to derive the conditional probability of H given J and K from (1) and (3):

$$P(H \mid J \wedge K) = \frac{P(H \wedge J \wedge K)}{P(J \wedge K)} = \frac{0}{0.5} = 0$$

Gemes's definition has the consequence that there is no evidence that can corroborate the falsification of a hypothesis, yet this is counterintuitive. Even knowing that H is false, it is intuitively possible to obtain corroborating evidence that the hypothesis is false.

One might question whether corroborating counterexamples are actually relevant evidence, but there are several reasons to see them as evidence. Firstly, imagine that the original falsifying counterexamples are withdrawn from the body of scientific evidence, but the corroborating statements are retained. For example, we might discover that the birds that were the initial apparent counterexamples to 'All swans are white' were another species of bird. However, provided that the corroborating counterexamples are retained, then the hypothesis will still be inconsistent with our total evidence. In such scenarios, it is natural to say that 'The hypothesis is inconsistent with what remains of our evidence', rather than that 'The counterexamples were never part of the evidence, but entered into the total evidence once the original observations were retracted.'

Secondly, though the corroborating counterexamples are probabilistically irrelevant to the hypothesis, it is natural to say that they indirectly provide support to the negation of the universal generalisation, because the initial counterexamples are more reliable given the corroborations. For instance, even if one accepts counterexamples to 'Every chemical element has a uniform melting temperature under laboratory conditions', it is still possible to increase the reliability of the initial counterexamples by replicating the falsifying experiments involving allotropes. Such corroboration of falsifications is one of the contributions that experimental evidence often provides to scientific knowledge.

Thirdly, it would be unfair to say that a scientist who has gathered observations of black swans has not gathered any evidence regarding ‘All swans are white’, assuming that her project managers asked her to carry out such an investigation. What she has done might be uninteresting, but it nonetheless constitutes gathering relevant evidence. There is an important distinction between ‘interesting evidence’ and ‘relevant evidence’. Corroborating counterexamples can be instances of the latter without being instances of the former.

4.4. Further Alternatives

There are other responses to the PIE that I have not discussed here. Peter Gärdenfors’s abandons the standard Bayesian framework and combines a modified definition of relevance with a nonstandard model of belief revision[67]. Ronald N. Giere briefly approaches the problem from a Neyman-Pearson version of classical statistics[68].

Others, such as Maya Bar Hillel, analyse the problem in terms of second-order probabilities, which are degrees of confidence in first-order probability assignments such as the assignment of 0.5 to ‘The *n*th toss will land heads’ in Popper’s scenario[69]. I am quite

[67] Gärdenfors (1990).

[68] Giere (1970) p. 357.

[69] Bar Hillel (1982).

sympathetic to the use of probabilities with multiple orderings (as in Chapter 3) but often they can solve a particular formulation of a problem only by creating a regress, and such an approach needs to answer versions of the PIE that use the second-order probability assignments. There is also the challenge of formulating an answer that would also address the Problem of Corroborating Intervals. While I do not have a strong argument against this approach to the PIE, I shall not pursue it further, for the reasons adumbrated.

Section Summary

Even with clever modifications, the standard probabilistic analysis of evidential relevance is unsatisfactory. However, this does not prove that a probabilistic analysis is impossible. In the next two sections, I shall present an alternative to Bayesianism and use this rival to develop a new definition.

SECTION 5: EVIDENTIAL PROBABILISM

In this section, I shall introduce Kyburg's theory of Evidential Probability. This will be the basis of my answer to the PIE in Section 6.

5.1 Basic Principles of Evidential Probability

Kyburg developed his theory over 40 years, from 1961 to 2001[70]. Throughout this evolution, two fundamental ideas remained constant: (1) epistemic probability should be

[70] From Kyburg (1961) to Kyburg and Teng (2001).

modelled as a formal evidential relation between sentences and (2) every probability relation must be based on information about relative frequencies[71].

Like Bayesianism, Kyburg's theory of Evidential Probability is an epistemic theory of probability. (Specifically, it is a logicist interpretation.) The fundamental difference is that there is no prior distribution; instead, Evidential Probabilities are derived by *direct inference* from relative frequency data[72]. Direct inference is also sometimes called "the statistical syllogism" or "the proportional syllogism". Such reasoning begins from (a) a premise about the relative frequency of some predicate in a population and (b) a premise that an individual or sample is a member of that population, and infers to (c) the assertion or denial that the individual or sample satisfies that predicate. For example:

(1) The proportion of red balls in the box is 95/100.

(2) Ball A is a ball in the box.

Therefore, probably, (3) Ball A is red[73].

There are a number of important points about direct inferences. Firstly, direct

---

[71] Kyburg (1990) p. 43.

[72] Kyburg and Teng (2001) p. 201.

[73] The positioning of 'probably' is important: the argument is a non-deductive inference of (3), rather than the ascription of a probability to (3), so 'probably' characterises the argument. The conclusion of this argument is not a probabilistic statement. This is importantly distinct from a similar form of inference, in which (3) would be a probabilistic statement, i.e. 'Ball A is *probably* red' or 'The probability that A is red > 0.5' etc. Put another way, 'probably' qualifies 'therefore', which in turn is the relation of the premises to the conclusion; 'probably' is *not* a qualifier in the conclusion itself. For further discussion of these aspects of direct inference, see Hempel (1960) p. 444-447.

inferences are (typically) deductively invalid. For example, it is consistent with the truth of (1) and (2) that Ball A is one of the 5% of balls in the box that are not red. Direct inferences will only be deductively valid when either:

(i) The premises assert that 100% of the population have the predicate in question and the conclusion asserts that the subject of the direct inference has that predicate.

- or -

(ii) The premises assert that 0% of the population have the predicate in question and the conclusion *denies* that the subject of the direct inference has that predicate.

When either (i) or (ii) holds, a direct inference is simply a type of deductive syllogism. Indeed, one can understand the logic of direct inferences as a generalisation of syllogistic logic to arguments that feature quantifiers like ‘Most’ or ‘15%’ rather than ‘All’ or ‘Some’[74].

Secondly, when (i) and (ii) do not hold, the argument is non-monotonic. In a non-monotonic argument, additions to the premises can alter the strength of the argument. In the example of the box of balls, one could add the premise that ‘Ball A is near the top of the box and nearly all of the balls near the top of the box are green’. Intuitively, the premises would no longer provide support for the conclusion.

In practice, our use of direct inference is usually so automatic and confident that it is

[74] Williams (1947) p. 36.

unreflective. Our everyday facility with direct inferences seems to have led many philosophers and logicians to overlook non-monotonic direct inferences, in comparison to more controversial forms of non-monotonic reasoning like enumerative induction and analogy.

5.2 Imprecise Probabilities

Before I describe Kyburg's theory of direct inference, there is a crucial part of Evidential Probability that I must explain. Kyburg incorporates imprecise probabilities into his system because our statistical information is typically imprecise: I might know that between 50% and 75% of the balls in a bag are green, but not know the exact proportion. Imprecise statistical information is even more familiar when we are not considering gambling apparatuses. For instance, I am confident that the average UK household owns at least one car, but I do not know the precise figure. Similarly, when discussing the probability of scientific hypotheses like 'The Solar System has more planets than have been discovered', it is more natural to use imprecise background data like 'Star Systems with the astrophysical characteristics of the Solar System generally have more planets than have been discovered' as opposed to point-valued figures.

Kyburg uses closed intervals within the co-domain of fractions from 0 to 1 inclusive to formalise imprecise statistical information and probabilities. A closed interval $[x, y]$ is an interval that includes x and y and everything in between. An open interval (x, y) includes everything between x and y, but excludes these values. For example, the closed interval of integers [0, 5] contains 0, 1, 2, 3, 4, and 5, whereas the open interval of integers (0, 5) contains 1, 2, 3, and 4.

In my notation for Evidential Probability, I shall deviate in some respects from Kyburg's formalism, because I want to keep as close as possible to standard formalisation of Bayesian probabilities, in order to make my discussion easily accessible for readers who are used to Bayesian notation. I shall use 'EP' for the Evidential Probability function. I shall use 'K' to represent background knowledge. (EP always takes K as one term in the conditional probability of a statement.) For the probability of H given K alone, I shall use EP(H | K) = [*x*, *y*], where *x* and *y* are fractions or inequalities. When precise limits are inappropriate, I shall use inequalities for *x* and/or *y*, like [> 0.5, < 1], instead of open intervals like (0.5, 1), because this approach allows the easy formalisation of expressions like 'the probability of H is at least 50%' as closed intervals like [> 0.5, 1]. An agent's background knowledge can be conjoined to their evidence E to determine the probability of a hypothesis H given that E has been learned. I shall represent such probabilities using the notation EP(H | E ^ K) = [*x*, *y*]. When the relative frequency information is rich enough to supply an exact evidential probability, such that *x* = *y*, then an Evidential Probabilist formalises such a probability as a degenerate interval[75] like EP(H | K) = [0.5, 0.5], which corresponds to a precise probability of 0.5.

For example, if K represents the knowledge that 65-75% of the balls in the bag are green and H represents the hypothesis that a randomly selected ball from that bag will be green, then this evidential relation is represented as EP(H | K) = [0.65, 0.75]. This formalises the claim that the probability of H given E is 65% to 75%.

Generally, one could remove K from the notation. However, sometimes it is important to consider different real or hypothetical bodies of background knowledge, and this can be formalised by distinguishing between K_1, K_2, K_3… Additionally, the explicit reference to

[75] A degenerate interval is any interval in which the limits are equal.

background knowledge makes it clear that evidential probabilities are always relational: there are no unconditional Evidential Probability values. By contrast, in standard Bayesianism, there are marginal probabilities like P(H) and P(E) that can be derived from the full distribution over the domain. The reference to K highlights this important difference between the two systems.

Returning to my illustration of direct inference:

(1) The proportion of red balls in the box is 95/100.

(2) Ball A is a ball in the box.

Therefore, probably, (3) Ball A is red.

Suppose one is constructing an Evidential Probability model of an agent's reasoning. Assume that (1) and (2) are her relevant background knowledge K. Let H represent the conclusion (3). On the assumption that the information in the premises is the agent's best data for the probability of H, then the function EP will assign the value [0.95, 0.95] as the value for EP(H | K)[76].

[76] To clarify, Kyburg is *not* saying that we can infer the evidential probability from the premises; instead, the information from the premises (along with any relevant background knowledge) is part of what determines the probabilistic relation *between* the premises and the conclusion. In other words, as in the original example, it is not the case that a probabilistic statement is being inferred; instead, Kyburg is proposing that the value of a conditional probability that relates two sets of statements should be calculated using the relative frequency data in the premises, provided that this is the best available evidence with respect to the conclusion. Thus, 'EP(H | K) = [0.95, 0.95]' is a statement that asserts a (non-deductive) logical relation *between* H and K, rather than a contingent relative frequency statement. Put another way, 'EP(H | K) = [0.95, 0.95]' describes 'therefore' in the argument; it is not a probabilistic statement that is inferred by the argument itself.

Intervals are Kyburg's own preferred means of representing imprecise probabilities, but there are many alternatives. Each approach has advantages in some contexts and disadvantages in others[77]. A pluralistic attitude is consistent with the basic principles of Evidential Probability, because the central objective for an Evidential Probabilist is the best possible representation of the available statistical information; plausibly, different formalisations are optimal for this objective in different contexts. Since Kyburg uses an interval-valued formalism and this formalism is generally satisfactory, I shall discuss Evidential Probability as an interval-valued system. I shall use 'imprecise' as a synonym of 'interval-valued'.

One interesting facet of Kyburg's system is that the probability that is defined for any statement Φ given any set of statements Γ, *even if* Γ contains no statistical information regarding Φ. For example, consider the hypothesis that the average annual rainfall will be greater in Borneo than in the Amazon in the year 250,000 AD. Such conjectures are cases of what Frank H. Knight called "uncertainty", for which there can be no precise probabilities that are determined by the available evidence. Knight contrasts such conjectures with cases of "risks" that *can* be measured, because there are known reasons to assign precise probabilities[78]. Someone who made a decision regarding the rainfall hypothesis would be making an 'uncertain' decision, in Knight's sense[79]. According to Evidential Probabilists, a precise probability distribution for such conjectures is inappropriate. However, even if *no* degree of precision (even a wide interval like [0.1, 0.9]) seems appropriate, an Evidential Probabilist still assign the maximally imprecise [0, 1] interval. (I shall explain this aspect of

[77] Walley (2000) provides a detailed critical survey.

[78] Knight (2006) p. 233.

[79] Knight (2006) p. 226.

Evidential Probability in Subsection 5.3) Even though the values must always be based on relative frequency data, an Evidential Probabilist can always assign an evidential probability, even when the evidence and background knowledge provides no relevant statistical information.

Despite the references to "relative frequencies" in Evidential Probability, the probabilities are not statements about long-run or short-run frequencies. Instead, they are normative claims about the relationships between statements[80]. The basic form of evidential probability is the assignment of a value to a single-case. The evidential probabilities of more general hypotheses are constructed from this foundation. If a hypothesis is a conjunction of different claims (like 'This unknown compound is highly radioactive, poisonous and conducts electricity') then its probability is determined by calculating the probability of the conjunction of those claims. If a hypothesis makes reference to a generic member of a class (like 'A man living in Bearsden has a life expectancy of 85 years') then its probability is the probability for an individual in that class for which it is only known that the individual is a member of this class. Consequently, there is no problem with single-case probabilities in Kyburg's system: these probability statements not only occur, but they are the bedrock of evidential probabilities.

To close this section, I shall state the general axioms for the EP function[81]. The domain of the probability function is a set of statements Ω that is weakly deductively closed and weakly consistent:

---

[80] One can also view them as purely about model-theoretic relationships between statements: see Kyburg and Teng (2001) Chapter 10. However, this does not eliminate the importance of normativity at some point in the theory, since one always needs bridge statements between facts about logical models and normative principles if the former are to have significance for methodology or the theory of rationality.

[81] Kyburg (1990) p. 49.

Weak Deductive Closure: If Ω contains Φ and Φ implies Ψ, then Ω contains Ψ.

Weak Consistency: There are no statements in Ω that are internally contradictory, like '$\Phi \wedge \neg\Phi$'.

The following axioms hold for EP:

(1) If $EP(\Phi \leftrightarrow \Psi \mid K) = [1, 1]$, then $EP(\Phi \mid K) = EP(\Psi \mid K)$.

(If K implies that statements have the same truth-value, then they have the same probabilities given K.)

(2) If $EP(\Phi \mid K) = [x, y]$ then $EP(\neg\Phi \mid K) = [1- y, 1- x]$

(Statements' probabilities vary in proportion to their contradictories' probabilities.)

(3) If $EP(\Phi \mid K) = [x, y]$ and $EP(\Psi \mid K) = [z, v]$ and K contains $\neg(\Phi \wedge \Psi)$, then $EP(\Phi \vee \Psi \mid K)$ is equal to the interval $[x + v, y + z]$ or the interval for $EP(\Phi \vee \Psi \mid K)$ is contained in $[x + z, y + v]$.

(This axiom adapts the axiom that P(A v B) = P(A) + P(B) when A and B are mutually exclusive to the Evidential Probability system. One can derive that there is always at least one additive measure function that is consistent with the constraints provided by the intervals.)

(4) If EP(Φ | K) = $[x, y]$ and EP(Ψ | K) = $[z, v]$ and Φ implies Ψ, then $z \geq x$.

(Statements cannot be less probable than the statements that imply them.)

(5) If T is tautologously true in Ω, then EP(T | K) = [1, 1].

(Statements that are necessarily true in Ω have a maximal probability.)

(6) If K contains no explicit or implicit contradictions and K contains Φ, then EP(Φ | K) = [1, 1].

(A statement has maximal probability given a body of statements that has no contradictions and that includes itself.)

From these axioms, Kyburg derives the theorem:

(7) For any finite set of statements, there is an additive probability function P such that its values are within the intervals supplied by Evidential Probability.

This theorem is important for understanding the relationship between Evidential Probability and Bayesianism. For instance, it means that one could use Evidential Probabilities as constraints on the choice of a Bayesian probability function; Jon Williamson and Gregory Wheeler take this path in their version of Objective Bayesianism[82].

---

[82] Wheeler and Williamson (2011) p. 327-329.

5.3 Reference Class Selection

Direct inference is an attractive basis for epistemic probabilities, because it seems very intuitive and uncontroversial, in comparison to both (a) other forms of ampliative inference or (b) in comparison to the choice of a prior distribution. However, direct inference has its own puzzles. One of the most studied issues is the Problem of the Reference Class, most influentially discussed by Hans Reichenbach[83]. Since an individual or sample will usually belong to many different reference classes (populations) there is a problem of selecting the appropriate reference class from among these rivals. For example, suppose that I am choosing whether to purchase an insurance policy for my laptop. I am estimating how likely it is that the laptop will last for more than four years. Should the evidential probability be an estimate of the relative frequency of laptops in general? Or this particular brand of laptop? Or laptops in general when used in my climate? The Problem of the Reference Class is the search for a systematic procedure for selecting a reference class from such a set of rivals.

To a large extent, the history of the development of Evidential Probability was a history of Kyburg's wrestling with the Problem of the Reference Class. While this selection process is usually intuitive and uncontroversial, Kyburg seeks to formalise our intuitions. The essence of his answer is simple: in direct inference, we should use the statistical information that provides the most information about the conclusion[84]. Taking this notion as his starting point, Kyburg develops a set of rules for reference class selection.

---

[83] Reichenbach (1949) p. 375.

[84] Kyburg and Teng (2001) p. 212.

All Evidential Probabilities are derived from single-case probabilities, so I shall begin with single-case statements. Suppose H is an assertion that a particular object has a particular predicate and that K contains a set of statistical statements {R*a*, R*b*, R*c* … R*n*} for which (i) K contains the statement that the object is a member of each of the reference classes described by each of these statements and (ii) each statement makes a claim about the relative frequency of the predicate in question in that reference class. I shall call these 'reference class statements'. For simplicity, I shall exclude statements that might be used (redundantly) in determining the evidential probabilities of statements that are known to be contradictory or tautologous from the extension of 'reference class statements'.

For example, consider the hypothesis H_1: 'My first toss of this £2 coin will land heads.' My background knowledge K states that the coin toss is a member of a variety of reference classes for which I have some statistical knowledge: my tosses of this coin, tosses of £2 coins, tosses of pound coins, tosses of contemporary coins, coin tosses in general, and so on. My estimates of the relative frequencies of landing heads in each of these coin tosses form a set of reference class statements. One can derive the evidential probability of the hypothesis given my knowledge by applying Kyburg's rules of Sharpening, which I describe below. Thus, I might eliminate all the reference class statements except 'Between 48% and 52% of £2 coin tosses land heads', and consequently deduce that $EP(H_1 \mid K) = [0.48, 0.52]$.

The selection of an evidential probability for H given K involves applying the following rules of Sharpening[85] in sequential order:

[85] Kyburg and Teng (2001) p. 218-219.

1. **Sharpening by Richness**: If two statements regarding reference classes, R*a* and R*b* give incompatible frequencies, and R*a* is statement about a joint distribution of random variables and R*b* is a statement about a marginal distribution of a random variable, then ignore R*b*[86]. At a broader level, if there a known full distribution for a set of random variables and this full distribution conflicts with a marginal distribution, then the information from the marginal distribution will be ignored in favour of the full distribution. Kyburg describes this rule as allowing Bayesian reasoning within Evidential Probability, provided that this reasoning is based on known relative frequencies[87].

Example

Imagine that I am making a selection from one of two piles of cards, Pile I and Pile II. I shall determine my choice of pile by tossing a fair coin. I know that there are 20 cards. I also know that 10 of the 20 cards are black.

Suppose I also know that Pile I has 16 cards, 10 of which are black, whereas Pile II has 4 cards, none of which are black. Assume that I know that the joint distribution of (i) selecting Pile I by using the coin *and* (ii) selecting a black card from Pile I is (1/2)(10/16) = 0.3125. Let R*a* be the reference class statement derived using my knowledge of this joint distribution: 'Black cards will be selected from the two piles with a relative frequency of

---

[86] A random variable is a function that takes statements (or propositions or sets etc.) as its domain and real values as its co-domain, or vice versa. A joint distribution for a random variable is a probability distribution that provides probabilities that each of a set of random variables gives a value in a particular interval or has a particular value. A marginal distribution for a random variable provides the probability that a specific random variable or subset of random variables is in a particular interval or has a particular value, without taking into account the values of the other random variables in the set.

[87] Kyburg and Teng (2001) p. 217.

0.3125.’ Let R*b* be the reference class statement that ‘Black cards will be selected from the two piles with a relative frequency of 0.5.’ Sharpening by Richness requires that I ignore R*b* in favour of R*a*.

2. <u>**Sharpening by Specificity**</u>: Having applied the previous rule to the set, apply the following rule: if R*a* and R*b* state incompatible frequencies and it is known that all the members of the reference class described by R*a* are members of the reference class described by R*b*, but not vice versa (so the members of R*a* are the elements of a proper subset of the set of members of R*b*) then ignore R*b*. Put simply and informally, more specific reference classes are favoured over less specific reference classes.

<u>Example</u>

Imagine that I know R*a*: ‘Lions attack their handlers in less than 1% of encounters.’ I also know R*b*: ‘Very hungry lions attack their handlers in over 75% of encounters.’ Suppose I am handling a lion and I know that it is very hungry. I know that hungry lions are a proper subset of lions. Sharpening by Specificity requires that I ignore R*a* in favour of R*b*.

If either of the first two rules has reduced the set {R*a*, R*b*, R*c* … R*n*} to a single reference class statement R*i*, then the relative frequency interval asserted in R*i* is the value of EP(H | K). If there is still a multi-member set of statements asserting different intervals, then one applies Sharpening by Precision:

3. **<u>Sharpening by Precision</u>**: If there is a statement R*a* whose relative frequency interval is a proper subinterval of the other statement's relative frequency intervals, then R*a* determines the Evidential Probability[88]. If there is no such statement, then one must use the cover of the remaining intervals: the limits of the interval are the lowest fraction and the highest fraction in the set of reference class statements.

<u>Examples</u>

As an example of clause (i), suppose that I am considering the probability that a particular respondent to a poll chose to vote at the last UK general election. I know that (a) the respondent is in full-time employment, and (b) they have a university degree. Suppose I know R*b*: 'Between 50% and 66% of people in full-time employment voted at the last general election', but I also know R*a*: 'Between 57% and 63% of university graduates voted.' Sharpening by Precision in this case requires that I select the interval [0.57, 0.63] as the evidential probability, because this information is consistent with what R*b* asserts, but it is also more precise.

As an example of clause (ii), consider the previous example, except that R*b* states a relative frequency of between 50% and 60%, while R*a* states a relative frequency of between 65% and 70%. Sharpening by Precision requires that I take the cover of the intervals: [0.5, 0.7], because this is the most precise interval that includes both of the intervals provided by R*a* and R*b*.

[88] I.e. the first two rules have created a situation in which, for any two members of this set, either they give the same interval or one is a proper subinterval of the other, such { [0, 1], [0.6, 0.9], [0.75, 0.8], [0.75, 0.75] }. By contrast, there is no proper subinterval of both [0.75, 0.75] and [0.8, 0.8].

As a further example of (ii), suppose there is unseen wooden carving in a black box. I have reached into the box and felt the carving. I know that (a) it has a rough texture, (b) it has a cubic shape and (c) it is near the top of the pile of carvings in the black box. I am wondering whether the carving is brown. The remaining reference class statements are (1) R*a*: 'Between 10% and 50% of cubic carvings are brown' (2) R*b*: 'Between 50% and 70% of the carvings with a rough texture are brown' and (3) R*b*: 'Between 70% and 90% of the carvings near the top of the box are brown.' Sharpening by Precision requires taking the cover of the intervals: [0.1, 0.9].

Followed sequentially, these rules determine a unique Evidential Probability for any statement Φ and any set of statements Γ. This uniqueness is a result of the fact that any set of reference classes that is not eliminated by Sharpening by Richness or Sharpening by Specificity will provide a single interval via Sharpening by Precision, which will require selecting the shortest interval that is consistent with all the surviving rival intervals. The resulting interval might be wide, but it is always unique for a particular hypothesis and particular total evidence.

This uniqueness obtains even if Γ provides no relative frequency data about Φ. For any statement ascribing a predicate F to a particular object *a*, there is the reference class U that is the unit set[89] containing only *a*. Either Fa or ¬Fa. If Fa, then 100% of objects in U are F. If ¬Fa, then 0% of objects in U are F. Consequently, one always knows the reference class statement R*u*: 'The relative frequency of objects that are F in U is between 0% and 100%' when considering the Evidential Probability that *a* is F. If any more precise information is available, then the statement R*u* will always be eliminated via Sharpening by Precision.

[89] A unit set is a set with a single member.

However, if more precise information is unavailable, then R*u* will remain, and Sharpening by Precision will require using Ru, so that the interval is the maximally wide [0, 1]. Accordingly, even in the special case where no precise interval can be assigned, one can assign the interval [0, 1].

For the application of Evidential Probability, Kyburg and Teng also restrict the use of direct inference to a set of reference classes that is closed under conjunction, but not closed under disjunction[90]. For instance, 'chemical element' and 'laboratory sample' might both be names of possible reference classes, but not 'chemical element or laboratory sample', unless the corresponding class is separately distinguished as a reference class. The choice of reference classes is decided on grounds that are exogenous to the formal model of Evidential Probability.

Plural-case probabilities, like the probability of 'The next ten tosses of this £2 coin will land heads' are derived from these single-case probabilities using the axioms that I described at the end of Subsection 5.2. The joint probabilities are calculated from the marginal probabilities, except that (if the intervals are non-degenerate) one must calculate two values[91]. For example, suppose that I know that my tosses of the £2 coin are independent. Let ($H_1 \wedge H_2 \wedge H_3 \wedge H_4 \wedge H_5$) be the hypothesis that 5 tosses of the coin will land heads. The lower limit of the evidential probability interval is the joint product of the lower limits for each of the hypotheses:

[90] Kyburg and Teng (2001) p. 204-207.

[91] Kyburg and Teng (2001) p. 263.

(4) $(0.48)(0.48)(0.48)(0.48)(0.48) = (0.48)^5 = 0.01$ (2 d. p.)

One calculates the upper limit in the same manner:

(5) $(0.52)(0.52)(0.52)(0.52)(0.52) = (0.52)^5 = 0.02$ (2 d. p.)

From (4) and (5):

(6) $EP(H_1 \wedge H_2 \wedge H_3 \wedge H_4 \wedge H_5 \mid K) = [0.01, 0.02]$ (2 d. p.)

In the special case of degenerate intervals, such a calculation is easier, since it will be simply a case of calculating the joint probabilities given the marginals, given the known logical and probabilistic relations between the statements about single-cases. For an idealized set-up, like a model of a gambling apparatus or an idealized model of a physical system with precise relative frequency claims about events in that system, one can use the standard probability calculus to determine the evidential probabilities. For instance, in a standard idealized model of tossing a die, one can calculate the evidential probability of *not* rolling a six in 10 tosses using the marginal probabilities for the individual tosses:

$$(5/6)^{10} = 0.16 \text{ (2 d. p.)}$$

Hence, using HNS for the hypothesis that die will not land on six is the following degenerate interval:

$$EP(HNS \mid K) = [0.16, 0.16] \text{ (2 d. p.)}$$

Thus, in the special case where the relevant background knowledge is sufficiently rich (perhaps because the hypothesis in question concerns an idealized model) the calculation of plural-case probabilities is identical to the calculation of a joint probability given the relevant marginals using the standard probability calculus. When the relevant background knowledge is sufficiently rich, there can be a high degree of similarity between Evidential Probability and Bayesianism.

The degree of similarity between Evidential Probability and Bayesianism in dynamic reasoning is also dependent on the richness of the background knowledge. Updating probabilities using Evidential Probability will sometimes, but not always, correspond to updating using Bayesian conditionalization. When Bayesians derived their probabilities from knowledge of relative frequencies, then an Evidential Probabilist can make corresponding inferences[92]. For example, suppose that I know that 90% of students who answered the final question in the logic exam passed. A student who is wondering about her mark, prior to her marks being released, tells me that she is fairly confident that she answered that question. Bayesian reasoning would require that I take into account both my evidence regarding the pass rate among those students *and* her degree of confidence about whether she answered that question. Analogously, Sharpening by Richness would favour reference class statements that combine these two pieces of information over those that simply used the 90% pass rate.

In the special cases in which all of the relative frequency data is precise and our new evidence takes the form of new joint distributions, Evidential Probabilist updating and Bayesian conditionalization give identical results. In a hypothetical world where all our data

[92] Kyburg (2007) p. 291.

was sufficiently precise and our evidence took the form of a simple linear accumulation of incorrigible statistical evidence, there would be no major differences between Bayesianism and Evidential Probability. The differences would only involve ancillary issues like deductive closure or the interpretations of the probabilities.

Bayesianism and Evidential Probability diverge when the Bayesian priors are not based on relative frequency data. For instance, imagine that Miss Marple knows that at least 2/10 of the guests at the dinner party are complicit in murder of the host, but she does not have any better information regarding the guilt of any particular guest. Miss Marple knows that a young woman, called Bundle, is one of the guests. Let H be the hypothesis that Bundle was complicit in the murder. Let K be Miss Marple's total relevant knowledge. The evidential probability that Bundle was complicit in the murder is $EP(H \mid K) = [0.2, 1]$, because Miss Marple's best information about Bundle's complicity is that (a) Bundle is one of the guests and (b) at least 2/10 of the guests are complicit in the murder. (It is consistent with Miss Marple's total evidence that *all* the guests were complicit.) Via direct inference, Miss Marple can infer that the evidential probability of Bundle's complicity is [0.2, 1]. Of course, if Miss Marple had some better information (such as knowing that Bundle is a morally upstanding woman) then the evidential probability could be very different. In contrast, with Bayesian reasoning Miss Marple can have a precise probability for H. It could be the case that $P(H \mid K) = 0.8$. As I shall soon discuss, the PIE is another example of a divergence between Bayesianism and Evidential Probability.

SECTION 6: EVIDENTIAL PROBABILISM AND THE PARADOX OF IDEAL EVIDENCE

I shall now analyse the PIE from the perspective of Evidential Probability. Firstly, I propose an analysis for relevance using Evidential Probability. Secondly, I apply this definition to some examples. Thirdly, I discuss the PIE using this analysis. Fourthly, I answer the Problem of Corroborating Evidence using my proposed definition. Fifthly, I contrast my proposal with some alternatives, before concluding with a general assessment of my definition and the goal that it is supposed to fulfil.

6.1 Relevance in Evidential Probability

The standard probabilistic definition of relevance uses the *values* of Bayesian probabilities as the means of distinguishing evidence from non-evidence. In contrast, my proposal will focus on whether conjoining E to K adds a 'reference class statement' that *could* be used to determine the value of an evidential probability.

As detailed in the previous subsection, one begins to determine evidential probabilities by identifying reference class statements. For a given object *i* and predicate F, it will typically be the case that K contains a number of statistical generalisations about reference classes which, according to K, include *i*. The evidential probability that 'Fi' is determined by applying the Rules of Sharpening to the set {R*a*, R*b*, R*c* … R*n*} that contains all of these reference class statements in K. Let the 'candidate reference class statements' be the set containing only those statements that (a) are potential reference class statements for statements that are not known to be tautologous or contradictory and (b) state an interval other than [0, 1], so that they are those that report interesting relative frequencies.

This concept forms the basis of my definition:

Evidential Probabilistic Definition of Evidential Relevance: E is relevant to H relative to K if and only if (E ^ K) implies a larger set of candidate reference class statements for H than the set of candidate reference class statements implied by K.

Put more simply and informally, if learning E results in richer relevant statistical information (according to the rules of Sharpening) regarding H, relative to K, then E is relevant to H, relative to K.

To explain my definition further, I shall begin with single-case hypotheses. Suppose that H states that Fi for some individual *i* and predicate F. Assume that K contains a number of candidate reference classes {R*a*, R*b*, R*c* … R*n*}, each of which state the relative frequency of F in a reference class that (according to K) includes *i*. If (E ^ K) implies a candidate reference class statement R*o* about a reference class that also contains *i*, then according to my definition E is relevant to H given K.

For instance, let F be ‘lands on a 6’ and ‘*i*’ be the label for a particular throw of a strangely shaped die with 42 sides. Let K be my background knowledge. Assume that I have a number of candidate reference class statements for H, like ‘Nearly symmetric objects with 42 labelled sides will land on a 6 in about 0.01% and 0.03% of throws in the long run’, ‘Gambling die with 42 labelled sides land on a 6 in about 0.015% to 0.025% of throws in the long-run'…

If E is the information that 'The die is one of the die from a particular board game and it is weighted to land on 6 in about 0.03% to 0.05% of throws in the long run', then (E ^ K) implies a new candidate reference class statement. Regardless of whether this statement is the appropriate basis for the values of the intervals for H given K, I have acquired new evidence according to my definition. Similarly, suppose that E is the information that the die has landed on a 6 in 30/100 throws. Now, (E ^ K) implies that the die is a member of the reference class of 42-sided die that land on 6 in 30/100 straight throws, which is relevant to the hypothesis that toss *i* will land on 6 according to both commonsense and my definition.

In the case where H describes a set of individuals, E will be relevant to H when it is relevant to one or more of the members of the set. This relevance occurs because the set of candidate reference class statements that determine the evidential probability of H is simply the set that contains all the candidate reference class statements that are relevant for all of the individuals described by H. For example, if H is the hypothesis that 'This set of 10 throws of this 42 sided die will land on a 6' and E is relevant to a hypothesis H_n that a 6 will be thrown on the *n*th throw in that set, then E is relevant to H given K. (E might be the statement that the first throw landed on a 6.) Since E is relevant to that throw, it follows from my definition that E is relevant to H. By the same reasoning, if H is the hypothesis that 'The male birth rate in humans is 0.51' and E is a report of a sample of human births, then E will be relevant to H given K.

In many cases, learning relevant evidence for a hypothesis will result in a change in the evidential probability of that hypothesis given the total evidence: if I did not know if a job candidate could speak Spanish and I subsequently learned that she was from South America, then that would raise my evidential probability that she could speak Spanish. (I shall discuss examples in which such a change does *not* occur when the set of candidate reference class statements is expanded towards the end of the next subsection.)

6.2 Additional Examples

Since all evidential probabilities are derived from the available candidate reference class statements, Evidential Probabilities can only change when the set of available candidate reference class statements has changed. In other words, whenever $EP(H \mid E \wedge K) \neq EP(H \mid K)$, it will be because $(E \wedge K)$ implies at least one candidate reference class statement that is not contained within K. Consequently, if learning E results in a change in the intervals of evidential probabilities, then E is relevant to H. I shall provide examples of the two ways in which the intervals can change.

Conjoining new evidence with the background knowledge can result in narrower evidential probabilities. Suppose that there is a very large treasure chamber that contains many items, including jewels, weapons, art, and apparently golden artefacts like cups, necklaces, and bracelets. You wonder about H, which is the hypothesis that a very inaccessible crown attached to the roof of the chamber is actually iron pyrite, i.e. fool's gold. Assume that you have no reasons in your background knowledge to think that H is likely or unlikely, such that the evidential probability of this hypothesis given your background knowledge K is $EP(H \mid K) = [0, 1]$. You examine a sample of the apparently golden artefacts and you discover that they are all genuinely golden. Assume that your sample report E and background knowledge provide suitable data for a classical statistical inference from this sample to the population of apparently golden items in the chamber in general[93]. In other words, (E ^ K) implies that in a long-run series of trials, such a sample would be unlikely to occur unless the chamber contained mostly golden artefacts. Thus, (E ^ K) implies R*a*, which states that over 50% of the artefacts in the chamber are golden. If this reference class statement is the statement that is ultimately chosen by the Rules of Sharpening for the probability of H given E and K, then EP(H | E ^ K) = [0.5, 1] and so the interval has narrowed.

Evidential Probability intervals can also widen because of new evidence. For example, upon learning the evidence, Sharpening by Specificity might require that the evidential probability is derived from a reference class statement whose intervals are wider than the initial intervals. Imagine that you are an entomologist and you are studying the spread of a virus in an artificial nest of ants. You have sampled an ant a_1 from the nest. You know that the object a_1 is an ant in the nest and that 45-55% of ants in the nest have the virus. Initially, this might be your best statistical information regarding the probability that a_1 has the virus, so that the evidential probability interval for the hypothesis 'a_1 has the virus' is [0.45, 0.55]. You subsequently learn that a_1 is a soldier ant and that only 25-40% of soldier ants in the nest are infected. Since you know that the soldier ants are a subcategory of ants in the nest, but not vice versa, Sharpening by Specificity requires that the new evidential probability of 'a_1 has the virus' is the wider interval [0.25, 0.40]. This new evidence is relevant according to my definition, because your evidence provides a new reference class

[93] Thus, you have no reason to think that your sample is highly atypical. For a more detailed discussion of inductive inferences using Evidential Probability, see Chapter 5.

statement ('25-40% of soldier ants in the nest are infected') to the set of candidate reference class statements to be used in determining the probability of 'a_1 has the virus'.

Furthermore, the intervals will not always change at all upon learning a reference class statement, because the putative relative frequencies for different reference classes do not have to differ: they can corroborate each other. Suppose that, in the example in the previous paragraph, that you learn that a_1 is a soldier ants and that 45-55% of soldier ants in the nest are infected. Thus, your new information (which I shall call "E_2") corroborates the interval that Sharpening selected from the earlier information (which I shall call "E_1") that a_1 is an ant in the nest and that 45-55% of ants in the nest have the virus and the rest of your initial background knowledge. Some people have the intuition that E_2 is relevant to the hypothesis that a_1 has the virus, relative to E_1 and your background knowledge K. ($E_1 \wedge E_2 \wedge K$) implies a reference class statement for the hypothesis that is not implied by ($E_1 \wedge K$), which is that '45-55% of soldier ants have the virus'. Consequently, my definition sides with those who have the intuition that E_2 is relevant to the hypothesis that a_1 has the virus, relative to E_1 and the background knowledge[94]. One justification for this intuition is the following: if the statistical information about the soldier ants had been different (such as '0-1% of soldier ants have the virus') then the rules of Sharpening would require a very different evidential probability interval, since Specificity requires using the new statistical information for the interval. Therefore, the additional information *could* have required a change in the evidential probability, even though it happened to repeat the existing interval. Additionally,

[94] If one does not have this intuition, and does not consider the Problem of Corroborating Evidence (discussed below) to feature a case of evidence that is relevant but does not change the evidential probability, then simply modifying the standard probabilist definition of evidence to use Evidential Probability rather than Bayesianism seems to be a satisfactory analysis of evidence. I am sympathetic to such a definition, because the use of changes in reference class systems was initially motivated by examples such as the example in this paragraph and the Problem of Corroborating Evidence. It is also pleasingly close to the definition of evidence that Bayesians have felt is intuitive.

a_1 might have been a worker ant, and your data about the relative frequency of the virus among worker ants could also be very different from 45-55%. Again, the additional information might have resulted in a different evidential probability, and simply happened to produce a value that is identical to the value given E_1 and K. These aspects of the scenario seem to be what motivates some people's intuition that E_2 is relevant given E_2 and K.

Finally, one interesting form of evidential relevance occurs when the additional evidence increases the reliability of existing evidence. In the previous example, you might merely learn evidence that supports the 45-55% figure for the relative frequency of the virus among soldier ants. More generally, information about the reliability of scientific instruments, historical sources, court witnesses, and other sources of knowledge can significantly corroborate our existing relative frequency information regarding a hypothesis, without altering the evidential probability.

6.3 The Paradox of Ideal Evidence

The essence of the PIE is that P(H | E ^ K) might be identical to P(H), even if E is evidentially relevant to H. In Popper's example, E is a report of a large number of tosses of a coin of unknown bias/fairness, in which the tosses are evenly balanced between heads and tails. There are possible Bayesian probability distributions such that E is probabilistically irrelevant to the hypothesis that the *n*th toss will be heads. Popper's example demonstrates that the standard probabilistic definition of evidential relevance is too narrow.

I shall now examine the PIE using the definition that I developed in Subsection 6.1. H is the hypothesis that the coin will land on heads in the *n*th toss. K is your background information, which tells you nothing about the fairness/bias of the coin. Prior to learning E, your evidential probability is EP(H | K) = [0, 1].

E is the report that the coin landed heads in 1,500 out of the 3,000 earlier tosses, where the *n*th toss is the 3,001st toss. Upon learning E, you learn that the coin is a member of the reference class of coins that land heads on 1,500 out of 3,000 tosses. By combining E with your background knowledge, you might be able to use classical statistics to infer that the long-run relative frequency of the coin landing heads is 0.5 ± a margin of error ε. If ε is 3%, then you have learned from (E ^ K) that:

R_1: The long-run relative frequency of heads for this coin is between 0.47 and 0.53.

Since the *n*th toss is a member of the reference class of long-run tosses of the coin, it follows that the set of reference class statements for H has expanded, because it now includes R_1 and this statement is one of the statements that must be considered when determining EP(H | E ^ K). (When R_1 is the appropriate reference class statement for H according to Sharpening, then EP(H | E ^ K) = [0.47, 0.53].) Therefore, E is relevant to H given K, according to my definition.

One can apply a similar analysis to the example of the team of economists studying whether Turkish membership of the EU would increase the GDP of the UK. Even if conjoining the economists' new data with the relevant background knowledge does not alter the evidential probability of the hypothesis, their data will be relevant if one could use it to infer reference class statements that include this event as a member of the class they describe. For instance, if they develop new statistical data on the effects of the regulatory harmonization between countries like Turkey and the UK that would result from Turkish membership of the EU, then they will have produced new reference class statements. In contrast, if there had been a bizarre miscommunication and the economists had instead produced data on the effects of the UK leaving the EU, with no data that one could use in calculating the evidential probability of Turkey's entering the EU, then they would have failed to provide any new reference class statements. Consequently, they would have failed to provide any new relevant evidence, according to my definition.

6.4 The Problem of Corroborating Evidence

In my discussion of Gemes's response to the PIE, I presented a kind of counterexample for his analysis, which I called the Problem of Corroborating Evidence. I shall now explain how my definition avoids this problem.

In the Problem of Corroborating Evidence, the task is to develop a definition of evidential relevance that includes the relevance that statements like 'This is a non-white swan' have to hypotheses such as 'All swans are white', given our background knowledge that there are counterexamples like non-white swans. In Bayesianism, the probability of the hypothesis given the evidence and the background knowledge will be equal to the probability given the background knowledge alone, and so many probabilistic definitions of relevance will fail to register the corroborating evidence as relevant.

The same phenomenon occurs in Evidential Probability: evidence that corroborates the falsification of a hypothesis will not alter the evidential probability.

Key

H: All swans are white.

¬Wb: b is a non-white swan.

K: The background knowledge, including the knowledge that there are non-white swans.

'There are non-white swans' implies ¬H. Since K is a weakly deductively closed set of statements (see Subsection 5.2) it will include ¬H. By Axiom (6), this inclusion entails that ¬H has maximal evidential probability relative to K:

(i) EP(¬H | K) = [1, 1].

By (i) and Axiom (2):

(ii) EP(¬H | K) = [1 – 1, 1 – 1] = [0, 0].

If ¬Wb is added to K, the interval values for EP(H | ¬Wb ^ K) cannot be any lower than zero[95]. However, if ¬Wb is conjoined with K and ¬Wb is not part of K, then this adds to

[95] Else ¬H would be more probable than a tautology T, since it would have a value [1 – (x < 0), 1 – (x < 0)] and T has a probability of [1, 1]. This contradicts Axiom (5) that states that no statement can be more probable than a statement it implies, since ¬H implies T. In short, the interval values cannot fall below zero because zero is a

the set of candidate reference class statements for H. One added reference class statement describes the unit set containing the swan *b*:

R*b*: All members of the reference class containing *b* alone are non-white.

Recall from Subsection 5.3 that the set of reference class statements for a plural-case hypothesis are just the set of reference class statements for the individual cases. Therefore, if the evidence ¬Wb has increased the reference statements for ¬Wb, then it has increased the reference class statements for H, and so it is evidentially relevant to H given K according to my definition, even when EP(H | ¬Wb ^ K) = EP(H | K) = [0, 0].

Since any evidence that corroborates a known falsification of a universal generalisation will be relevant according to my definition, my definition avoids the Problem of Corroborating Evidence. A report of the most recently observed black swan is as intuitively relevant to ‘All swans are white’ as the first such report. It seems to be an unusual strength of my definition that it captures this intuition.

6.5 Comparison with Some Alternatives

6.5.1 The Standard Definition

general minimum value for the value of either the lower limit or the upper limit of an Evidential Probability interval, just as it is a minimum value for a Bayesian probability.

My definition appears to be wider than the standard definition. Probabilistic relevance can only occur in Evidential Probability when the set of candidate reference class statements has changed. Consequently, all cases of probabilistic relevance in Evidential Probability will be cases of evidential relevance on my definition. Insofar as it is narrower than the standard definition (when the probability function in the definition is a normal Bayesian function) then it will presumably be because Bayesianism allows probabilistic relevance in cases where there is no foundation for such probabilistic relevance in the relative frequency data. (Bayesian reasoning that has grounds in relative frequency data is also possible in Evidential Probability.) Whether such cases are problematic depend on broader questions in formal epistemology; I shall not discuss them in this chapter.

At least under circumstances like the PIE, my definition is broader than the standard probabilistic definition, because it allows for (what in Bayesianism would be) probabilistically irrelevant statements to be evidentially relevant to a hypothesis. Furthermore, it includes corroborating evidence, as I noted in my discussion of the Problem of Corroborating Evidence in Subsection 6.4. By contrast, in the standard definition, if H and K are inconsistent, then E will be probabilistically irrelevant to H given K, as proven in Subsection 4.2.3. The PIE and the Problem of Corroborating Evidence are two cases in which my definition differs from the standard probabilistic definition; in both cases, the divergence is a strength of my proposal.

6.5.2 Keynes's Strict Definition

My definition also differs from Keynes's strict definition. In particular, it avoids Carnap's trivialization proof. On Keynes's definition, almost any arbitrarily selected statement E is relevant to any hypothesis H. The problem was that Keynes's strict definition entails that E is relevant to H, relative to our background knowledge K, if (E ^ K) implies any statement that is relevant to H, and Carnap proved that this led to excessive breadth.

In my definition, there is no such clause. If (E ^ K) does not provide us with a new candidate reference class statement for H, then E is not relevant to H given K, and by this requirement my definition avoids Carnap's trivialization proof. For example, for my actual background knowledge, 'Tweety can fly' provides no statements about the relative frequency of landing heads that could be used for the evidential probability of 'The next toss of this £2 coin will be heads', and thus their evidential relation is irrelevance according to both common usage and my definition, whereas it would be relevant on Keynes's strict definition.

6.5.3 Hempel's Theory of Relevance

In this chapter, I have not discussed Hempel's theory of evidence. (I shall discuss his theory in Chapter 4. A version of the PIE is a special case of one of the objections that Carnap raises to Hempel's theory[96].) However, one contrast that is interesting in this context is that my analysis favours Keynes and Hosiasson over Hempel: evidence can be relevant to a hypothesis, even though it does not confirm or disconfirm that hypothesis. In the PIE, the report of coin tosses does not confirm the hypothesis that the *n*th toss will land heads, because the report is no more favourable towards this hypothesis than that the *n*th toss will land tails. Nevertheless, it is evidentially relevant on my definition. Put another way, in my analysis, evidence regarding a hypothesis can be (1) favourable, (2) unfavourable, or (3) neutral. Therefore, it is possible to distinguish equivocal evidence from an absence of evidence.

---

[96] Carnap (1962) p. 480.

6.6 General Assessment

I have not proven that my proposal is free of counterexamples. There are two possible types of counterexample to my definition of evidential relevance: (a) examples in which the definition fails to include some intuitively relevant evidence or (b) examples in which the definition includes some intuitively irrelevant evidence. Popper's PIE is an instance of (a) for the standard definition, whereas Carnap's trivialization proof is an instance of (b) for Keynes's strict definition. I have not proven that my definition avoids either form of counterexample. However, it addresses the standard problems in the literature, as well as some new problems that I have developed. The key positive challenge that I have not addressed is a very general project for an Evidential Probabilist: to establish that all (or at least most) reasoning can be modelled as reasoning within Kyburg's formalism. For instance, it would be interesting to see if testimonial reasoning can be modelled using Evidential Probability and the definition of relevance I have proposed in this chapter, because testimonial reasoning is a crucial part of scientific, legal, and commonplace reasoning. This project is naturally vast and beyond the scope of this thesis. Nonetheless, the potential scope is broad: for my definition of relevance, the evidence statements that are relevant to a hypothesis do not have to *be* relative frequency statements, provided that they imply reference class statements when conjoined with the background knowledge.

Furthermore, I am not anxious to avoid all counterexamples, provided that they are not too severe and/or general. For example, if there are counterexamples from (arguably) unusual areas of reasoning, like metaphilosophy or mathematics, then that would be consistent with using my analysis for confirmation theory and normative decision theory. My definition is part of a tradition of philosophical analysis in which the goal is a formalisation of an ordinary concept that is useful for a specific purpose in a particular context. It is not intended to be a discovery of necessary and sufficient conditions for a purpose-independent objective concept[97]. If I have improved on the standard probabilistic alternatives, then I have accomplished my objective.

Finally, I have not discussed the question of evidence whose relevance consists in eliminating existing data. The removal of data from the body of science is part of scientific practice, but confirmation theorists generally ignore this practice. Thus far, those who have analysed evidential relevance have almost always discussed it in terms of background knowledge that stays constant over time. This is not an accurate model of how science works: exposing fraudulent observation reports and experimental reports is a valuable contribution to scientific knowledge; such debunking is intuitively relevant to the hypotheses for which the debunked 'evidence' was relevant.

My definition would have to be modified to be applied to a more realistic model of scientific knowledge, but it seems that the essential approach could be the same: relevant evidence for a hypothesis H consists of statements that change the statistical basis for determining the evidential probability of H. Short of discussing issues of contracting evidence and engaging with the literature on that subject, I cannot provide a detailed discussion of this issue. For this reason, I have limited my analysis to the standard *analysandum* in the literature on relevance: evidential relevance where the background knowledge stays constant over time.

[97] Carnap (1962), Chapter I, provides a discussion of this approach to analytic philosophy.

CONCLUSION

Probabilists can answer the PIE by using Evidential Probability as the epistemic probability function in a modified probabilistic definition of evidential relevance. Aside from the use of Evidential Probability, the salient difference between my answer to the PIE and the standard probabilist alternatives is that my definition uses changes in the *input* for the determination of a probability function, rather than change in the *value* of a probability function. One could use changes in the values of evidential probabilities as a sufficient but not necessary indicator that there is new relevant evidence.

Insofar as reasoning can be modelled using Evidential Probability, my answer provides a formal analysis of evidential relevance. Furthermore, for informal contexts, my answer provides a general answer to the question, ‘What is evidence?’ The answer is that evidence for a hypothesis is information that can be used (with the available background knowledge) to infer novel reference class data that could be used to determine the probability of the hypothesis via direct inference. It follows that evidential relevance is a three-place relation, because the putative evidence, the hypothesis, and the background knowledge are all vital terms in my definition. Furthermore, evidential relevance is not a matter of opinion, but instead it is determined by a logic of direct inference that provides unique values given specific inputs.

An additional feature of my analysis is that it helps clarify the concept of the quantity of relevant evidence, which Keynes and some other philosophers of science have regarded as important. On my analysis, this quantity will increase when learning evidence enables the inference of a new candidate reference class statement. This theory of changes in the quantity of relevant evidence suggests the possibility of using imprecise probability systems like Evidential Probability to create a measure for this quantity. I shall explore this possibility in the next chapter.

CHAPTER 2: IMPRECISE PROBABILITY AND THE MEASUREMENT OF THE WEIGHT OF ARGUMENT

Philosophers like Charles Sanders Peirce and John Maynard Keynes have argued that there is an important dimension of scientific reasoning that Keynes calls the "weight of argument". As I shall use the term, the weight of argument is the *quantity* of relevant evidence cited in the premises of an argument, rather than the evidence's *balance* for or against the conclusion of the argument.

Some imprecise probabilists have claimed that their systems can be used to measure this dimension of reasoning[98]. Kyburg seems to have been the first to propose such an approach[99]. Though imprecise probabilities are not the only proposed approach to measuring weight, they offer the prospect of a relatively simple measure via their degree of imprecision. These measures will be my focus in this chapter[100].

In Section 1, I describe Keynes's original (informal) concept of weight and explain its importance within epistemology and decision theory. In Section 2, I critically examine a measure developed by Walley, who used a system that I call "Imprecise Bayesianism". In Section 3, I examine Kyburg's proposal for measuring weight using imprecise probabilities. I

---

[98] Schmeidler (1989) p. 571.

[99] Kyburg (1961) p. 63.

[100] There are alternative measures of weight in Carnap (1962) p. 554-555 and Gärdenfors and Sahlin (1982).

argue that this method has fewer flaws than Walley's measure, but it nonetheless has some very severe problems. I finish by briefly suggesting an alternative approach.

SECTION 1: KEYNES AND THE WEIGHT OF ARGUMENT

Keynes uses the term 'weight of argument' to refer to his concept, but the term 'weight of evidence' is often used in the literature. I shall stick to Keynes's terminology, because the phrase 'weight of evidence' also sometimes refers to other concepts, such as the degree of confirmation of a hypothesis by the evidence. Additionally, in law, the phrase 'weight of evidence' is typically used to mean the balance of evidence[101], so my usage avoids inconsistency with a large and relevant literature.

The reference to "arguments" also suggests a conception of evidence as the premises of an argument and the hypothesis as the argument's conclusion. This matches Keynes's conception of evidence: to say that "E is evidence for H" is to claim that an argument from E to H would have a premise that is relevant to the conclusion. (See Chapter 1 Section 2.2 for a formal discussion of Keynes's theory of relevance.) Probabilist confirmation theorists like Keynes, Bayesians, and Kyburg often formalise these arguments using conditional probabilities. For example, in conditional probabilities like $P(H \mid E)$, $P(H_1 \wedge H_2 \mid E)$, and $EP(H \mid E \wedge K)$, statements to the left of '|' can be interpreted as the conclusions of arguments, while the statements to the right of the vertical bars can be interpreted as the premises.

[4] Nance (2016) p. ix.

1.1 Keynes's Concept of the Weight of Arguments

Chapter VI of Keynes's 1921 book *A Treatise on Probability* is the first extended discussion of weight. The "weight of argument" is the quantity of relevant evidence in the premises of an argument for a hypothesis. Keynes uses 'weight of argument' in a number of other senses, as Jochen Runde has clarified[102], but I shall focus on this particular meaning of the term.

Suppose that there is an argument from K to H, where H is a hypothesis and K is the relevant background knowledge. If some additional relevant evidence E is added to K, so that the new premise is (E ^ K), then the weight has increased. For example, imagine that H is the claim that a type of steel bar will be able to bear α newtons of force in a proposed Alaskan bridge and K is a conjunction of relevant background knowledge, such as descriptions of controlled experiments in which (under somewhat analogous conditions) this type of steel bar withstood this force. E is a description of experiments in which engineers were able to replicate more closely the conditions under which the bridge will be used (such as extreme cold or heavy precipitation) than in past experiments. E will add to the quantity of relevant evidence, regardless of whether it reports that the steel bar was able or unable to bear α newtons in the experiments, because it is relevant to H given the background knowledge.

Another important feature of Keynes's concept is that weight is distinct from probability. Adding some relevant evidence to the background knowledge always increases the weight, *even if* the conditional probability of the conclusion given the total evidence is unchanged. For example, imagine that you hear a weather report stating that tomorrow will

---

[102] Runde (1990) p. 279-281.

be windy. The conditional probability that tomorrow will *not* be windy, given your total evidence, has decreased. In contrast, the weight of argument has *increased*.

The informal concept of weight has a long history. I. J. Good notes that, as a metaphor, one can trace it as far back as the Ancient Greek goddess Themis and her "scales of justice"[103]. Peirce provides one of the earliest philosophical discussions. He uses a simple example of its apparent significance: imagine that there is a bag with an unknown proportion of red beans and/or black beans. Assume that we expect a red bean or black bean with equal credence. There is an apparent difference between founding our expectation on a sample of 1,000 beans drawn from the bag, rather a sample of only 2 beans. If we are extrapolating from the former bag, then we seem to be entitled to greater confidence (in some informal sense) than from the latter bag[104]. Peirce argues that the difference is the greater quantity of evidence that the larger sample provides.

James Franklin argues that weight is important in legal reasoning. According to his analysis, the legal concept of "proof beyond reasonable doubt" should be interpreted as two distinct requirements: (1) the prosecution has proven that the guilt of the accused is highly probable given the evidence *and* (2) the prosecution has provided a sufficiently large quantity of evidence[105]. For instance, a prosecutor might provide several eyewitness reports that all give *hints* that the accused murdered the victim in their home: one eyewitness claims she saw the accused near the house that night; another eyewitness saw them driving away quickly

[103] Good (1985) p. 249.

[104] Peirce (1932) 2.677.

[105] Franklin (2012) p. 237.

within 20 miles of the crime scene; and so on. However, the cumulative weight of these different pieces of evidence might be jointly insufficient to prove guilt beyond reasonable doubt. Franklin also notes that weight seems to be an almost explicit part of some legal reasoning. One example is that there is a "burden of production" that is required before civil cases can begin, which means that plaintiffs must present a significant amount of relevant evidence ahead of the case[106].

As O'Donnell emphasises, the adjective "relevant" in front of "evidence" is very important[107]. Keynes is not referring to the sheer number of statements on the right hand side of a conditional probability like P(H | E) or the sheer bulk of information that these statements contain. By "relevant evidence", Keynes is only referring to those statements contained in E that meet a suitable definition of evidential relevance. For instance, the argument from 'This 300-fold sample of bees are all female' to 'It will be windy tomorrow', relative to our actual background information, has no weight, because the premise is irrelevant to the conclusion. Only *relevant* evidence will add to the weight of argument.

Summarising, I shall understand weight as the quantity of relevant evidence for a hypothesis. It is a two-place relation between a conjunction of premises and a conclusion, or between evidence and a hypothesis. (These are equivalent in Keynes's theory of evidence.) The weight of an argument is neither reducible to the absolute amount of information in the premises, nor to the balance of the premises in favour or against the conclusion.

[106] Franklin (2006) p. 161.

[107] O'Donnell (1989) p. 69.

1.2 The Significance of Weight

Keynes does not doubt that weight is an important concept, but he is unsure about precisely *how* weight has practical significance[108]. In the previous subsection, I described some examples of the concept in action from Peirce and Franklin. To highlight the importance of weight, I shall now give some further instances of philosophers who have applied the concept.

In confirmation theory, Hosiasson uses Keynes's concept to explain the value of acquiring evidence that does not confirm or disconfirm a hypothesis. As I discussed in the previous chapter, Hosiasson argues that weight can increase even with evidence that is equivocal with respect to a hypothesis. For instance, a sample report of 50 coin tosses of a coin with unknown bias/fairness that landed heads in 25/50 tosses and tails in the other 25 tosses might not alter the probability of the hypothesis that the *n*th toss will land heads, but it will add to the weight for this hypothesis given the total evidence[109].

In decision theory, Franklin notes that there are trade-offs between (1) the benefits of acquiring more evidence and (2) the cost of acquiring evidence. The choice of a particular balance between (1) and (2) involves Keynes's concept of weight, because both refer to the notion of a greater or lesser quantity of relevant evidence[110]. Imagine that you are a member of a UK Treasury committee that the government has established to assess whether merging National Insurance and income tax will result (*ceteris paribus*) in an increase in revenues.

[108] Keynes (1921) p. 76-77.

[109] Hosiasson (1931) p. 36.

[110] Franklin (1998) p. 112-113.

You aim to study a large variety of econometric, historical, and theoretical research that is relevant to this hypothesis before coming to a judgement. While you are amassing the evidence, you will make decisions about whether you have acquired a sufficiently large quantity of evidence to give an informed judgement. Since you cannot acquire all the relevant evidence, there must be a point at which you judge that "enough" evidence has been acquired. This process involves decisions about weight: at a given time *t*, have you sufficient relevant data to form an informed judgement?

In the philosophy of law, Barbara Davidson and Robert Pargetter use weight in their analysis of the legal phrase "beyond reasonable doubt"[111]. Like Franklin, they argue that this notion involves more than just the probability of guilt. They argue that it involves two other aspects of the evidence that the prosecution and defence present at a trial: (1) the reliability of the evidence and (2) the quantity of evidence presented. Davidson and Pargetter use the following example to illustrate their point: there are 10 suspects and the jury knows, via reliable evidence, that 9/10 of the suspects are guilty and one suspect is innocent. Furthermore, the suspects confirm this information, but they refuse to confess who is innocent. For any given suspect, the probability of guilt is high. Yet this high probability does not suffice to prove their guilt beyond reasonable doubt. Furthermore, a higher threshold of probability than 90% will not help either, because we can simply consider cases in which 99/100 suspects, 999/1000 suspects etc. are known to be guilty. It is implausible that the problem in this case is that the threshold for guilt is a 90% probability, rather than a 99.9% probability or a 99.99% probability. Davidson and Pargetter justify this requirement for more evidence on the grounds that just one additional piece of evidence (e.g. nine of the suspects

[111] Davidson and Pargetter (1987) p. 187.

confessing without coercion or incentives) would radically alter the probability of any particular suspect being guilty, so that a small quantity of weight is sufficient to change the balance of the evidence in the case[112]. They also note that judges can dismiss cases on the grounds of insufficient evidence, regardless of the balance of the evidence or the evidence's reliability[113]. Davidson and Pargetter argue for explicitly incorporating weight into juries' reasoning and they suggest that the standards for weight in determining guilt should vary with the seriousness of crimes, just as the seriousness of crimes affects the standards for the reliability of evidence and the probability of guilt.

Keynes doubts that weight can be measured quantitatively, though he nonetheless thinks that it can be incorporated into decision theory[114]. Nonetheless, given weight's many applications, a quantitative formalization of weight could be useful. Imprecise probabilities offer one possible formal basis for such measurement. In the rest of this chapter, I shall critically examine two quantitative measures of weight. Insofar as such measures are successful, they offer the prospect of better theories of evidence and rational decision-making. However, I shall argue that the proposed measures have several exigent problems.

[112] Davidson and Pargetter (1987) p. 183.

[113] Davidson and Pargetter (1987) p. 184.

[114] Keynes (1921) p. 72-73.

SECTION 2: IMPRECISE BAYESIANISM AND THE WEIGHT OF ARGUMENT

In this section, I examine Walley's measure of weight, which uses a probability theory that I shall call 'Imprecise Bayesianism'. I describe an existing objection to this measure, which derives from a problem called "dilation". I also adapt a standard general objection to Imprecise Bayesianism into an objection against Walley's measure. I finish by developing several novel objections to Walley's measure.

2.1 Imprecise Bayesianism

I provided an outline of Bayesianism and its different forms in Chapter 1 Section 2.1. As I shall use the term, 'Imprecise Bayesians' are a subgroup of Subjective Bayesians who retain much of standard Bayesianism, but with two important differences. Firstly, they drop the requirement that all probabilities must be additive. Secondly, they allow probability functions to take values that are not real numbers. I outline the formal framework of Imprecise Bayesianism, before describing the updating process within this system.

2.1.1 The Formal Framework

In Imprecise Bayesianism, the probability of a statement is determined by a continuous set of ordinary Bayesian probability functions, instead of via a single function. Each individual function produces a distribution that satisfies the axioms of additive probability. Imprecise Bayesians interpret these values as subjective degrees of belief, where this magnitude is represented by a pair of real numbers, rather than a single-number. This set of probability functions is sometimes called a "representor"[115], a "credal set", or a "committee"[116]. I shall use the last term for the set. This committee, *as a whole*, represents the credences of the agent being modelled.

Before explaining how the membership of the committee is determined, I shall outline the terminology that I shall use. As in previous chapters, Ω is the domain of the function, which is a set of statements, while other Greek letters represent arbitrarily selected members of that domain. I shall use 'P' to refer to a function whose values are determined by a committee. I shall add labels to 'P' to refer to individual members of that committee. In particular, P_i will refer to an arbitrarily selected member of the set of functions. I shall use '$\underline{P}$ (Φ)' to denote the lower bound of the interval for a statement Φ, which is given by the function in the committee that gives Φ the lowest probability. I shall use '$\overline{P}$(Φ)' to denote the upper bound of the interval for Φ, which is given by the function in the committee that gives Φ the highest probability. Thus, when P(Φ) = [0.3, 0.7], then the function in the committee that has the lowest value for Φ's probability assigns Φ a value of 0.3, while the function that has the highest value for this statement assigns Φ a value of 0.7. When there is only one value that the members of the set assign to Φ, I shall simplify the equations by using a single real number *r* for the value of the Imprecise Bayesian probability, instead of a degenerate interval, so that 'P(Φ) = *r*' obtains when every member of the committee assigns Φ a value of *r*. Finally, I shall add the mark ' ´ ' to 'P' to indicate that the distributions in the members of the committee have been updated.

[115] Van Fraassen (1990) p. 347.

[116] Bradley (2015).

The following axioms govern Imprecise Bayesian probabilities:

(i) $\underline{P}(\Phi) = x$ and $\overline{P}(\Phi) = y$ in the committee that determines $P(\Phi) = (x, y)$.

(This axiom defines how P's value is determined by the set of probability functions that compose the committee.)

(ii) If $P(\Phi) = [x, y]$, then $0 \leq x \leq y \leq 1$.

(The values for x and y are an ordered pair of real numbers between 0 and 1 inclusive; they are ordered by their numerical value.)

(iii) $\underline{P}(\Phi) = 1 - \overline{P}(\neg\Phi)$.

(The lower bound of a statement and the upper bound of its negation vary in proportion.)

(iv) If $P(\Phi \wedge \Psi) = 0$, then $\underline{P}(\Phi) + \underline{P}(\Psi) \leq \underline{P}(\Phi \vee \Psi) \leq \overline{P}(\Phi \vee \Psi) \leq \overline{P}(\Phi) + \overline{P}(\Psi)$.

(This axiom adapts axiom (iii) from Chapter 2 Section 2 for imprecise probabilities. For each of the extreme functions in the set, the sum of the probabilities of two statements cannot exceed the probability of their disjunction.)

(v) If $P(\Phi) = [x, y]$ and $x = y$, i.e. the interval is a degenerate interval, then $P(\Phi)$ obeys the axioms of additive probability.

(Precise Bayesian probabilities are a special case of Imprecise Bayesian probabilities.)[117]

Subject to these axioms, the choice of P(Φ) for any given statement in the domain is subjectively determined by the Imprecise Bayesian. One approach to calculating these values is to let *x / y* characterise the least favourable betting odds at which, under special conditions[118], you will buy a bet on a statement. The value of $\underline{P}(\Phi)$ is determined by the following definition:

Implied Lower Bound Probability: *z* / *w* are your minimum buying odds for Φ if and only if $\underline{P}(\Phi) = \frac{w}{w+z}$.

In contrast to precise Subjective Bayesianism, the minimum buying odds and the minimum selling odds can be unequal. When they are unequal, the Imprecise Bayesian interval for a statement Φ will be imprecise. When the minimum buying odds are equal to the minimum selling odds, then the intervals are degenerate, such that $\underline{P}(\Phi) = \overline{P}(\Phi)$[119]. For example, suppose that the minimum odds at which your will buy a bet on Φ are 4/1, so that $\underline{P}(\Phi) = 0.2$, but that the minimum odds at which your will sell a bet on Φ are 1/4, so that $\overline{P}(\Phi) = 0.8$. Your Imprecise Bayesian probability for Φ is P(Φ) = [0.2, 0.8].

Full distributions are determined in the same way: for a given committee S, the value

[117] These axioms are adapted from Howson and Urbach (1993) p. 88.

[118] These are the same conditions that precise Bayesians often use to characterise precise probabilities in terms of fair betting odds.

[119] Bradley (2015), formal appendix.

of each joint distribution is determined by (1) the function that is a member of S and which gives the *lowest* value for that joint distribution and (2) the function that is a member of S and which gives the *highest* value for that joint distribution. For both marginal probabilities and joint distributions, it is the committee that determines the values of P. Membership of the committee is a matter of the Imprecise Bayesian's choice; the procedure I outlined above that uses minimum buying/selling odds is one method for determining such choices.

If an Imprecise Bayesian has a joint distribution from which they want to obtain a marginal distribution, then they can do so by deriving the marginal probabilities for the lower and upper members of the committee. Suppose that $\Psi \in \Omega$, $\Phi \in \Psi$, and Ω is the Cartesian product of Ψ and Γ, denoted as $(\Psi \times \Gamma)$[120]. Let $\underline{P}1$ be the agent's probability distribution defined over Ω and $\underline{P}2$ be the probability distribution for the agent defined over Ψ. Under such circumstances, $\underline{P}2(\Phi) = \underline{P}1(\Phi \times \Gamma)$. The same holds, *mutatis mutandis*, for the value of $\overline{P}2(\Phi)$[121].

The goal of the Imprecise Bayesian formalism is a more faithful and philosophically fecund representation of psychological states than a single precise probability function. Imprecise Bayesians view the intervals as *bona fide* epistemic probabilities, just like precise probabilities. They do not regard the intervals as mere computational expedients or as a method for handling the problems of eliciting precise credences from an agent. The intervals are intended to be a superior formalism for epistemic probability theory, not a second-best option that must be adopted for pragmatic reasons.

[120] The Cartesian product of two sets is the set that contains all the possible ordered pairs of members of those sets. For example, the Cartesian product $(\Phi \times \Psi)$ where Φ = {Heads, Tails} and Ψ = {1, 2} is {(Heads, 1), (Tails, 1), (Heads, 2), (Tails, 2)}.

[121] Walley (1991) p. 181-182.

It is also important to distinguish Imprecise Bayesian probabilities from interval-valued estimates of a single number, such as a relative frequency or value of a propensity. For both uses of interval-valued probabilities, we might articulate a probability of P(H | E) = [0.4, 0.6], as “The probability is between 0.4 and 0.6”. However, in Imprecise Bayesianism, the interval value is not a range that includes the value of a precise objective probability, where this precise value is the “real” probability. In addition, the interval value is *not* an estimate of a real-valued number that is the actual value of the probability for that agent. Instead, it is an interval that describes the values that different members of the committee give for a particular statement.

2.1.2 Updating with Imprecise Bayesian Probabilities

In Imprecise Bayesianism, conditionalization of a statement H takes the form of conditionalizing for the entire committee, using the equation:

$$P(H \mid E) = \{\, P_i(H \mid E), P_i \in P, P_i(E) > 0 \,\}$$

- which says that the conditional probability for the committee is determined by the conditional probability for each member P_i of the set, provided that $P_i(E)$ exceeds 0.

Naturally, examining every probability function in the continuous distribution is normally out of the question. (The exceptions occur when the interval is degenerate.) Instead, Imprecise Bayesian conditionalization begins via examining $\overline{P}$ and $\underline{P}$. Suppose that an agent is conditioning H on E. They must undertake the following two procedures in sequential order:

(1) Does $\overline{P}$(E) > 0? If not, determine the remaining function in the set that has the highest prior probability for H; this function is the new upper limit function in the set. They must then performs the same process, *mutatis mutandis*, for $\underline{P}$.

(2) Let $\overline{Pr}$ and $\underline{Pr}$ be the remaining upper/lower limit functions for H after carrying out procedure (1). The agent's conditional probabilities for H given E are the new upper/lower values for H in the new distribution P´ that has been conditioned on E. The other functions in the new set are the functions that provide the continuum of intermediate values between $\overline{Pr}(H \mid E)$ and $\underline{Pr}(H \mid E)$.

For example, imagine that you are making a random draw of a card from a deck of uncertain composition:

Key

H: An Ace of Spades will be selected.

E: A black card will be selected.

You feel that the deck might be slightly stacked in favour of Aces of Spades. You decide that it plausible that there are 3 Aces of Spades in the deck, but no more than this amount. You also feel that it is plausible that the deck has the normal ratio of 1/52 Aces of Spades, but not any lower proportion. 1/52 is your lower bound prior probability for H. The 3/53 is your upper bound prior probability for H. An Imprecise Bayesian represents this state of belief as $P(H) = \left[\frac{1}{52}, \frac{3}{52}\right]$.

Suppose that E has the following probability according to $\underline{P}$ and $\overline{P}$:

$$\underline{P}(E) = \frac{26}{52} = \frac{1}{2}$$

$$\overline{P}(E) = \frac{28}{52} = \frac{7}{13}$$

For both functions, the likelihood of E given H is 1. By Bayes's Theorem:

$$\underline{P}(H \mid E) = \frac{\underline{P}(E \mid H)\underline{P}(H)}{\underline{P}(E)} = \frac{1}{26}$$

$$\overline{P}(H \mid E) = \frac{\overline{P}(E \mid H)\overline{P}(H)}{\overline{P}(E)} = \frac{3}{26}$$

Hence:

$$P(H \mid E) = \left[\frac{1}{26}, \frac{3}{26}\right]$$

Thus, after updating with E:

$$P'(H) = \left[\frac{1}{26}, \frac{3}{26}\right]$$

As the example illustrates, conditionalization in Imprecise Bayesianism is somewhat more complex than in precise Bayesianism, but it is not always computationally difficult. (There are other general computational problems in Imprecise Bayesianism: Kyburg notes that assigning a prior distribution requires assigning probabilities to each subset of state descriptions, in addition to each individual state description[122].) The salient mathematical difference is that Imprecise Bayesian updating does not involve conditionalization for a function. Instead, one eliminates functions for which $P_i = 0$ and updating the remaining members of the committee.

Updating with conditionalization is not the only possible updating method for imprecise probabilities. Evidential probabilities are updated by conditionalization under some circumstances and by different rules in other circumstances, as described in Chapter 1 Subsection 5.3. It would also be possible to develop systems that were similar to Imprecise Bayesianism, but had weaker or stronger updating rules. A very weak updating rule could turn membership of the set into a matter of personal caprice. A stronger updating rule could involve conditionalization plus additional requirements that a function must satisfy in order to remain in the committee. I shall discuss some modifications of Imprecise Bayesian updating in Subsection 2.4.

Finally, I shall state the principal contrasts between Imprecise Bayesianism and Evidential Probability. Firstly, the two systems use different interpretations of probabilities. Imprecise Bayesian probabilities are psychological statements about the credences of an

[122] Kyburg (1992) p. 192-193. A state description is a conjunction of assertions/negations of each statement in the domain of P.

imaginary agent, where their states of belief are represented by intervals determined by the set. In contrast, evidential probabilities are normative claims about purely formal evidential relations between statements. Secondly, there are formal differences. Imprecise Bayesian probabilities are intervals that range over a set of precise probability functions. Evidential Probability intervals range over values from the relative frequency statements that survive the rules of Sharpening. Imprecise Bayesian probability statements describe sets of functions, whereas Evidential Probabilities are akin to deductive logical entailment relations because they are derived via formal rules. Thirdly, there is a difference in the updating method. Imprecise Bayesian updating is always a modified form of conditionalization. Evidential Probability updating involves adding E to an agent's body of statements and checking, for each statement H, whether E contains information about relative frequencies that alters the evidential probability of H relative to the agent's total evidence. Therefore, despite the fact that both systems are imprecise probability systems, they have many important differences. These differences will be crucial to my arguments in this chapter.

2.2 Imprecise Bayesianism and Weight

Using Imprecise Bayesianism, Walley proposes the following measure for the weight of E with respect to H:

$$Degree\ of\ Imprecision = DI = \overline{P}(H \mid E) - \underline{P}(H \mid E)$$ [123]

[123] Walley (1991) p. 522.

Informally, Walley's DI measure of weight involves using the difference between the lower bound value and the upper bound of the Imprecise Bayesian probability to measure the weight of an argument from E to H.

Part of the appeal of this measure is that the representative richness of imprecise probabilities is one of primary arguments for Imprecise Bayesianism. For instance, the theory seems particularly suitable for modelling differences of ignorance. An Imprecise Bayesian can argue that imprecise values seem most natural when one is in a position of ignorance: if I am confronted with the question, 'Are tardigrades carnivorous?', when I know very little about tardigrades (or microscopic organisms in general) then it seems strange to assign a precise probability to the hypothesis that they are carnivorous, whereas it is relatively intuitive for me to claim that the hypothesis's probability is greater than 1% and less than 99%. If I were to acquire more information and reduce my ignorance, a more precise probability would arguably become more intuitive.

Conversely, many of the circumstances in which precise probabilities seem most natural are situations in which there is a rich body of evidence. People typically learn about precise probabilities via gambling devices like coins, cards, balls in urns, and so on, where these devices and their historical antecedents have been used for hundreds or even thousands of years. The behaviour of these devices is fairly well-grasped by ordinary folk physics, while the behaviour of the users of these devices is also typically well-grasped by ordinary folk psychology, and when issues that could require acquiring more evidence are raised (e.g. the possibility of a coin landing on its side) these are typically removed by assumption or idealization. At least in some cases, there is a *prima facie* relationship between the degree of imprecision of the intuitively appealing probabilistic statements and the quantity of relevant evidence, so that Walley's measure of weight has some initial plausibility.

Before going further, I shall make an important formal point. Usually, an agent's background knowledge will provide some relevant evidence about a hypothesis. In precise Bayesianism or Imprecise Bayesianism, P(H) can be regarded as the probability of H given the background knowledge. If I feel that H has a probability of [0.5, 0.9] given what I already know, then this value can serve as both my prior probability *and* my probability for H given my background knowledge. In my discussion below, the DI measure for P(H) will be the value for the weight of H given the agent's background knowledge.

I shall now look at an example using Walley's measure. Imagine that an Imprecise Bayesian has inherited a specially designed Magic 8-Ball. She is wondering what answer the ball will give to her next question when she shakes it. H is the hypothesis that 'This unusual Magic 8-Ball will give an equivocal answer'. Suppose that she knows that the Magic 8-Ball is designed to give equivocal answers at a long-run relative frequency of ≥ 0.5. Furthermore, she has already shaken it several times and its answer was equivocal in each case. For all she knows, the Magic 8-Ball's internal design causes it to *always* give equivocal answers. Assume that her prior probability for H is:

$$P(H) = [0.5, 1]$$

Using the DI measure, the weight for P(H) is:

$$DI(H) = 1 - 0.5 = 0.5$$

Subsequently, she discovers a manual for the 8-Ball. The manual tells her that the ball's relative frequency of equivocal answers is about 70% to 80%. Let E represent what she has learned from the manual. Suppose that:

$$P(H \mid E) = [0.7, 0.8]$$

Intuitively, E is relevant evidence for H, given her background knowledge. The DI measure registers this increase in the quantity of relevant evidence, because the interval has narrowed:

$$DI(H \mid E) = 0.8 - 0.7 = 0.1$$

However, Walley's formalism for his measure can be a little confusing, because the degree of imprecision has *fallen* as relevant evidence has increased. The values of DI and the quantity of relevant evidence will vary inversely, so that Walley's measure is linguistically awkward. A trivial modification makes the measure more verbally felicitous:

$$Weight = DI = 1 - (\overline{P}(H \mid E) - \underline{P}(H \mid E))$$

Thus, the DI measure's values are 1 minus the degree of imprecision. DI will be a large fraction when weight is high according to the measure and a small fraction when weight is low according to the measure.

Walley provides an informal example of this sort of measure in action[124]. Let H be a prediction about the next trial in a series of probabilistically independent trials, like a toss of a coin of unknown bias or fairness. Let E_1, E_2 … E_n be a set of reports of individual trials, such as the tosses of the coin. Under appropriate values for the full probability distribution, it is intuitive that the weight of argument will vary in proportion to *n*, where *n* is the number of reports. Similarly, the DI measure will increase (in the limit) as the number of trials increases, because the conditional probabilities of the $n + 1$ toss will converge (again, given an appropriate full distribution) towards a single value. Therefore, *if* the probability distributions for the functions in the set have suitable values, then weight will increase (in the limit) according to both intuition and the DI measure.

However, the DI measure does not always perform so adroitly. In the next five subsections, I shall present problems for this method of measuring weight.

2.3 Dilation

Dilation is a well-established paradox in Imprecise Bayesianism, in which imprecise posteriors can be wider than their priors, because of the updating procedure in the system.

[124] Walley (1991) p. 211.

Some philosophers use dilation as an objection to Imprecise Bayesianism in general[125] and defenders of imprecise probabilities have made rebuttals to this criticism[126]. In contrast, in this chapter I shall only discuss dilation as a problem for the DI measure, because I am concerned with the measurement of weight, as opposed to providing an overall evaluation of Imprecise Bayesianism. In this subsection, I shall explain how learning new evidence can widen the probabilities in Imprecise Bayesianism and why this presents a problem for the DI measure. Seamus Bradley has also made this point[127].

For example, suppose that I am about to randomly select a card from a pile of 40 cards. I know that 20 of the cards are red and 20 of the cards are black. However, I do not know exactly which cards are in the pile. In particular, I do not know the proportion of odd-numbered cards to even-numbered cards. Assume that the following Imprecise Bayesian distribution characterises my degrees of belief:

Key

H: The card is red.

X: The card is even.

Y: The card is red if and only if it is even.

(1) $P(H) = 0.5$

(2) $P(X) = [0, 1]$

(3) $P_i(X \mid H) = P_i(X)$ (For any function *i* in the set, X is independent of H.)

[125] White (2010).

[126] Pedersen and Wheeler (2014).

[127] Bradley (2015).

The value of P(Y) can be determined from the probabilities above. I shall prove this by deriving the value for an arbitrarily selected function P_i and generalising via mathematical induction that the entire set of functions must have this value.

Proof 1

Claim: That (1) and (3) imply that P(Y) = 0.5.

By the symmetry of probabilistic independence, it follows from (3) that:

(4) $P_i(H \mid X) \ = \ P_i(H)$

(H ↔ X) is logically equivalent to ((H ^ X) v (¬H ^ ¬X)). By Axiom (iii) in Chapter 1 Subsection 2.1.2, which I shall call the 'Additivity Axiom':

(5) $P_i((H \wedge X) \ v \ (\neg H \wedge \neg X)) = P_i(H \wedge X) \ + \ P_i(\neg H \wedge \neg X)$

The challenge is now to determine the value of each disjunct without assigning a particular value to P_i(X). H is logically equivalent to ((H ^ X) v (H ^ ¬X)). Hence, from (1):

(6) $P_i(H) = P_i((H \wedge X) \ v \ (H \wedge \neg X)) = 0.5$

From (6) and the Additivity Axiom:

(7) $P_i((H \wedge X)\ v\ (H \wedge \neg X)) = P_i(H \wedge X) + P_i(H \wedge \neg X) = 0.5$

Since (3) and (4) state that H and X are both independent of each other, it must be the case that:

(8) $P_i(H \wedge X) = P_i(H \wedge \neg X)$

- else an assertion of X would be relevant to the probability of H. Therefore, $P_i((H \wedge X)$ v $(H \wedge \neg X))$ is the sum of two equal numbers, one of which is $P_i(H \wedge X)$. From (7) and (8):

(9) $P_i(H \wedge X) = \frac{0.5}{2} = 0.25$

This provides the probability for the first disjunct in (5). Since H and X are independent:

10) $P_i(H \wedge X)\ v\ P_i(H \wedge \neg X) = P_i(\neg H \wedge X)\ v\ P_i(\neg H \wedge \neg X)$

- else the truth of H would be probabilistically relevant to X. From the Additivity Axiom:

(11) $P_i((\neg H \wedge X)\ v\ (\neg H \wedge \neg X)) = P_i(\neg H \wedge X) + P_i(\neg H \wedge \neg X)$

From (7), (10), and (11):

(12) $P_i(\neg H \wedge X) + P_i(\neg H \wedge \neg X) = 0.5$

Since H and X are independent:

(13) $P_i(\neg H \wedge X) = P_i(\neg H \wedge \neg X)$

- else the truth of H would be relevant to the probability of X. Hence, $P_i((\neg H \wedge X) \vee (\neg H \wedge \neg X))$ is the sum of two equal numbers, one of which is $P_i(H \wedge X)$. From (12) and (13):

(14) $P_i(\neg H \wedge \neg X) = \frac{0.5}{2} = 0.25$

From (5), (9), and (14):

(15) $P_i((H \wedge X)\ v\ (\neg H \wedge \neg X)) = P_i(H \wedge X) + P_i(\neg H \wedge \neg X) = 0.5$

By (15) and the logical equivalence of $(H \leftrightarrow X)$ and $((H \wedge X) \vee (\neg H \wedge \neg X))$:

(16) $P_i(H \leftrightarrow X) = 0.5$

Finally, substituting Y for (H ↔ X):

(17) $P_i(Y) \ = \ 0.5$

As P_i was arbitrarily selected from the committee, it follows that Y's probability is 0.5 in all members of the set that determines the values of the function P. Therefore, (1) and (3) imply that P(Y) = 0.5.

The problem of dilation concerns the value for P(H | Y) in a distribution for which (1), (2), and (17) all hold. Imagine that you select the card from the pile behind a screen, and give me only the cryptic clue that 'Either (a) the card is red and even or (b) black and odd'. This is logically equivalent to telling me that Y is true. If Y is true, then H and X must have the same truth-values:

(18) $P(H \mid Y) = P(X \mid Y)$

The problem is now to determine the values of the probabilities in (17). As neither $\underline{P}$ nor $\overline{P}$ assign a probability of zero to Y, both will remain in the committee. The other functions in the committee give the continuum of values between the values for these extreme functions. Thus, to calculate P(H | Y), it is sufficient to calculate $\underline{P}(H \mid Y)$ and $\overline{P}(H \mid Y)$.

From (2):

(19) $\underline{P}(X) = 0$

The conditional probability of a statement with an extreme prior probability is identical to the statement's prior, so from (19):

(20) $\underline{P}(X \mid Y) = 0$

Y states that H is true if and only if X is true. Hence, from (20):

(21) $\underline{P}(H \mid Y) = \underline{P}(X \mid Y) = 0$

For $\overline{P}(H \mid Y)$, the proof is almost identical. From (2):

(22) $\overline{P}(X) = 1$

For reasons analogous to the derivation of (20):

(23) $\overline{P}(X \mid Y) = 1$

For reasons analogous to the derivation of (21):

(24) $\overline{P}(H \mid Y) = 1$

Finally, from (21) and (23):

(25) $P(H \mid Y) = [0, 1]$

Hence, in a new committee P´ that is conditioned on Y:

(26) $P'(H) = [0, 1]$

This completes a proof that, in Imprecise Bayesianism, learning that a statement Φ that has a precise probability is equivalent to a statement Ψ with a relatively imprecise probability can lead to an imprecise value for the posterior probability of Φ.

Dilation presents a problem for using DI as a measure of weight, because the imprecision of the probabilities can increase without removing relevant evidence. (This is problematic regardless of whether one believe that Y is intuitively relevant to H.) In the proof above, P(H) initially has a precise value, meaning that I consider its probability to be exactly 0.5 given my background knowledge. According to the DI measure, this entails that my total evidence (prior to learning Y) has maximal weight with respect to H:

$$DI(X) = 1 - (0.5 - 0.5) = 1$$

However, the derivation of (26) proves that learning Y results in *minimal* weight for H given my new total evidence:

$$DI(H \mid Y) = 1 - (1 - 0) = 0$$

Consequently, according to the DI measure, learning E has *reduced* my total quantity of relevant evidence regarding X.

Dilation presents several problems for the Imprecise Bayesian. Firstly, one might have the intuition that Y ought to have the opposite effect on the degree of precision: the precision of P(H) should dominate the imprecision of P(X). Secondly, it is strange that adding information can lead to a decrease in weight. For my purpose, it is the second aspect of dilation that is important.

An Imprecise Bayesian might respond by forbidding any precision in the prior distribution. If one assigned [0, 1] to every statement as its prior probability, then dilation would be impossible. In the example above, if it were the case that P(H) = [0, 1], then there is no value that P(E) might have such that learning (H ↔ E) increases the imprecision of H's probability, because H's initial probability (prior to learning E) is already maximally imprecise. As a result, it would not be possible to dilate the probability for H.

There are two problems with this suggestion. The first is that, assuming it is possible to narrow the [0, 1] interval, then the DI measure would still have problems with sequences of updating events.

Proof 2

Claim: Dilation is possible even without precision in the prior probabilities.

Again, consider the random draw of a card from a pile of unknown composition. Suppose that my Imprecise Bayesian prior distribution is the following:

Key

H: The card is red.

X: The card is even.

Y: The card is red if and only if it is even.

E: Exactly 1/2 of the cards in the deck are red.

(1) $P(H) = [0, 1]$

(2) $P(X) = [0, 1]$

Suppose, for the sake of argument, that it is possible to conditionalize H on E to narrow it, such that:

(3) $P(H \mid E) = 0.5$

Hence, conditioning H on E produces the new distribution P´, such that:

(4) $P'(H) = 0.5$

The probability of X is unchanged. As a result, P´(X) is still [0, 1]. However, (4) can simply be substituted for (1) in the derivation of (25) in Proof 1. Therefore, by the same reasoning, dilation will occur upon learning Y[128]. Forbidding precise prior probabilities (even prior probabilities with *any* precision) does not avoid the problem of dilation.

Secondly, even aside from the problem of sequences of conditionalization, the [0, 1] interval poses its own problems for the DI measure of weight, due to the problem of inertia. I shall discuss this problem in the next subsection. For this subsection, the significance of these problems is that there is a price for avoiding initial precision in response to dilation problems.

2.4 The Problem of Inertia

Like dilation, the phenomenon of inertia is a well-established challenge to Imprecise Bayesianism. However, unlike dilation, the challenge that inertia poses for the DI measure does not seem to have been previously noted. Walley provides an early discussion in his

[128] If one learns (E ^ Y) together, then there is neither a narrowing nor a dilation of the precision of H's probability, because the probability values for precise Bayesian probability functions in the committee are invariant to the order in which evidence is acquired, such that learning E and Y together is equivalent to learning them one-by-one in either order. (See Howson and Urbach p. 26-27 for a discussion of the time-invariance of Bayesian probability functions.) Again, regardless of whether this is problematic for the system of Imprecise Bayesianism, it is problematic for Walley's measure, since E will not add to H's weight when it is learned alongside Y, even though it is assumed that the probability distributions in the committee are such that E would have added to the weight had Y not been learned.

study of imprecise probabilities[129].

To set up the problem of inertia, it is helpful to examine the Imprecise Bayesian updating of the [0, 1] interval via conditionalization on deductively conclusive evidence. Consider the following probability distribution:

Key

H: There are Edwardian buildings on Old Elvet.

E_1: There are no Edwardian buildings on Old Elvet.

E_2: The Old Shire Hall is an Edwardian building and it is on Old Elvet.

(1) $P(H) = [0, 1]$

(2) $P(H \mid E_1) = 0$

(3) $P(H \mid E_2) = 1$

If an Imprecise Bayesian with this distribution learned E_1, then she should eliminate the function $\overline{P}$, because (1) implies that:

(4) $\overline{P}$(H) = 1

From (4) and the logical incompatibility of H and E_1, it follows that:

129 Walley (1991) p. 367-369.

(5) $\overline{P}(E_1) = 0$

Thus, we must exclude $\overline{P}$ from the committee upon learning E_1. For the remaining probability functions, it follows from (2) that $P_i(H \mid E_1) = 0$. Since the conditional probability for the new committee of functions is $P(H \mid E_1) = 1$, the probability of H in the Imprecise Bayesian distribution P´ that has been conditioned on E_1 is P´(H) = 0.

Similarly, consider E_2, which implies H. Upon learning E_2, we can eliminate the function $\underline{P}$, because (1) implies that:

(6) $\underline{P}(H) = 0$

From (6) and the logical incompatibility of H and E_2, it follows that:

(7) $\underline{P}(E_2) = 0$

Thus, we must exclude $\underline{P}$ from the committee upon learning E_2. For the remaining probability functions, it follows from (2) that $P_i(H \mid E_2) = 1$. Since the conditional probability for the new committee of functions is $P(H \mid E_2) = 1$, the probability of H in the Imprecise Bayesian distribution P´´ that has been conditioned on E_2 is P´´(H) = 1.

Inertia occurs when the following assumptions hold:

(i) P(H) = [0, 1].

(ii) E neither implies nor contradicts H. In other words, E is not deductively conclusive evidence for or against H.

(iii) $P_i(E) > 0$ for $\underline{P}$ and $\overline{P}$. This condition entails that neither extreme member is eliminated when conditioning on E.

Suppose that $\overline{P}(H) = 1$ and $\overline{P}(E) > 0$. According to the updating procedure in Imprecise Bayesianism, this function cannot be eliminated from the committee. Furthermore, $\overline{P}$ cannot be altered by conditionalization, because a statement with an extreme probability is probabilistically independent of any other statement. Thus, when $\overline{P}$ is conditioned on E, the probability of H is unchanged, such that $\overline{P}'(H) = 1$. Analogously, $\underline{P}$ cannot be eliminated, because $P(E) > 0$. Since probability of H in $\underline{P}$ cannot be altered by conditionalization, the value is unchanged: $\underline{P}'(H) = 0$. The committee is a continuous set between these values. Consequently, the overall picture is unchanged: P′(H) = P(H | E) = [0, 1]. Hence the name "inertia", because an Imprecise Bayesian who starts in a state of absolute ignorance with respect to a hypothesis will be stuck there, except in special circumstances.

One way to understand inertia is to take the metaphor of the "committee" very literally: imagine that the committee consists of a group of Bayesians who choose intermediate opinions between two extreme members. When the probability of H is [0, 1], the extreme member $\underline{P}$ takes H to be utterly impossible and the other extreme member $\overline{P}$ takes H to be proven beyond any doubt. When the evidence either deductively implies or contradicts the hypothesis, one of the extreme members leaves the committee and the remaining members change their opinions to fit the other extreme member's view. However, if the evidence neither implies nor contradicts the hypothesis and neither member regards the evidence as impossible, then the extreme members will stay in the committee, and they will not change their views. The remaining committee members are still balanced across the continuum of degrees of belief that exist between the two extreme members' opinions, such that the committee's view on the hypothesis remains unchanged.

I shall use examples of direct inferences, since these raise fewer ancillary controversies than other non-deductive inferences like induction or analogy. If P(H) = [0, 1] and H is 'If I see a swan, then it will be white' and E is '75% of swans are white', then the posterior value for H given E will still be [0, 1]. One particularly counterintuitive aspect of inertia is that the same result obtains with alternative non-deductively conclusive evidence, like '95% of all swans are white' or '10% of swans are white'. Additionally, inertia means that (3) in Proof 2 must be false, because E is deductively independent of H, and accordingly an Imprecise Bayesian with such a maximally imprecise prior distribution would be 'stuck' unless they acquired deductively conclusive evidence regarding H.

The problem for the DI measure is that such intuitively relevant information will not add to the weight on this measure, because the degree of imprecision will not fall. For instance, in the swan example above, 'Jupiter is larger than the Earth' will be just as (ir)relevant to H as '95% of swans are white' according to the DI measure, because neither will narrow the [0, 1] interval. Our intuitions about quantities of evidence and the DI measure diverge in cases of inertia.

Susanna Rinard has suggested a possible Imprecise Bayesian response to inertia, when this problem is presented as an objection to the overall system. She suggests that the Imprecise Bayesian should completely avoid the [0, 1] interval and uses inertia as a reason to forego assigning [0, 1] values[130]. Her response seems plausible for *most* uses of imprecise probabilities in statistics and philosophy. In practice, there is typically some relevant evidence regarding any particular statement that interests us. Even if we are reluctant to assign precise probabilities in some context, we might still be willing to make an assignment with at least *some* precision.

For example, if there is a very strangely shaped Ancient Egyptian coin, I would be reluctant to assign P(H) = 0.5, where H is the hypothesis that it will land heads. However, if it has a shape that is quite similar to normal coins, then I might comfortably assign [0.05, 0.95] as my credence, because I feel that it is unlikely that tossing this coin will result in a very high long-run frequency of either heads or tails. Such an interval will not generate the problem of inertia, because the limits are non-extreme and so they can be altered by conditionalization. By assumption, I know very little about the Ancient Egyptian coin; *a fortiori*, in scientific investigations with rich background knowledge, it is very plausible that there will be some basis for an interval that is narrower than [0, 1].

However, Walley suggests a number of interesting applications for the [0, 1] interval[131]. Firstly, it is a plausible way of representing the epistemic state of an agent who has never considered the sample space before. For example, most people have never considered the possibilities associated with the Urim and the Thummim in the Old Testament.

---

130 Rinard (2013) p. 4-5.

131 Walley (1991) p. 227.

The Ancient Hebrews used these objects for divination, but their method is unknown. In the absence of more information about the Urim, the Thummim, and the divination techniques used by the Ancient Hebrews, a [0, 1] interval is a plausible representation of many people's beliefs about hypotheses like 'The Urim and the Thummim would tend to counsel against going to war'. Secondly, the [0, 1] interval is useful when there are words with unknown meanings in the hypothesis, as in 'All Fregons are zull'. In such circumstances, an agent trying to set prior probabilities would lack any sort of information about "Fregons" and "zull" with which to choose any degree of precision for their priors, and the [0, 1] interval offers a plausible means of representing this ignorance. One might argue that there is always *some* background evidence available when one encounters new words, such that weight never reaches a value of zero, but this position does not entail that this background information is sufficient to determine precision in one's priors.

Additionally, following Rinard's proposal would be incompatible with the claim that the [0, 1] interval is an example of the expressive power of Imprecise Bayesianism. For example, Joyce argues that [0, 1] is the proper probability value to assign when there is no evidence (and thus no weight) instead of the value of 0.5, because it is well-known that assigning 0.5 will lead to paradoxes[132]. Thus, forbidding the [0, 1] interval comes at a price for Imprecise Bayesians, since it requires abandoning one reason why their system is attractive for measuring weight: in an Imprecise Bayesianism that includes the [0, 1] interval as an option, one can consistently represent a total lack of weight using a value that is

[132] Joyce (2003) p. 171. It is easy to generate such paradoxes where the [0, 1] interval is a better assignment. For example, suppose that $\{\Phi, \Psi, \chi\}$ are a set of contraries. I cannot consistently assign 0.5 to all of them. In contrast, I can coherently assign [0, 1] to each of them. Such a maximally imprecise assignment simply asserts that, for each member of this set, the committee includes functions that assign a value r to that statement, for all values r such that $0 \leq r \leq 1$.

unavailable to precise probabilists.

One might add/subtract a constant C from the limits of [0, 1] to represent the sort of strong ignorance that Joyce is discussing when using the DI measure, so that it is represented as [0 + C, 1 – C]. However, this method would arbitrarily conflate the total absence of weight with instances of the value [0 + C, 1 – C] when this value involves some quantity of evidence. Furthermore, assuming that C is a real value, there will be some value r such that r < C. Thus, according to a DI measure that has [0 + C, 1 – C] as its minimum value, a hypothesis whose probability given the total evidence is [0 + r, 1 – r] will have negative weight, and this seems nonsensical. One might try to use an infinitesimal value for C, but this would come at the cost that the co-domain of the DI function would no longer be the set of real numbers, and it is not clear how the outputs of such a function could then be incorporated into formal epistemology and decision theory.

Finally, it is plausible that there is a minimum value for weight that corresponds to absolute ignorance. Rinard's proposal would eliminate the possibility of using DI = 1 – (1 – 0) = 0 as this minimum value. Whatever the general merits of Rinard's proposal as a defence of Imprecise Bayesianism from the problem of inertia, it would not help the DI measure, because it would avoid inertia at the cost of losing the use of a minimum value for weight.

Another Imprecise Bayesian response to inertia is to modify the update rule. Such a change requires abandoning normal conditionalization. Joyce considers the possibility of simply ignoring the probability functions at the extreme ends of the set when one updates the [0, 1] interval[133]. This response has several problems. Firstly, this sort of response has not yet

[133] Joyce (2010) p. 291-292

been developed into a comprehensive and tested updating rule; as Bradley notes, it constitutes a leap in the dark[134].

Secondly, Joyce himself notes that modifying conditionalization to sometimes allow the exclusion of extreme members is an *ad hoc* departure from normal conditionalization, since it lacks any motivation other than the avoidance of inertia. It would be preferable to have some non-*ad hoc* means of escaping this problem.

Thirdly, Joyce raises the problem that such an updating method would involve an additional and significant element of arbitrariness. There is no formal reason to remove the "extreme" probability functions rather than the other, "moderate", probability functions in the set. If we excluded the moderate functions, then there would still be two extreme probability functions $\underline{P}(\mathrm{H}) = 0$ and $\overline{P}(\mathrm{H}) = 1$; the committee containing these functions still provides an interval of [0, 1]. Joyce argues that such updating constitutes a non-evidence based inference, because there is no reason in the evidence itself (from an Imprecise Bayesian standpoint) for excluding the extreme members.

An Imprecise Bayesian who adopts this new updating rule might respond to Joyce's arbitrariness criticism in the following way: the reason to exclude $\underline{P}$ and $\overline{P}$ in this case is that E is intuitively relevant to H, but according to these extreme functions, it is irrelevant. After all, the apparent relevance of the evidence is why there is a problem of inertia in the first place: when E is not intuitively relevant to H, there is no paradox.

134 Bradley (2015).

However, this appeal to intuition does not avoid the charge of arbitrariness. Firstly, modifying the updating rule to take intuitions about evidence into account introduces a new role for exogenous judgements into the framework. By “exogenous judgements”, I mean decisions that are not determined by the formalism of the Bayesian model. Of course, there is always some degree of exogeneity in formal epistemology. For standard Subjective Bayesians, the most notable exogenous judgement is the choice of a prior distribution, since neither coherence nor conditionalization uniquely determines a prior distribution. Yet Bayesians can correctly assert that, once a prior distribution has been set and conditionalization has been adopted as the updating rule, then every judgement involved in updating a hypothesis by that evidence is determined by mathematics and logic.

For some philosophers, this seems to be one of the attractions of Bayesian confirmation theory. If a Bayesian is applying their confirmation theory to problems in areas like the philosophy of economics or medicine, then they might hope to reach a point where all the philosophical issues become computable. There are obviously many other reasons to become a Bayesian, but rendering important philosophical issues into questions of calculation would be a rare and precious prize in philosophy. Indeed the computability of important issues in Bayesianism, once certain conditions have been satisfied, can give the attractive impression that these questions can be resolved objectively and uncontroversially.

However, suppose that Imprecise Bayesians adopt an updating rule according to which one excludes the extreme members in inertia when (and only when) it is *intuitive* that they should be excluded. This means that exogenous judgements are involved in both the prior distribution and the updating rule. To call something ‘arbitrary’ typically means that it is a matter of personal judgement, rather than a matter of rules, and so introducing an updating rule that involves intuition would be arbitrary. This arbitrariness is not trivial: introducing arbitrariness into the updating rule would undermine one of the appeals of Bayesianism. For the project of measuring weight, it would entail that some values of the DI measure would be arbitrary: if $P(H) = [0, 1]$ and $P(H \mid E) = [0 - C_1, 1 - C_2]$, where C_1 and C_2 are arbitrarily selected constants, then the weight for H given E would be arbitrary, since different choices of C_1 and C_2 would provide different values on the DI measure. Finally, even if a systematic version of such a rule is possible, the DI measure as it currently exists would still be inadequate, since it uses probabilities from the Imprecise Bayesian system.

Another possible response would be to contest the intuition: if there is no relevant background knowledge regarding H, then perhaps it is intuitive that non-deductive evidence will fail to modify H's probability. Many confirmation theorists now believe that background knowledge is essential to non-deductive inference. However, this response is an acceptance of inertia, rather than an elimination of the phenomenon, because it does not deny that the evidence in inertia scenarios is *evidentially* relevant (though not probabilistically relevant) to the hypothesis. The problem for the DI measure remains: even if inertia is an acceptable part of the Imprecise Bayesian formalism, it still entails that intuitively relevant evidence is unregistered by the DI measure. Regardless of whether inertia is a genuine flaw in Imprecise Bayesianism, it remains a severe problem for Walley's measure of weight.

2.5 The Problem of Corroborating Evidence

I discussed the Problem of Corroborating Evidence as a challenge for probabilistic definitions of relevance in Chapter 1. The essential idea is that sometimes a piece of evidence is relevant because it corroborates existing data, yet it does not alter the conditional probability of the hypothesis. For instance, imagine that scientists test the universal generalization H ‘All metal rods expand when heated’ via an experiment in which a rod is gradually compressed by a machine while it is heated. Suppose that E states that ‘The tested rod did not expand when heated’. Assume that P(E) > 0. Since E is inconsistent with H, it follows that P(H | E) = 0. The weight, using the DI measure, is:

$$DI(H \mid E) = 1 - (1 - 0) = 0$$

However, the scientists could repeat the experiment; intuitively this would provide more evidence. E´ might state that another rod also did not expand when heated and compressed. The probability of H would be unchanged, since an extreme probability cannot be changed by adding evidence, but E´ still seems to be relevant to H, because corroborating evidence with independent sources of information will often increase the reliability of the original evidence. In such circumstances, the DI measure will not reflect the increases in weight.

An Imprecise Bayesian might respond that while E´ does not increase H’s probability, it does increase the probability of E. Thus, even assuming that P(H | E) = 0, we can still alter the probability of H by learning E´. If P(E) < 1, then it is possible that P(E | E´) > P(E) and P(H | E´) < P(H). However, this claim could only be true if the scientists have not yet accepted E. Once E has a probability of 1, the probability of H will be zero and thus it cannot be altered by conditioning on E´. This is not necessarily problematic for Bayesian updating: it is plausible that if one accepts E and E is inconsistent with H, then one should assign a probability of zero to H and this should not alter when additional evidence E´ is acquired. However, the problem for measuring weight remains, because E´ adds to the total quantity of relevant evidence regarding H, even though it does not alter the probability of H, but the DI measure does not register this fact.

One alternative Bayesian approach to modelling learning is Jeffrey Conditionalization[135]. In this method of updating, P(E) is altered to an exogenously determined value r such that $0 \leq r \leq 1$. For any statement H in the domain of P, the posterior probability P´(H) is determined by the following rule:

Rule of Jeffrey Conditionalization: $P´(H) = P(H \mid E)\ P´(E) + P(E \mid \neg H)\ P´(\neg E)$, where $P´(\neg E) = 1 - P´(E)$.

P´(E) and P´(¬E) are the probabilities of E and ¬E in the new probability distribution P´. Thus, upon exogenously determining that E has a probability r, $P´(E) = r$ and $P´(\neg E) = 1 - r$.

This rule contains ordinary conditionalization as a special case in which $r = 1$. When $r < 1$, it offers an interesting alternative to conventional Bayesian updating. For example, suppose that $P(E) = 0.5$. Something exogenous changes, such that $P´(E) = 0.9$. Using Jeffrey Conditionalization, we can alter the probability distribution to reflect this change, and thus we can model ‘learning E’ in a weaker sense than ‘becoming absolutely certain that E is true’.

[135] Jeffrey (1965).

This will enable the Imprecise Bayesian to avoid the Problem of Corroborating Evidence under some assumptions. Suppose that H and E are inconsistent. Suppose that E confirms E´ in all members of the committee. Assume the following values:

(1) $P(E) = [0.75, 0.85]$

(2) $P(E') = [0.1, 0.9]$

(3) $P(H) = [0.3, 0.5]$

(4) $P(H \mid E) = 0$

(5) $P(E \mid E') > P(E)$

Assume that, as a result of an experiment in which the event described by E´ occurs, its probability shifts to [0.8, 0.9]. From (1), (2), and this new value for the probability of E´, Jeffrey Conditionalization requires that this change must result in an increase in the probability of E in the new distribution:

(6) $P'(E) > P(E)$

From (3), (4), and (5), we know that an increase in the probability of E must reduce the probability of H, so that:

(7) $P'(H \mid E') < P(H)$

Therefore, by modelling the process of learning using updating via Jeffrey

Conditionalization rather than ordinary conditionalization, it is possible to corroborate existing evidence that is inconsistent with a hypothesis and thereby reduce the probability of that hypothesis.

While this will work for some cases of corroborating evidence, it will not always work. There is nothing in Jeffrey Conditionalization that excludes the possibility of assigning $P'(E) = 1$ when modelling the learning event. It would still be permissible, in Imprecise Bayesianism, to assign $P'(E) = 1$. Thus, since $P(H \mid E') = 0$, it follows that $P'(H) = 0$. From this point, the Problem of Corroborating Evidence re-emerges: no amount of corroborating evidence will increase the weight of H given the total evidence, but the quantity of relevant evidence has increased. Consequently, there is a type of evidence that the DI measure does not register.

SECTION 3: EVIDENTIAL PROBABILITY AND WEIGHT

I shall now examine an alternative proposal for using imprecise probabilities to measure Keynes's concept of weight. Kyburg discusses weight on several occasions: for example, he makes the negative claim that the probability theories of Keynes and Carnap cannot provide a formal measure of weight; he argues that this shortcoming is a weakness of their systems[136]. He also makes the positive claim that his own system of Evidential Probability *can* serve as the basis of a measure of weight. Originally, in *Probability and the Logic of Rational Belief* (1961), he claims that evidential probabilities represent both (1) the

[136] Kyburg (1961) p. 52.

support that a set of statements Γ gives to Φ and (2) the quantity of relevant evidence that Γ provides regarding that Φ[137]. If this claim were true, then evidential probabilities would be comparable to vectors. (A vector has both a magnitude and a direction, in contrast to a scalar, which has only a magnitude.) Analogously, evidential probabilities would state both an evidential relation and the quantity of relevant information in that relation..

In his article "Bets and Beliefs" (1968), Kyburg makes the more modest claim that differences in the quantity of "evidence of the same sort" could be measured using the degree of imprecision of evidential probability intervals[138]. In this section, I shall mostly criticise his 1968 claim. Since his 1961 claim entails the 1968 claim, my reasons for rejecting the 1968 claim will also constitute reasons for rejecting the 1961 claim.

Kyburg's proposed measure uses the imprecision of evidential probabilities to measure weight:

$$Kyburg's\ measure\ of\ weight = WK(H \mid K) = 1 - (y - x)$$

- where *x* is the lower bound of the evidential probability of H given K and *y* is the upper bound. I have modified Kyburg's proposal in the same way that I modified Walley's in order to make the discussion more verbally felicitous.

[137] Kyburg (1961) p. 225.

[138] Kyburg (1968) p. 63.

I shall now review the problems from the previous section. I shall conclude that the WK measure is better than DI measure for some of these problems, but not all of them.

3.1 Dilation

In Imprecise Bayesianism, the problem of dilation occurs when an Imprecise Bayesian has a precise probability for Φ, a precise probability for Ψ, and they learn that (Φ ↔ Ψ). The imprecise probability will dominate the precise probability, so that P(Φ) becomes imprecise. In contrast, the precise probability will always be chosen in Evidential Probability, such that X will stay precise and Y will become precise. Thus, unlike Imprecise Bayesianism, dilation is not a feature of Evidential Probability. Some imprecise probabilists have noted that Evidential Probability avoids dilation[139]. However, it has not been previously noted that this is an advantage for the WK measure.

This contrast is a result of a difference in the method of updating. Imprecise Bayesians update by using conditionalization for each member of the continuous set of probability functions, as described in Section 2. By comparison, Evidential Probabilists update via direct inference using the relative frequency data that has been selected by the rules for Sharpening reference classes, which I described in Chapter 1 Subsection 5.3. One of these rules is Sharpening by Precision. This rule requires that, given the choice between two otherwise unsharpenable reference class statements[140] R_1 and R_2 for which R_1 provides an interval that is a proper subinterval of R_2, an Evidential Probabilist will use R_1 rather than R_2

[139] Pedersen and Wheeler (2014) p. 1307.

[140] A reference class statement is either a claim about the relative frequency of a predicate in a reference class or a claim that an object is a member of a reference class.

to determine the evidential probability. For example, imagine that you are calculating car insurance premiums for Mr. Smith. You know that Mr. Smith is a young man and that 2-10% of young men crash their car within 5 years. You also know that Mr. Smith has undergone special driving courses and he has acquired driving qualifications; among those with these qualifications, only 3-5% of people (of any age or gender) crash their car within 5 years. However, your full distribution does not contain the joint distribution for young men who have obtained these qualifications. Therefore, you cannot use a known full distribution to Sharpen by Richness. Furthermore, neither reference class is a subclass of the other. Therefore, you cannot Sharpen by Specificity. If there are no rival reference classes, then Sharpening by Precision requires that you use the second reference class, so that the evidential probability is [0.03, 0.05].

To illustrate how this rule addresses the problem of dilation, I shall return to the example from Subsection 2.3. Imagine that I am about to randomly select a card from a pile of 40 cards:

Key

H: The card is red.

X: The card is even.

Y: The card is red if and only if it is even.

K: My total relevant evidence, in which the best reference class information for X's probability is that 1/2 of the cards in the pile are red, whereas the proportion of even-numbered cards r in the pile is only known to be $0 \leq r \leq 1$.

(1) $EP(H \mid K) = [0.5, 0.5]$

(2) $EP(X \mid K) = [0, 1]$

Imagine I learn Y. I must now decide how the probabilities of H and X will change in response to this information. I can either use my information about the proportion of red cards in the pile or my information about the proportion of even-numbered cards. I must choose between the following equations:

(3) $EP(H \mid Y \wedge K) = [0.5, 0.5]$

(4) $EP(H \mid Y \wedge K) = [0, 1]$

Since [0.5, 0.5] is a proper subinterval of [0, 1], Sharpening by Precision requires selecting (3), so there is no dilation in this scenario and the WK measure provides the intuitively sound output that learning Y has not affected the weight of H with respect to the total evidence:

$$WK(H \mid K) = 1 - (0.5 - 0.5) = 1$$

$$WK(H \mid Y \wedge K) = 1 - (0.5 - 0.5) = 1$$

By the same reasoning, *mutatis mutandis*, the evidential probability for X becomes precise, because I now know that the card is even-numbered if it is red and that 1/2 of cards in the pile are red. Hence:

(5) $EP(X \mid Y \wedge K) = [0.5, 0.5]$

Adding Y to my total evidence has increased the weight of argument according to the WK measure:

$$WK(X \mid K) = 1 - (1 - 0) = 0$$

$$WK(X \mid E \wedge K) = 1 - (0.5 - 0.5) = 1$$

These are both intuitive results. The WK measure tracks the intuition that Y is relevant to X, but irrelevant to H. By using Evidential Probability to measure weight, it is possible to avoid one of the problems of the DI measure.

3.2 Inertia

Inertia is the feature of Imprecise Bayesianism that an Imprecise Bayesian reasoner might learn some intuitively relevant evidence, but their updated probability will be equal to their prior probability. For the DI measure, the problem is that there will be no increase in the intervals' precision and thus no increase in weight according to the DI measure. One possible answer is to create an exception to standard conditionalization, but the suggested alternative updating methods are arbitrary and *ad hoc*.

Unlike the Imprecise Bayesian, an Evidential Probabilist can have a rule-governed and independently motivated escape from inertia. Indeed, even rival imprecise probabilists agree that Evidential Probabilists avoid the problem of inertia[141]. Imagine that there is a card that will be drawn from a deck using a randomized procedure. Prior to learning the composition of the deck, your best statistical information regarding the card is the statement:

K: The card will be drawn from a deck of cards. Either 0%, 100%, or some intermediate proportion of these cards are red.

H is the hypothesis that the card is red. Thus:

$$EP(H \mid K) = [0,1]$$

Subsequently, you learn:

E: 50% of the cards in the deck are red.

Since [0.5, 0.5] is a proper subinterval of [0, 1], Sharpening by Precision requires that:

$$\mathrm{EP}(\mathrm{H} \mid \mathrm{E} \wedge \mathrm{K}) = [0.5, 0.5]$$

[141] Levi (2007) p. 265.

Kyburg's system thus avoids inertia in this example, because any precise reference class data will be more precise than the [0, 1] interval and Sharpening by Precision requires that we must use this data for the new probability. A further virtue of this solution to inertia is that it is non-*ad hoc*, because Kyburg's rules are the product of an independent project of developing a formal logic of direct inference.

I shall return to the Swans example from Subsection 2.4 to further illustrate this point. Here, H is 'If I see a swan, then it will be white' and E is '75% of swans are white'. Assume that EP(H | K) = [0, 1], which corresponds to P(H) = [0, 1] in the Imprecise Bayesian version of this scenario. Suppose that I learn E and that (E ^ K) provides no better basis for the probability of H other than the reference class statement that '75% of swans are white'. Sharpening by Precision requires that:

$$\mathrm{EP}(\mathrm{H} \mid \mathrm{E} \wedge \mathrm{K}) = [0.75, 0.75]$$

The interval has narrowed, so that the weight has increased on the WK measure. In contrast to the DI measure, the WK measure can register increases in the quantity of relevant evidence that occur when non-deductively conclusive evidence is discovered in a state of absolute ignorance.

As a final note on inertia, Isaac Levi criticises the Evidential Probabilist's avoidance of inertia as "*creatio ex nihilo*"[142]. His worry seems to be that Kyburg's system makes it possible to move from a state of complete ignorance to a precise probability distribution for H, without making a presupposition about the joint distribution of H and E in a prior

[142] Levi (2007) p. 265.

distribution. If we interpret imprecise probabilities as sets of probability functions (as Imprecise Bayesians interpret them) then Levi has a point. In Imprecise Bayesianism if I assign P(H) = [0, 1], then I am not dismissing any possible fraction as the prior probability of H. For the reasons I described in Subsection 2.4, I cannot conditionalize upon some non-deductively conclusive evidence to shift my credence to a more committal state. Thus, Levi would be right that Kyburg's avoidance of inertia is counterintuitive, *if* evidential probabilities were the same as Imprecise Bayesian probabilities and *if* conditionalization is the only possible method for updating a probability distribution.

Neither antecedent is true. As I have described, Evidential Probability updating does not proceed by conditionalization. (I discuss arguments that conditionalization is the only rational form of updating in Chapter 3 Subsection 5.1.) Furthermore, evidential probabilities are very different from Imprecise Bayesian probabilities. Despite the lexicographical similarity between (i) and (ii):

(i) $P(H) = [0, 1]$

(ii) $EP(H \mid K) = [0, 1]$

- these are actually two very different claims. The claim (i) makes an assertion about a set of precise Bayesian functions. The claim (ii) makes an assertion about the formal relations between the statements H and K, where these formal relations are determined by Kyburg's rules for selecting reference classes. For the two systems of imprecise probability, there is a difference both in how these equations are obtained and how they are used, so that there can be an evidence-based reason to change an evidential probability that would not suffice to change an Imprecise Bayesian probability. Consequently, Levi's objection to the Evidential Probabilist solution of inertia is mistaken.

Returning to the topic of weight, the WK measure performs better than the DI measure in both dilation scenarios and inertia scenarios. The contrast in performance is due to fundamental differences between the relevant imprecise probability systems, and in particular in the updating method.

3.3 The Hollow Cube

While WK overcomes some of the challenges of DI, there are still serious problems with the measure. One flaw involves Teddy Seidenfeld's Hollow Cube scenario[143]. Seidenfeld presents his scenario in order to demonstrate that Evidential Probabilists cannot always make use of Bayesian statistical methods and that the intervals of Kyburg's system can widen. However, he does not note the Hollow Cube scenario also proves that WK is a problematic measure of weight.

In Seidenfeld's scenario, we are measuring the volume of a hollow cube. We hypothesise that the cube has a volume of V millilitres, where V is an interval-value. Assume that we have two available measurement methods:

(1) We could fill the cube with a liquid of a known density. We can then calculate the conditional probability that the cube will have a volume V given that it has been filled by the measured quantity of that liquid.

[143] Seidenfeld (2007) p. 276-277.

(2) We could cut a rod that has a length equal to the cube's edge and measure the length of this rod. We can then calculate the conditional probability that the cube will have a volume V given the results of this cutting and measuring procedure.

Seidenfeld notes that Bayesians can always combine these results to calculate a posterior probability for the hypothesis: it is simply a matter of using the relevant priors and likelihoods from our existing full distribution to calculate the conditional probability of the hypothesis given the conjunction of the measurements' results. In contrast, using Evidential Probability, we can only use these Bayesian methods if there is relative frequency data that provides a conditional probability for the hypothesis given *both* measurements. However, it is possible that such rich information will not be available. Potentially, a very wide interval might result from applying the rules of Sharpening. If the report of (2) is added after the report of (1) or *vice versa*, then the evidential probability can become wider as more evidence is acquired.

For example, it is possible that using method (1) produces a measurement that strongly indicates that the cube's volume is in the interval V, whereas using method (2) produces a measurement that strongly indicates that the volume is *not* in the interval V. I shall use the following abbreviations:

Key

E_1: The estimate for the volume of the cube from the measurement that used the liquid.

E_2: The estimate for the volume of the cube from the measurement that used the rod.

H: The hypothesis that the volume of the cube lies within a particular interval V.

K: Our background knowledge. By assumption, K does not provide sufficiently rich statistical information to use Bayesian methods to calculate a conditional probability for H given E_1 and E_2. However, it does contain enough information to enable the calculation of probabilities for H given E_1 and H given E_2 via classical statistical methods.

Since we have assumed that Bayesian inference that only uses known frequencies is impossible, Evidential Probability requires that we use confidence-interval methods from classical statistics (as the necessary background knowledge is available in K) as an alternative[144]. Suppose that H is extremely unlikely given E_1 and that we are using a confidence level such that the inference that ¬H has a ± 2% margin of error. Assume that:

(1) $EP(H \mid E_1 \wedge K) = [0.01, 0.03]$

With comparable assumptions, suppose that E_2 provides an evidential probability that is almost a mirror image, such that H is very likely given our measurement using the rod:

(2) $EP(H \mid E_2 \wedge K) = [0.95, 0.97]$

Sharpening by Richness is not available in this case, since we have assumed that there is no available joint distribution for ($H \wedge E_1 \wedge E_2 \wedge K$). Sharpening by Specificity is not available, since neither (1) nor (2) is based on a reference class that we know to be a subset of the other. Finally, the Sharpening by Precision rule requires that the cover of the intervals becomes the new evidential probability, since neither the interval in (1) nor the interval in (2) is a proper subinterval of the other. Hence:

[144] Kyburg and Teng (2001) p. 263-265.

(3) $EP(H \mid E_1 \wedge E_2 \wedge K) = [0.01, 0.97]$

Therefore, if we learned E_1 and obtained (1) as the probability of H, then subsequently learning E_2 would result in a much wider imprecise probability.

Seidenfeld's example demonstrates that is possible to acquire new evidence that is intuitively relevant, but which increases the degree of imprecision of the evidential probabilities. If we use the WK measure, we can obtain counterintuitive results: learning E_2 after E_1 can actually reduce the weight of argument. This presents a powerful counterexample to the WK measure of weight, since combining different types of measurement methods is a perfectly normal part of science.

Kyburg accepts Seidenfeld's example without objection[145]. Seidenfeld's example is not problematic *in itself* for an Evidential Probabilist: if we discover that our initial probability for H was radically dependent on our choice of measurement method, then a less precise interval seems to provide an appropriate representation of our greater uncertainty upon learning this dependence. Informally, our evidence is providing a much murkier picture regarding the hypothesis. Insofar as Kyburg is aiming to formalise evidential relations, there does not seem to be anything problematic here: sometimes, learning new evidence can create greater uncertainty, and this is represented in this case by the increase in imprecision. The Hollow Cube nevertheless poses a problem for Kyburg's measure, because the measure is registering what is clearly new evidence as a reduction in the quantity of evidence.

[145] Kyburg (2007) p. 289-290.

The Hollow Cube involves multiple forms of evidence. Kyburg seems to have anticipated such problems, because he claims that evidential probabilities can only be used to measure weight in the context where the evidence is the “same sort”. Seidenfeld’s scenario therefore goes beyond Kyburg’s intended area of application of the WK measure, at least under his 1968 restriction. However, in the next section I shall argue that Kyburg’s restriction does not avoid all the problems for the WK measure.

3.5 The Problem of Corroborating Evidence

The WK measure does not avoid the Problem of Corroborating Evidence, where it behaves in a very similar way to the DI measure. Suppose that, as a result of many experiments, scientists accept that there exists at least one metal rod that was heated but did not expand. For example, they might have used magnets or vices to prevent the expansion of rods. There will be statements that are inconsistent with this existential claim; one will obviously be the denial of existence of such a rod. Repetitions of this experiment seem to be relevant to this denial of existence.

Key

H: There are no metal rods that are heated and do not expand.

E_1: The exists a metal rod that was heated but did not expand.

E_2: In additional experiments, metal rods were headed and sometimes were measured as not expanding.

Since E_1 is inconsistent with H, the scientists' acceptance of E_1 reduces the evidential probability of H to zero:

(1) $EP(H \mid E_1 \wedge K) = [0, 0]$

- and the weight is -

(2) $WK(H \mid E_1 \wedge K) = 1 - (0 - 0) = 1$

Learning E_2 and conjoining it with (E_1 ^ K) will not provide a different value to (1). Therefore, the WK measure value for H given (E_1 ^ E_2 ^ K) will not be different from (2). The WK measure once again fails to match our intuitions about evidential relevance: acquiring corroborating evidence will not register as an increase in evidence. The Problem of Corroborating Evidence remains. Additionally, the evidence here is entirely of the same sort, because it consists simply of repetitions of a type of experiment.

One might object that, even on the basis of multiple experiments, scientists would never accept a statement like E_1, so the probability of H given the total evidence would never drop to zero. However, unless one forgoes the idea of accepting evidence altogether[146], there will be *some* statements that are inconsistent with the evidence that one has accepted, and in either Evidential Probability or Bayesianism, the statements will have a conditional probability of zero given the evidence. Such statements need not constitute anything that one

[146] Kyburg, at least, would not take such a position: see Chapter 5 Section 2.

might call a “scientific theory”, because the problem is that even after one has accepted evidence that is inconsistent with these statements, one can still acquire further evidence about them.

Kyburg might respond that the value of ‘1’ for the WK measure is merely an idealization; it just reflects a quirk of how accepted evidence is treated in many probabilistic models of scientific reasoning. Certainly, the Problem of Corroborating Evidence is not a great problem in itself. In many interesting applications of the measure, such as testing unrestricted statistical generalizations like ‘51% of human births are male’ or ‘95-100% of applications of this insecticide will result in the infertility of wheat in the area of application’, finite samples cannot falsify the hypotheses in question. Similarly, in experimental contexts, where scientists must use fallible instruments, the distribution of measurement errors in the instruments means that while experimental results might be extremely improbable given a theory and auxiliary hypotheses about the instruments involved, the results are still *logically* consistent with the theory. If the WK measure is applied to such contexts, then the Problem of Corroborating Evidence will not arise. The problem only occurs in relatively trivial scenarios, where one statement (which need not be a ‘scientific theory’ in any but the broadest sort of sense of this term) is accepted and another is inconsistent with the accepted statement.

Yet this response does not dismiss the Problem of Corroborating Evidence entirely. Instead, it gives us reason to regard the problem as a minor flaw of the WK measure. In particular, someone using the WK measure will have to be alert to the idealization involved in the value of ‘1’. It would be preferable to have a measure of the weight of argument that did not have this feature (and was otherwise at least as satisfactory as the WK measure) but if such a measure is unavailable, then very limited scope of the Problem of Corroborating Evidence does not give sufficient reason to reject the WK measure. In short, the Problem of Corroborating Evidence is a small imperfection of the measure, in contrast to the more worrying problems that the WK measure faces in the Hollow Cube case where multiple sorts of evidence are involved.

Using Evidential Probability as the basis of our quantitative measure of weight can avoid some of the problems that the DI measure faced, but not all of them. As such, it performs better in some applications, but does not offer a general quantitative formalization of the weight of argument, even within the restricted field of application that Kyburg suggested. An alternative Evidential Probabilist measure might involve combining a quantitative measure from information theory with the definition of relevance that I developed in Chapter 1.

CONCLUSION

Keynes's concept of weight is an interesting notion, but it is extremely difficult to formalize. The results of my discussion confirm Keynes's scepticism that weight is quantitatively measurable. It might be possible to measure the quantity of relevant evidence using imprecise probabilities, but neither Walley's measure nor Kyburg's measure succeeds. The DI measure works well in some simple contexts, but it faces a number of problems that make it unsatisfactory as a measure of weight. Kyburg's WK measure avoids some of the strange outputs of the DI measure, but it is still imperfect. In particular, as Kyburg hinted, it has a limited scope of acceptable application: it does not work well when applied to different sorts of evidence, as I illustrated using Seidenfeld's Hollow Cube scenario.

An alternative option is to combine the qualitative definition of evidential relevance that I developed in Chapter 1 with a suitable metric from information theory. Such a measure of weight would first identify relevant data in the premises using my qualitative definition. We could apply a measure of information to this data and thus obtain a measure of the weight of an argument. Assuming my analysis of evidential relevance in Chapter 1 is adequate, the only remaining task in developing this measure is the selection of a satisfactory and applicable measure of information.

Despite my criticisms of their measures, I agree with Walley and Kyburg that there is *something* epistemically significant that is reflected in the width of imprecise probability intervals. I disagree that this "something" is the weight of argument, but there are other possible uses for degrees of imprecision. In the next chapter, I shall argue that we can use imprecise probabilities to measure the extent that our choice of expected utilities for decision-making involves "going beyond the evidence". I shall use this measure to develop a novel answer to a paradox in decision-theory.

CHAPTER 3: EVIDENTIAL PROBABILITY AND THE ELLSBERG PARADOX

In the previous chapter, I argued that the degree of imprecision of either Imprecise Bayesian or Evidential Probabilist intervals does not provide a general measure of the weights of arguments. However, I agree with philosophers like Kyburg and Walley that there is *something* important tracked by the width of these intervals. In this chapter, I shall argue that the degree of imprecision of evidential probabilities can be used to formulate a novel answer to the Ellsberg Paradox. This paradox challenges the standard assumption that expected utility is sufficient to identify rationality in all decisions, because it involves a type of decision problem in which most people consistently violate the axioms of maximizing expected utility, yet there is no received account (independent of those axioms) as to why their decision-making is problematic. I shall focus on providing reasons for someone attracted to the standard decision theory framework to adopt the decision theory that I develop in this chapter.

In Section 1, I introduce the standard approach to normative decision theory. In Section 2, I explain the Ellsberg Paradox. In Section 3, I provide a brief overview of the main types of response to the paradox. In Section 4, I critically assess Kyburg's answer to the Ellsberg Paradox, which takes advantage of the imprecision of his system. In Section 5, I propose an alternative Evidential Probabilist decision theory. Finally, in Section 6, this decision theory is applied to the Ellsberg Paradox. I argue that my answer avoids the paradox, while also preserving standard decision theory for most static decision problems. Thus, I provide a fairly conservative answer to the Ellsberg Paradox. My answer will be interesting for decision theorists who are attracted to standard decision theory, but who are dissatisfied by its treatment of the Ellsberg Paradox. Like Kyburg's response to this problem, my answer uses Evidential Probability, and I argue that it improves on his answer in some respects.

SECTION 1: MEU DECISION THEORY

1.1 Normative and Descriptive Decision Theory

Decision theory is the study of rational choice. In particular, it can be broadly defined as the theory of rational choice[147]. It is usually divided into two categories: (1) descriptive decision theory, the investigation of how people *actually* make decisions, and (2) normative decision theory, the investigation of how people *ought* to make decisions. One can also view this split as a difference in the interpretations of the same theory[148].

Whether a topic is treated as a descriptive or normative issue is extremely significant. Nevertheless, empirical facts about people's actual decisions can be relevant in either form of decision theory. If people's behaviour conflicts with a normative decision theory and they do not regard their choices as irrational when they are given an evaluation using that decision theory, then this empirical fact is a *prima facie* problem for that theory. Furthermore, our cognitive capacities set a limit on what constitutes rationality for us: it would be problematic if a normative decision theory were a mere counsel of perfection and lacked clear

[147] Steele and Stefánsson (2015).

[148] Michael D. Resnik (1987, p. 4) argues that "abstract" and "experimental" are better distinctions, with the former tending (non-uniformly) towards normative questions and the latter tending (non-uniformly) towards descriptive inquiries. However, the normative/descriptive distinction is the standard bifurcation in the literature.

implications for computationally restricted reasoners. Finally, as Hempel points out, calling an action (like a decision) "rational" involves both (a) an empirical conjecture about people's reasoning and (b) a normative appraisal of that reasoning, because such an assertion entails both that those particular reasons explain the action *and* that the reasons were sufficient for their choice to be rational in that context[149].

In the Ellsberg Paradox, an empirical fact about most people's decision-making under uncertainty poses a problem for the standard framework of normative decision theory. I shall discuss this paradox as a normative problem. My proposals will be logically compatible with a very large variety of descriptive explanations of the Ellsberg Paradox.

1.2 MEU Decision Theory

Decision theorists use idealized scenarios. A common idealization is that a decision-maker is considering an algebra of statements that is strongly deductively closed, so that it contains every deductive implication of every set of members of the algebra. Each statement in the algebra describes possible circumstances. In standard decision theory, there are the further idealizations that (1) the decision-maker's beliefs can be represented using an additive probability function that has the algebra as its domain and (2) the decision-maker has preferences that can be formalised using a utility function that takes the statements of the algebra as its domain and provides cardinal utilities as its range. (The range of a function is the set of its outputs for all the members of its domain.) A cardinal utility is a real number;

[149] Hempel (1965) p. 463.

they can be contrasted with ordinal utilities, which have only a ranked order[150].

To determine a cardinal utility, one ranks the statements in the algebra (to order one's preferences for the different possible circumstances) and arbitrarily chooses two numerical points. The chosen numbers are unique up to a positive affine transformation, which means that one could obtain different numbers by multiplying, adding, or subtracting constant values without altering the cardinal utility. Formally, this means that if U is a cardinal utility function of an agent and *u* is the value of each of the function's outputs, then a utility function U´ where $U(x) = u$, $U' = a + bu$, and $(b > 0)$ is just as satisfactory a formalisation of the agent's preferences[151]. For instance, suppose that I am standing on a river and I have dropped two sticks into the water; I might represent my utility for successfully predicting which of two sticks will travel underneath the bridge as U(S) = 0.7 and my utility for being unsuccessful as U(H) = 0.3, where S is the circumstance of making a successful prediction. However, I could perform a positive affine transformation by multiplying both numbers by a constant value of 10, such that U´(S) = 7 and U´(H) = 3, and my cardinal utilities would be unchanged.

It is important to stress that, in decision theory, cardinal utilities are not measures of intensities of pleasure, but rather they are formal representations of people's preferences. You and I might have the same cardinal utilities for each of two possibilities, but our psychological states could nonetheless be very different. Like ordinal utility functions, cardinal utility functions are still fundamentally concerned with orderings of preferences, but

[150] John von Neumann and Oskar Morgenstern, in (1953), develop a method by which such utilities could be derived by an agent who could not introspect them directly, but who had suitably rich preferences among the different circumstances and possible lotteries.

[151] Reiss (2013) p. 45.

unlike ordinal utility functions, they provide a numerical scale for these orderings[152].

The Maximizing Expected Utility (MEU) approach requires that the agent maximize expected utility, as defined below, by choosing the action with the highest expected utility. If two actions have equal expected utilities, such that the agent is *indifferent* between them, then MEU allows either action to be chosen.

Key

$X_{1\text{-}n}$: Possible states of the world.

$A_{1\text{-}n}$: Possible alternative actions that the decision-maker can perform in a particular decision-problem.

E: The total evidence available to the decision-maker.

$U(X_i, A)$: The cardinal utility of the action A given that X_i obtains.

Expected Utility

$$\sum_{i=1}^{n} U(X_i, A)P(X_i \mid E)$$

Maximizing Expected Utility: An action A_1 maximizes expected utility if and only if there is no alternative action A_2 such that the expected utility for A_2 is higher than the expected utility for A_1.

Using this rule, one can always identify a unique action or a set of actions that will be

[152] Kreps (1990) p. 30.

rational for an individual possessing the functions U and P.

I shall illustrate the MEU theory of rationality in a simple and uncontroversial example. Imagine that you have a choice between two actions:

$\underline{A_1}$ – Betting that a fair die will land on a 1, 2, 3 or 4.

$\underline{A_2}$ – Betting that the die will land on a 5 or 6.

It is assumed that your algebra distinguishes the following states of the world that are mutually exclusive and exhaustive given your evidence E:

$\underline{X_1}$: The die lands on a 1.

$\underline{X_2}$: The die lands on a 2.

$\underline{X_3}$: The die lands on a 3.

$\underline{X_4}$: The die lands on a 4.

$\underline{X_5}$: The die lands on a 5.

$\underline{X_6}$: The die lands on a 6.

Assume that your utility from winning either bet is 100 units. Among other things, E states that the die is a fair die; thus, there is a 1/6 probability of the die landing on each side, such that $P(X_i \mid E) = 1/6$ for each outcome X_i. Recalling the definition of expected utility:

$$\sum_{i=1}^{n} U(X_i, A)P(X_i \mid E)$$

- the expected utilities are 100(4/6) = 100(2/3) for A_1 and 100(2/6) = 100(1/3) for A_2. MEU requires that you choose A_1 rather than A_2, because 100(2/3) > 100(1/3). Clearly, in this scenario MEU theory corresponds to ordinary intuitions about rationality. Furthermore, the theory seems fruitful, since we can clearly extend it to situations in which most people's intuitions are not so clear or where different people disagree.

There is a gargantuan critical literature on the MEU framework, but I shall put these debates to one side and generally assume that it is worth attempting to retain either MEU or a similar decision theory, if possible. I shall also assume that, while people obviously do not always act in accordance with MEU theory, it can sometimes provide valuable formal rationalisations of what is ordinarily regarded as reasonable behaviour.

SECTION 2: THE ELLSBERG PARADOX

2.1 The Ellsberg Paradox

Daniel Ellsberg's paradox has been one of the most influential challenges to conventional MEU theory[153]. He developed the paradox as a PhD student and published in an article one year prior to completing his thesis[154]. Ellsberg presents two situations that are paradoxical for MEU. However, since a single example can illustrate the basic problem, I shall focus on his second scenario, which I shall call 'the Ellsberg Scenario'.

---

[153] Al-Najjar and Weinstein (2009) p. 249-250.

[154] Ellsberg (1961).

<u>THE ELLSBERG SCENARIO</u>

Suppose you know that there are 90 balls in an urn. Exactly 30 of these balls are red. The other 60 balls are black and yellow in some unknown proportion. In this scenario, a decision maker has two choices among bets about the colour of a ball that will be randomly selected from the urn:

Red: Betting that the ball will be red.

Black: Betting that the ball will be black.

Black or Yellow: Betting that the ball will *not* be red.

Red or Yellow: Betting that the ball will *not* be black.

The utilities are the same for winning with each bet. In experiments, most people prefer **Red** to **Black** and **Black or Yellow** to **Red or Yellow**, when given each choice between each pair of alternative actions[155]. Thus, in either choice, most people prefer the bet for which there is a precisely known relative frequency of the type of ball they are anticipating.

However, there is no probability distribution under which their behaviour satisfies the axioms of maximising expected utility under the standard axioms, because **Red** is preferable to **Black** if and only if **Red or Yellow** is preferable to **Black or Yellow** in MEU[156]. This can be demonstrated formally. By assumption, the expected utilities for each bet are equal, so that

---

[155] Ellsberg (1961) p. 653-654.

[156] Ellsberg (1961) p. 655.

the only way that an action could have a greater expected utility than another is via a higher probability. The following is a list of the relevant circumstances:

R: The ball will be red.

B: The ball will be black.

¬**R**: The ball will not be red.

¬**B**: The ball will not be black.

From the probability calculus:

(1) $P(R) = 1 - P(\neg R)$

From the composition of the urn and the random selection of the ball:

(2) $P(R) = \frac{1}{3}$

- hence -

(3) $P(\neg R) = \frac{2}{3}$

If R has a greater expected utility than B and they each have equal utilities, then P(R) > P(B). Hence, from (2):

(4) $\frac{1}{3} > P(B)$

By the probability calculus:

(5) $P(B) = 1 - P(\neg B)$

From (4) and (5), it follows that:

(6) $\frac{1}{3} > 1 - P(\neg B)$

From (6):

(7) $-\frac{2}{3} > -P(\neg B)$

From (7):

(8) $P(\neg B) > \frac{2}{3}$

Finally, from (3) and (8), it follows that:

(9) $\mathrm{P}(\neg\mathrm{B}) > \mathrm{P}(\neg\mathrm{R})$

(9) and the assumption that the utility of correctly betting that ¬B is equal to the utility of correctly betting that ¬R entail that **Red or Yellow** must be preferable to **Black or Yellow**, no matter the particular values of P(¬R) and P(¬B). This is contrary to most people's preferences, as they prefer betting that ¬R to betting that ¬B.

Thus, if someone is maximizing their expected utility and they prefer betting that R to betting that B, then they must also prefer betting that ¬B to betting that ¬R. The same follows, *mutatis mutandis*, for preferring betting that ¬R to betting that ¬B: one cannot simultaneously prefer betting that R to betting that B. Therefore, a preference for both **Red** over **Black** and **Black or Yellow** over **Red or Yellow** is inconsistent with MEU theory.

I shall demonstrate this point using a decision matrix. Firstly, to simplify the discussion, I shall use monetary values and assume (counterfactually) that these monetary values correspond unproblematically to utilities. The outcome of winning any of the bets shall be £10. From Matrix I, it can be seen that **Red** is more likely to yield £10 than **Black** if and only if **Red or Yellow** is more likely to yield £10 than **Black or Yellow,** because if the proportion of red balls to black balls is high, such that R is probable relative to B, then there must be fewer black balls relative to red balls in the urn, so that the probability of (R v Y) is greater than the probability of (B v Y).

Matrix I

	R	B	Y
Red	£10	0	0
Black	0	£10	0
Black or Yellow	0	£10	£10
Red or Yellow	£10	0	£10

Key

R: The ball will be red.

B: The ball will be black.

Y: The ball will be yellow.

Red: Betting that the ball will be red.

Black: Betting that the ball will be black.

Black or Yellow: Betting that the ball will not be red.

Red or Yellow: Betting that the ball will not be black.

Contrary to some people's intuitions, a proponent of MEU cannot use risk-aversion to explain this divergence between their theory and people's behaviour. In MEU, a difference of "risk" is a difference of probabilities, because MEU decision theories identify risk is with an agent's personal probabilities[157]. I have assumed that monetary values and utilities correspond, but this is not an essential assumption for the use of MEU. In particular, a utility-maximising agent can have preferences for both monetary values and the risks involved in decisions. For example, suppose that you are offered a choice between (a) betting that a fair die will land on a 5 or 6 with a payout of £24 and (b) a guaranteed £2, so that the expected value of (a) is 2/6(£24) = £8 and the expected value of (b) is £2[158]. A preference for (b) over (a) is consistent with MEU, because you might assign a sufficiently low utility value to the risk of losing the bet that this overwhelms the £6 difference between the bets' expected values. If your utilities had this structure, then you would be maximising expected utility by "playing it safe" and choosing (b), because you are risk-averse.

[157] Reiss (2013) p. 48.

[158] The expected value of a bet is its probability-weighted mean value in an indefinitely long series of trials.

However, as I demonstrated above, **Red** can only be riskier than **Black** if **Red or Yellow** is riskier than **Black or Yellow**, because $P(R) > P(B)$ can only be true if $P(\neg B) > P(\neg R)$. Similarly, **Red or Yellow** can only be riskier than **Black or Yellow** if **Red** is riskier than **Black**. Therefore, even if people are being risk-averse (in the MEU sense) their behaviour is still inconsistent with MEU. There does seem to be something "cautious" or "risk-averse" about the standard preferences, but if this characteristic exists, then the MEU formalisation of risk cannot capture it.

From a normative perspective, the most interesting feature of Ellsberg's scenario is that people's behaviour in the Ellsberg Scenario is (at least superficially) rationally permissible. Firstly, there does not seem to be anything *prima facie* strange with such preferences. Secondly, there are no formal problems with the Ellsberg preferences that are independent of the axioms of MEU. For example, there is no proof that people with the standard Ellsberg preferences are certain to lose money, or certain to not win money, or even less likely to win money that an MEU player. Even at a superficial level, there seems to be something that A_1 and A_3 have in common that makes them different from A_2 and A_4, so that it is conceivable that someone could reasonably regard this difference as significant for their preferences. (This is a much weaker claim than that *only* the standard Ellsberg preferences are reasonable.) In contrast, for some other violations of MEU, one can point to formal problems with the recalcitrant preferences. To summarise, there is neither a pre-theoretical intuition against the standard Ellsberg preferences, nor an independent (in the sense of not presupposing MEU theory) formal argument in favour of rejecting them as irrational. Consequently, we have a case where many people's behaviour does not match the MEU theory, but it is plausible that the problem might be the theory rather than the behaviour.

SECTION 3: RESPONSES TO THE ELLSBERG PARADOX

3.1 Ambiguity Aversion Responses

A notable empirical explanation of people's choices in the Ellsberg Scenario is the Comparative Ignorance Hypothesis of Craig R. Fox and Amos Tversky[159]. According to this hypothesis, many people have a comparative preference against ambiguous decisions. An "ambiguous" decision is one that is made without the statistical information to create a relative frequency-based probability distribution. According to this hypothesis, when people have a choice between a relatively ambiguous decision and a relatively unambiguous decision, people will choose the latter. The preference is 'comparative' in the sense that people have a linear ordering for the actions.

Another related explanation is the hypothesis that most people have an *absolute* preference for relatively unambiguous decisions. Such a preference is different from a mere comparative preference, because it implies that people will be willing to bet greater sums on the less ambiguous bets. In the context of the Ellsberg Scenario, this explanation in terms of absolute preference has the same observational prediction as the Comparative Ignorance Hypothesis: people with the preference for unambiguous decisions will opt for the less ambiguous bets, which are **Red** over **Black** and **Black or Yellow** over **Red or Yellow**.

The ambiguity aversion explanations are empirical. However, there is a similar literature of normative answers to the Ellsberg Paradox, which holds that normative decision theory should be reconciled with ambiguity-averse decision-making. Nabil I. Al-Najjar and

[159] Fox and Tversky (1995*).*

Jonathan Weinstein call this approach the “ambiguity aversion literature”[160]. This Ambiguity Aversion response to the Ellsberg Paradox combines several positions: (1) the standard preferences are rational; (2) ambiguity aversion is a factor in rational decision-making; and (3) MEU theory is modified to reconcile ambiguity aversion with the decision theory by incorporating a preference for information into the utility function. James Dow and Sérgio Ribeiro da Costa Werlang[161] provide an example of how one can carry out part (3) of this strategy: they adopt a decision theory in which ambiguity aversion is quantified via imprecise probabilities. These values are incorporated into a decision theory that is a strong departure from MEU, in the sense that the ambiguity aversion factor is a non-MEU element of decision-making that is relevant to a large variety of choices.

William A. Huber motivates such uses of an ambiguity parameter in decision theory by arguing that, when decisions are ambiguous, then there is uncertainty (in an intuitive, pre-theoretical sense) regarding the correspondence between the probabilities one uses in decision-making and the real world[162]. This distinction between probabilities based on knowledge and probabilities based on ignorance does not play a role in MEU. However, as Huber argues, there is a commonsense distinction here that an MEU analysis of the Ellsberg Paradox does not capture: the MEU framework contains probabilities as a parameter for risk, but no parameter for ignorance about the relative frequencies. In the Ellsberg Paradox, most players are averse to the choices for which the relevant proportions of balls are unknown. An ambiguity aversion approach might seem to provide a natural means of addressing this *lacuna* in MEU theory by providing the ‘missing’ parameter for such ignorance and by using

[160] Al-Najjar and Weinstein (2009) p. 250.

[161] Dow and Werlang (1992).

[162] Huber (2010) p. 374.

an apparently relevant feature of Ellsberg's scenario.

However, allowing ambiguity aversion in this way creates new paradoxes when we examine *dynamic* settings with series of choices, in comparison to the *static* situation in which there is just one choice, like the Ellsberg Paradox[163]. For instance, allowing ambiguity aversion can result in a decision theory that permits decision-makers to take sunk costs into account, which is the "Sunk Cost fallacy". A sunk cost is something that has been given up in the pursuit of a goal and that cannot be recovered. For example, in the Robert E. Howard story "The Tower of the Elephant", the adventurer Conan is breaking into a sorcerer's tower to retrieve a famous gem called the "Heart of the Elephant". Having braved several lethal perils and just survived a close fight against a giant spider, Conan finds himself in a treasure chamber. He can proceed into the next room and continue to pursue the Heart of the Elephant or he can take the wealth that he can carry (which is presumably less valuable than the Heart) from this chamber. Conan reasons that, given all the danger that he has risked thus far to steal the Heart, he will continue in his adventure and step into the next chamber in the tower[164]. Since Conan is taking past costs into account, he has committed the Sunk Cost fallacy.

To illustrate this problem, I shall slightly adapt an example from Al-Najjar and Weinstein. The set-up of the scenario is similar to the Ellsberg Paradox scenario. A randomly selected choice ball will be drawn from an urn. The player knows that the urn contains 30 red balls, 0-60 black balls, and 0-60 yellow balls. In this scenario, the player is assumed to be risk-neutral, in the sense that they do not have a preference for taking risks, nor are they risk-averse. As in the Ellsberg Paradox, we assume counterfactually that monetary values

[163] Al-Najjar and Weinstein (2009) p. 258-271.

[164] Howard (2006) p. 94.

correspond to utilities.

Unlike the Ellsberg Paradox, the ball is drawn behind a screen, so that the player cannot see which ball was selected. Before this occurs, the player can choose to invest a sum G, such that $G \geq 0$. The value of G is fixed ahead of their choice and the selection of the ball. If the player chooses to invest G and the ball is yellow, then they will receive their investment back, plus a sum of (£10 – G). If they invest G and the ball is not yellow, then they receive their investment back. In essence, G is a hedge against yellow balls.

If the ball is not yellow, then the player is offered a second choice between (1) betting that the ball is red and (2) betting that the ball is not black. If the player has invested G, then she receives (£10 – G) for correctly guessing the colour of the ball. If she guesses incorrectly, then she loses her initial investment. If the player has *not* invested G, then she receives £10 for guessing correctly and she gains/loses nothing if she guesses incorrectly.

It is never rational to invest G ≥ £10, because not investing G yields an equal or greater sum under any possible outcome when G takes this value. Similarly, it is always rational to invest G = 0, because if the ball is yellow then investing G provides a gain of £10 and not investing G provides no gain, whereas if the ball is not yellow then investing G provides equal gains or losses to not investing G. The rationality of choosing to invest a sum 0 < G < £10 will depend on a player's probabilities and utilities.

The player might face two decision problems:

Choice 1: Whether to invest a given value of G as a hedge against yellow balls.

Choice 2: If the ball is not yellow, then whether to bet (a) that the ball is red or (b) that the ball is black.

I shall represent this scenario using a decision matrix for a player who has chosen to invest G. It is assumed that the players' preferences are (weakly) *dynamically consistent*, so that their ordinal preferences for different series of decisions do not change upon learning probabilistically irrelevant information. This entails that their preferences between **Red** and **Black** will be the same after learning that the ball is not yellow. It is also assumed that their preferences are *additive invariant*, so that the addition of a constant G does not change the ordering of their preferences. Therefore, if they prefer betting £10 on an outcome Φ to £10 on an outcome Ψ, then they prefer betting £10 ± G on Φ to £10 ± G on Ψ. A player who chooses not to invest G faces Matrix II, which is equivalent to the top two rows of Matrix I above. If someone has the standard Ellsberg Paradox preferences and they have not invested G, then they prefer **Red** to **Black**, because this part of Matrix II is identical to the relevant part of Matrix I.

Matrix II

	R	B	Y
Red	£10	0	0
Black	0	£10	0

However, suppose that they have chosen to invest G. The player thus faces a choice that is characterised by Matrix III.

Matrix III

	R	B	Y
Red	£10 – G	0	£10 – G
Black	0	£10 – G	£10 – G

Returning to the standard Ellsberg Paradox, we know that the player in question prefers ‘Black or Yellow’ to ‘Red or Yellow’ in Matrix IV. The only differences between Matrix III and Matrix IV are the information that the ball is not yellow and the constant G. Given the assumptions of dynamic consistency, the information that the ball is not yellow does not change the order of the preferences for the rows of the matrices. Furthermore, the constant G in Matrix III cannot change the player’s ordinal preferences, because their preferences are assumed to be additive-invariant. Since the ambiguity-averse agent now faces an identical choice problem in Matrix IV to Matrix III and they prefer **Black or Yellow** to **Red or Yellow** in Matrix IV, they must also prefer **Black** to **Red** in Matrix II. Thus, a player who has chosen to invest G and who knows that the ball is not yellow will prefer **Black** to **Red**.

Matrix IV

	R	B	Y
Red or Yellow	£10	0	£10
Black or Yellow	0	£10	£10

However, from their preference in Matrix I, we know that they would prefer **Red** to **Black** had they not invested G. Therefore, their choice to bet **Red** or **Black** *depends purely on their earlier decision to invest G*. This means that a dynamically consistent ambiguity-averse player with additive-invariant preferences will commit the Sunk Cost fallacy in Al-Najjar and Weinstein's scenario.

By contrast, a standard MEU player will prefer **Red** to **Black** in one of the matrices above if and only if they prefer **Red** to **Black** in all the matrices. For example, if P(R) > P(B), then they will prefer **Red** to **Black** in Matrix II, but they will also prefer betting **Red** to **Black** in Matrix III and Matrix IV. Their choices will be independent of the decision to invest G.

A supporter of ambiguity aversion responses might try to avoid this consequence by dropping the assumption of additive invariance. However, at least in this case, the assumption that investing G should not modify one's ordinal preferences is a plausible assumption. Furthermore, to drop it would be to accept outright that investing G makes a difference to the player's decision-making, because their preference for **Red** or **Black** would depend on the value of the constant G, which is a sunk cost. In short, this strategy would commit the ambiguity-averse player to the Sunk Cost fallacy from the outset.

Alternatively, they might contest the status of the Sunk Cost fallacy as a genuine "fallacy". For instance, Robert Nozick argues that taking sunk costs into account can be rational under *some* circumstances[165]. The circumstances which he discusses can be classified into two general categories: (1) taking sunk costs can be advantageous for an irrational agent, because they can provide a counterbalance to irrational tendencies like being tempted by short-term benefits that have greater long-term costs, and (2) our commitments to past projects is part of what provides structure to our sense of selfhood and meaning to our lives, so that it is arguably permissible for a rational agent to take them into account, even if this involves the Sunk Cost "fallacy". While (1) is not *straightforwardly* relevant to normative decision theory as a theory of rational behaviour, (2) is very relevant.

However, even if one believes that taking sunk costs into account is not fallacious in general, it seems irrational in this case, because there is nothing in the notion of ambiguity aversion that motivates a concern for sunk costs. There are no long-term projects, promises, or principles involved; the influence of the earlier decision to hedge G does not have the same significance as commitments like promising to avenge a family member's death or deciding to devote one's life to medicine; the Sunk Cost reasoning in this scenario is merely an unmotivated consequence of such a preference structure. It is one thing to believe that the Sunk Cost "fallacy" is non-fallacious when a player has an explicit rationale for such behaviour, but quite another thing to think that it is non-fallacious without such a rationale.

Nevertheless, it would be possible for me to explore this debate in more detail. I shall

[165] Nozick (1993) p. 21-25.

simply note that the scenario developed by Al-Najjar and Weinstein proves that the Ambiguity Aversion response comes at a price: a normative decision-theorist who adopts this response to the Ellsberg Paradox (and accepts dynamic consistency as a norm) no longer has to address an inconsistency between apparently rational behaviour and their theory, but they must then defend a further controversial position. Regardless of one's views on the Sunk Cost fallacy, it would be attractive *ceteris paribus* to have a response to the Ellsberg Paradox that did not incur this obligation.

Furthermore, the result in this second scenario is comparable to the Ellsberg Paradox: it seems like it should be rationally permissible to have most people's Ellsberg Paradox preferences, but still prefer **Red** to **Black** in Matrix II. However, such preferences are excluded by the Ambiguity Aversion response that Al-Najjar and Weinstein criticise. Consequently, their example suggests that both forms of decision theory face the same sort of problems with Ellsberg-type scenarios.

3.2 Conservative Responses

Supporters of the MEU theory must argue that, contrary to most people's intuitions, the standard Ellsberg Paradox choices are actually mistakes. Such responses cannot be straightforwardly formal, because one cannot prove that the standard Ellsberg preferences are always produced by a mathematical error, such as miscalculating the expected values of the various bets. Additionally, there is no proof that people who make standard choices are more likely to lose money in the long-run.

Some conservative decision theorists have attempted to explain people's behaviour as heuristic reasoning. Al-Najjar and Weinstein have recently developed a response of this nature[166]. They argue that the experimental subjects with the standard preferences are misapplying heuristics for avoiding deception. These heuristics are reasonable in non-experimental settings, but people can lead to irrational behaviour in a deception-free experimental context. For instance, if the bookie could change the proportion of the black balls from any proportion from 0 to 2/3, whereas the proportion of red balls must be 1/3, then it would make sense for people to take a precaution against this possibility by avoiding betting on the black balls. According to the conjecture of Al-Najjar and Weinstein, people fail to modify this heuristic for the circumstances in which they know that there is no deception, and this causes them to violate MEU theory.

There are strong advantages to such conservative responses. As Al-Najjar and Weinstein note, they are consistent with the experimental data, but they do not require any revision of the MEU theory[167]. Furthermore, they seem just as good (or bad) explanations, *prima facie*, as the explanations in the ambiguity aversion literature, since both accounts make an appeal to internal psychological features of the test subjects' minds in the Ellsberg Paradox experiments.

From the standpoint of normative decision theory, there are even stronger advantages of the response of Al-Najjar and Weinstein. While it is widely regarded as counterintuitive to say that people are being irrational in the Ellsberg Paradox, there are many cases in which MEU theory either (a) formalises our intuitions about rationality or (b) enables an

[166] Al-Najjar and Weinstein (2009) p. 276-277.

[167] Al-Najjar and Weinstein (2009) p. 278.

unproblematic extension of these intuitions to contexts in which our intuitions are unclear. We should be reluctant to abandon such a powerful framework, even if this means sometimes making counterintuitive judgements about rationality.

However, the Ellsberg Paradox is still worth avoiding, *ceteris paribus*. If we use the definition of 'rational' that Al-Najjar and Weinstein themselves use, then a decision is rational if it is immune to introspection: if I choose an action A and you can make me regret A by using an analysis of my reasoning that does not involve providing any new non-formal information, then it was irrational for me to choose A[168]. Here, 'formal' information involve be mathematics, logic, or the acceptance of a decision theory like MEU. 'Non-formal' information could be the result of a bet or learning a scientific fact. For example, suppose that I am playing Blackjack. I am aiming to win money, rather than enjoy risky decision-making or any other motive. Suppose that I have the 10 of Diamonds and the 2 of Clubs. I choose to "Hit" and I receive the 10 of Hearts. I shall regret my decision, but my regret is not due to learning that I made a formal error in my reasoning; I am simply disappointed that I was unlucky. In contrast, suppose that, instead of choosing "Hit", I accept a bet from the dealer that I can win by having five cards, two of which are 10's. It is easy to prove that I cannot win money with such a bet, because this outcome is impossible.

What counts as an adequate 'formal' reason to regret one's decision will vary with the context, since it will depend on the accepted formal decision theory. In the Ellsberg Paradox, MEU theory is the subject of the debate, and therefore, the formal information cannot include the fact that the standard preferences violate MEU theory. Consequently, it is insufficient to note that the standard Ellsberg preferences violate the MEU Axiom of Independence:

[168] Al-Najjar and Weinstein (2009) p. 252.

Axiom of Independence: If A_x and A_y have the same consequences when X is believed to be false, then an agent will prefer A_x or A_y if and only if these actions have different expected utilities given X.

Put informally, the Axiom of Independence requires that only a difference in expected utility can provide grounds for a difference in preference, so that the ordering of an MEU agent's preferences is *independent* of information that does not alter either the probability or expected utility of the items on the ranking. Such an axiom is a necessary condition of representing someone's preferences using an expected utility function[169].

Unfortunately for the MEU approach, there are no known formal grounds (other than MEU) to regret having the standard Ellsberg preferences. There is no proof that is analogous to the second Blackjack example above. The conservative response might be a correct empirical explanation, but it does not solve the Ellsberg Paradox as a normative problem. Additionally, even if telling someone with the standard Ellsberg preferences about the heuristic explanation of Al-Najjar and Weinstein will induce regret in that person, the explanation is psychological information, rather than formal information. A conservative decision theorist might be willing to tolerate the Ellsberg Paradox and retain MEU theory, but Al-Najjar and Weinstein's response does not provide any reasons to do so.

Neither the ambiguity aversion literature, nor the conservative responses are entirely satisfactory. Contrary to MEU theory, the standard choices seem to be rationally *permissible*, even if they do not seem to be *mandatory*. However, incorporating an absolute preference

[169] Steele and Stefánsson (2015).

against ambiguity into decision theory creates a new set of paradoxes. In the next two sections, I shall discuss two Evidential Probabilist answers to the Ellsberg Paradox.

SECTION 4: KYBURG AND THE ELLSBERG PARADOX

Modifying MEU theory is natural for Evidential Probabilists, because MEU probabilities are always precise, whereas the probabilities in Evidential Probability can be imprecise. However, Evidential Probability as a theory of epistemic probability does not entail any particular normative decision theory. In this section, I shall outline Kyburg's own decision theory. I shall describe his answer to the Ellsberg Paradox. His response preserves the intuition that the standard preferences are permissible but not mandatory, while avoiding the Sunk Cost fallacy, but it has its own challenges. In the next section, I shall develop a more conservative response.

4.1 Kyburg's Decision Theory

In Kyburg's decision theory, the expected utility of an action A given an agent's total evidence K is determined by multiplying the utility of performing A under circumstance C_i for each of the *n* possible circumstances *i* by the probability of C_i given A. Formally, this involves calculating:

$$\sum_{i=1}^{n} U(C_i, A)EP(C_i \mid E \wedge K)$$

However, since the probabilities in Evidential Probability are interval-valued, it is possible that the expected utilities will also be interval-valued[170].

For example, imagine I know that:

(1) Between 0% and 50% of the balls in a bag are red and the evidential probability that a randomly selected ball is red given K is EP(R | K) = [0, 0.5], where R is 'The ball is red'.

(2) There is a payout of £10 for correctly betting that the ball is red.

(3) Between 50% and 100% of the balls in the bag are oblate spheroids and the evidential probability that a randomly selected ball is an oblate spheroid given K is EP(O | K) = [0.5, 1] where O is 'The ball is an oblate spheroid'.

(4) There is a payout of £30 for betting that the ball is an oblate spheroid.

Assume that utilities correspond to monetary values. The expected utilities are [0, 5] for betting that the ball is red and [15, 30] for betting that it is an oblate spheroid.

Kyburg endorses an intuitive principle for circumstances when the *minimum* expected utility of an action exceeds the *maximum* expected utility of the other action:

[170] Kyburg (1990) p. 231.

Principle of Dominance: If there is a choice between A_1 and A_2, where the *minimum* expected utility of A_1 is greater than the *maximum* expected utility of A_2, then a rational agent will choose A_1[171].

In the example of the balls in the bag, this principle requires that I choose to bet that the ball will be an oblate spheroid, rather than betting that it is red.

Kyburg discusses but does not endorse any further decision rules for when the Principle of Dominance does not fully determine a choice[172]. However, he suggests a number of possible decision rules for imprecise expected utilities:

Maximin: Choose the action with the highest minimum expected utility. For example, if the expected utility of betting that the ball is red is [0, 1] and the expected utility of betting that the ball is green is [0.5, 0.75], then a maximin rule would require betting that the ball is <u>green</u>.

Maximax: Choose the action with the highest maximum expected utility. For example, in the above situation where the expected utility of betting that the brick is spherical is [0, 1] and the expected utility of betting that the brick is green is [0.5, 0.75], then a maximax rule would require betting that the brick is <u>spherical</u>.

[171] Kyburg uses ‘dominance’ to refer to an action having a minimum expected utility that exceeds another action’s maximum expected utility. It is important to distinguish his usage from the use of this term in other fields, like game theory.

[172] Kyburg (1990) p. 233.

Combining the Limits: Choose the action that has the highest value for some function that combines the upper and lower limits of the expected utility interval. One simple function is the mean of the limits. If the expected utility of betting that the brick is spherical is [0, 1] with a mean of 0.5 and the expected utility of betting that the brick is green is [0.5, 0.75] with a mean of 0.625, then this rule would opt for betting that the brick is green.

4.2 The Ellsberg Paradox

Kyburg claims that the Ellsberg Paradox is an example of the advantages of an imprecise probability theory, because it illustrates how precise probabilities do not always provide the right inputs for decision-making[173]. His answer begins with the assumption that people with the standard choices are acting in accordance with a Maximin rule: choose the option with the greatest minimum expected utility.

If people's probabilities are interval-valued and based on the statements of the relative frequencies in the set-up, then a Maximin strategy requires choosing the standard choices. As before, suppose that I know that 1/3 of the balls in the urn are red, that between 0 and 2/3 are black, and that I am betting on a randomly selected ball. Returning to Matrix I, my Evidential Probability for R is [1/3, 1/3], so my minimum expected utility is 1/3(£10) = £3.33. My Evidential Probability for B is [0, 2/3], so the minimum expected utility is 0(£10) = 0. If I am following a Maximin rule, then I shall choose **Red** over **Black**, because the minimum expected utility of **Red** is greater. Consequently, if I am a Maximin Evidential Probabilist player, then my preferences will correspond to most people's behaviour for the first choice.

[173] Kyburg (1990) p. 238.

Matrix I

	R	B	Y
Red	£10	0	0
Black	0	£10	0
Black or Yellow	0	£10	£10
Red or Yellow	£10	0	£10

The second choice in the Ellsberg Paradox is between **Black or Yellow** and **Red or Yellow**. My Evidential Probability for (B v Y) is [2/3, 2/3]. Hence, my minimum expected utility is 2/3(£10) = £6.67. My Evidential Probability for (R v Y) is [1/3, 3/3] and the minimum expected utility is 1/3(£10) = £3.33. If I am following a Maximin rule, then I shall choose **Black or Yellow** over **Red or Yellow**, because the minimum expected utility is greater. Thus, Kyburg's response replicates the standard responses in the Ellsberg Paradox, if the player has adopted a Maximin strategy.

4.3 Evaluation

Kyburg's answer to the Ellsberg Paradox has some apparent advantages. Some are real, whereas others are questionable. One real advantage is that a player using Kyburg's version of the Maximin scenario will avoid the Sunk Cost fallacy in the dynamic choice context that I discussed earlier. Suppose that the player has been offered the choice to hedge G against the possibility of the ball being yellow. The ball is drawn behind a screen and the player is told that the ball is not yellow. They are now offered a choice between Red and Black in Matrix III. Since the player knows that the ball is not yellow, the evidential probability of Y is [0, 0]. Consequently, it is irrelevant to the probability of a disjunction that includes it. The probability of R is [1/3, 1/3], whereas the probability of B is [0, 2/3]. Hence:

$$EP(R\ v\ Y) = [\frac{1}{3}, \frac{1}{3}]$$

$$EP(B\ v\ Y) = [0, \frac{2}{3}]$$

Matrix III

	R	B	Y
Red	£10 – G	0	£10 – G
Black	0	£10 – G	£10 – G

The minimum expected utility for Red is 1/3(10 – G) and the minimum expected utility for Black is 0(10 – G) = 0, so that a player following an Evidential Probabilist Minimax strategy will prefer Red, just as they did in Matrix II. The earlier choice to invest G makes no difference to their choice in this case. Therefore, the Minimax Evidential Probabilist player avoids the Sunk Cost fallacy, while also avoiding the Ellsberg Paradox. There is more to say here, but it would be very similar to my discussion of this scenario in

Subsection 6.3. At the very least, Kyburg's answer does not face a *prima facie* problem from Al-Najjar and Weinstein's scenario.

Kyburg's answer might seem to have another advantage, which is that (unlike standard decision theory) it seems to capture that most people's choices in the Ellsberg Paradox seem more "risk-averse" or "cautious", in some informal sense of these terms. One might think that Maximin is a cautious decision rule, because a Maximin agent tries to optimize the worst-case scenario. Kyburg's decision theory might seem to capture an aspect of cautiousness and a sense of 'risk-aversion' that cannot be formalised in MEU.

However, this claim for Kyburg's answer is debatable. Williamson argues (in a different context from the Ellsberg Paradox) that a different method for determining expected utilities using evidential probabilities produces a more cautious long-run sequence of decisions[174]. If Williamson is correct, then it is arguable[175] that the Minimax rule with evidential probabilities does not characterise an ideally rational cautious agent, and one could not claim that Kyburg's answer captures the intuition that players with the standard preferences are following an (ideally rational) cautious rule.

Significantly, Williamson's method produces a different result in the Ellsberg Paradox from Minimax. He is an Objective Bayesian (as described in Chapter 1 Section 2.1) and so he interprets the probabilities in the calculation of expected utilities as "rational degrees of belief". In Williamson's form of Objective Bayesianism, these probabilities are

[174] Williamson (2007) p. 170.

[175] It follows if one adds the additional premise that the asymptotic properties of a decision-rule are suitable for characterising it in a single-case decision like the Ellsberg Scenario. Exploring this premise further would lead me far away from the central thrust of this chapter.

determined in two steps: (1) calibration, in which a rational agent determines the constraints that their available statistical information places on their degrees of belief and (2) equivocation, in which the agent selects a probability function that minimizes the distance between P(Φ) and 0.5 for every Φ in the domain of the function given the constraints from calibration. The distribution for this function becomes the rational agent's degrees of belief.

In the Ellsberg Paradox, the relevant statistical information regarding the statement R 'The ball will be red' is that 1/3 of balls in the urn are red, so that calibration requires that P(R) = 1/3. For B, calibration requires that P(B) must be in the interval [1/3, 2/3], since this is the player's best statistical information regarding B. For the statement Y, calibration also requires that P(Y) must be in the interval [1/3, 2/3]. The player knows that R, B, and Y are a set of mutually exclusive and exhaustive statements. A value of 1/3 minimizes the distance between P(B) and 0.5, as well as P(Y) and 0.5. Therefore, a rational agent who uses Williamson's Objective Bayesianism to determine probabilities for calculating expected utilities will be indifferent in the Ellsberg Scenario: **Red and Black** have equal expected utilities, as do **Black or Yellow** and **Red or Yellow**[176].

In the absence of a convincing reason to consider Maximin or Objective Bayesianism as more cautious (and thus indifference or the standard preferences as more "cautious" in the Ellsberg Paradox) it seems that cautiousness is still a concept without a clear explication. Perhaps no general formalisation of this concept is possible. It is unproven that Kyburg's response captures the intuition that the standard preferences are cautious, because the concept of cautiousness is still not clearly explicated and plausible alternative explications can

[176] For precise expected utilities, the expected utilities of **Red** and **Black** are equal if and only if the expected utilities of **Black or Yellow** and **Red or Yellow** are equal, since the expected utility of **Red** is equal to 1 minus the expected utility of **Black or Yellow** and the same *mutatis mutandis* for **Black** and **Red or Yellow**.

produce different results.

Additionally, Kyburg's answer to the Ellsberg Paradox faces some positive objections. Firstly, it is relative only to a Maximin strategy, so that if one thinks that it is *always* rational to have the standard preferences in the Ellsberg Paradox, then Kyburg's account will not justify this belief. For example, if we combine the limits to compute an expected utility, then the expected utilities are as follows:

(1) Expected utility of betting that the ball is red: $\frac{1}{2}\left(\frac{1}{3}+\frac{1}{3}\right)£15 = £5.$

(2) Expected utility of betting that the ball is black: $\frac{1}{2}\left(0+\frac{2}{3}\right)£15 = £5.$

(3) Expected utility of betting that the ball is not red: $\frac{1}{2}\left(\frac{2}{3}+\frac{2}{3}\right)£15 = £10.$

(4) Expected utility of betting that the ball is not black: $\frac{1}{2}\left(\frac{1}{3}+1\right)£15 = £10.$

- so that one should be indifferent in both choices. On the other hand, the Ellsberg Paradox is usually understood to be the problem that the standard choices are *permissible*, not that they are *mandatory*. It does not seem to be a problem that Kyburg's decision theory only enables us to establish the possibility of a rational agent making the standard choices.

A more serious objection is that Kyburg's use of decision rules to characterise the choice of actions (when such choices are not determined by the Principle of Dominance) is unnecessary given the rest of his formal epistemology. Kyburg also has a notion of “practical probabilities”, which are precise probabilities that are introduced for when precise probabilities are needed for calculating mathematical expectations[177]. For example, the Evidential Probability that a normal coin will land tails in a long series of trials is not *exactly* [0.5. 0.5], but this degenerate interval is a subinterval of the actual evidential probability, and so it is suitable for practical purposes. The only endogenous constraints on the choice of practical probabilities are that they must be within the relevant Evidential Probability intervals and they must be consistent with the axioms of the precise probability calculus. Since neither the choice of practical probability nor the choice of decision rule (like Minimax or Maximax) is determined within Kyburg’s model, there seems nothing to be gained by introducing the latter for decision-making. Decision rules are no more objective than practical probabilities, especially since Kyburg does not require that the choice of rule stays constant over time.

This point is pertinent, because there are reasons why Minimax should not be a decision rule for all choices. In particular, when the potential losses are very small and the potential losses are very great, Minimax can become a very implausible rule[178]. To use an extreme example, imagine you are offered a choice between (a) investing a £1 coin in a lottery with a 99% chance of a £1 billion prize and (b) a guaranteed sum of eleven 10p coins. Since the worst-case scenario if you choose (a) is a loss of £1, whereas the worst (and best) case scenario for (b) is a gain of £1.10, if you follow the Minimax decision rule, then you must choose (b). Yet such a choice is highly counterintuitive. Even if someone chooses to

[177] Kyburg (1990) p. 66.

[178] Resnik (1987) p. 27.

follow a Minimax rule in the Ellsberg Scenario, it is implausible that ordinary people could follow such a rule in all possible decision problems[179].

The issue of different decisions over time raises the further problem of dynamics. Kyburg does not provide an account of how different rules should be combined over a series of decisions. If I am playing a strategy game and I face a large number of optimization decisions over the course of the game, then it is unclear whether Kyburg's decision theory allows me to switch mid-game between Minimax, Maximax, Combining the Limits, a Williamson-style approach, or other rules. My point here is not that Kyburg's theory of dynamic decision-making is positively flawed, but that it is undeveloped. It would be preferable to have a decision theory that provided a similarly intuitive answer to the Ellsberg Paradox, while also having an account of decision-making over time. Kyburg's suggestion of practical probabilities that are within the Evidential Probability intervals could provide such a theory, because he has a well-developed theory of updating these intervals.

Finally, the redundancy of rules introduces inelegance into Kyburg's decision-theory. He requires that a decision-maker must check the coherence of their choice of betting-odds, but this is a distinct step from the choice of decision-rule[180]. Yet, if an agent acts in accordance with a practical probability distribution, then their choice of odds will be coherent (see the end of Subsection 5.2 in Chapter 1) and they will not require any additional choice of rule. It would be more elegant to remove the redundant element and simply use practical

[179] There are other versions of the Maximin rule, such as the Maximin Regret rule, that avoid this particular problem, but a discussion of such rules would be extensive and would not address Kyburg's actual position, which involves Maximin.

[180] Kyburg (1978) p. 160-161.

probabilities.

However, Kyburg's approach might be attractive as a decision theory that is less computationally demanding than MEU or my proposal in the next section. Imagine that I have evidential probabilities for a set of different possible outcomes given a statistical database, but calculating a coherent additive probability distribution over the range of possibilities is impractical. Assume that a particular computationally undemanding rule like Minimax or Maximax is acceptable in this context: for example, if I am working in a high-risk investment fund, Maximax might be appropriate, whereas Minimax might be a better rule when the 'utilities' I am maximizing are the number of people who survive a possible landslide. As a normative theory of ideal decision-making *given computational constraints*, an expanded version of Kyburg's theory might be an attractive option[181].

I shall not explore this suggestion further in this thesis, and continue on the assumption that computational constraints are unimportant. I have no fundamental objections to Kyburg's theory, but I shall argue that an Evidential Probabilist can do better in the realm of ideal normative decision theory. Practical probabilities and a principle like MEU seem capable of doing everything that Kyburg's approach can do for modelling ideal rational agents, but with greater clarity and elegance. I shall develop this proposal in the next section.

[181] It might also have useful applications when the utilities in a context do not have a proper structure to be represented by anything richer than an interval-valued function.

SECTION 5: EVIDENTIAL PROBABILITY, PRACTICAL PROBABILITY AND DECISIONS

In this section, I shall propose a new Evidential Probabilist decision theory. Firstly, I shall offer a method to determine precise expected utilities with imprecise evidential probabilities. This provides a simulacrum of orthodox decision theory. Secondly, I shall propose a quantitative measure of how far we have to speculate beyond the evidence when determining these expected utilities. Thirdly, I shall use this measure to make an addition to orthodox decision theory, which will be vital for my answer to the Ellsberg Paradox in Section 6.

5.1 Practical Probabilities

Practical probabilities are Kyburg's device for determining precise probabilities given imprecise evidential probabilities. He does not provide an interpretation for what they mean, but I shall interpret them as the postulation of the relative frequency information that would be required for a precise probability distribution given the background knowledge.

I shall illustrate this idea with an example. Imagine that an Evidential Probabilitist is betting on a toss of a £1 coin. She is calculating her expected utilities for choosing heads or tails. Her best statistical information is that £1 coins land heads with a relative frequency that is very close to 50%. The same holds for £1 coins landing tails. H is the hypothesis that the £1 coin will land heads and T is the hypothesis that it will land tails. The evidential probabilities could be:

$$EP(H \mid K) = [0.49, 0.51]$$

$$EP(T \mid K) = [0.49, 0.51]$$

Neither of the above evidential probabilities is sufficient to calculate an expected utility for her bets, even if precise utilities are available. However, evidential probabilities do not have to represent someone's actual knowledge: they can be a hypothetical body of knowledge[182]. Thus, she could calculate that the evidential probability that she *would* have if she had the statistical information to calculate a precise probability. This probability would be speculative, in the sense that she has used more information about relative frequencies than she has in her evidence. We can represent this additional speculative information as S and conjoin it with K to generate a practical probability statement. For example, let S_1 be the conjecture that the long-run relative frequency of £1 coins landing heads is exactly 1/2 and S_2 be the conjecture the long-run relative frequency of landing tails is exactly 1/2. If these are suitable reference class statements for H and T, then the following equations hold:

$$EP(H \mid S_1 \wedge K) = [0.5, 0.5]$$

$$EP(T \mid S_2 \wedge K) = [0.5, 0.5]$$

These degenerate intervals are the agent's practical probabilities, which can be used as inputs for calculating expected utilities. I shall require that the speculated relative frequency must be within the relevant Evidential Probability intervals for the hypothesis in question, because these intervals represent what guidance the Evidential Probabilist has from her evidence. Additionally, the relative frequency statement must describe a reference class that will provide a precise probability in these circumstances: it must be in a class that contains this particular £1 coin as a random member.

[182] Kyburg (1974) p. 317.

Formally, suppose that x and y, which are the limits of the evidential probability, have values such that $x \neq y$, so that the interval is not degenerate. Then, when:

$$EP(A \mid K) = [x, y]$$

- an Evidential Probabilist can calculate a practical probability by supposing an S such that -

$$EP(A \mid S \wedge K) = [z, z]$$

- where $x \leq z \leq y$. By doing so, the Evidential Probabilist can always obtain a practical probability for decision-making.

In this thesis, I shall take the choice of speculated relative frequency to be exogenous, except for (1) the constraints above and (2) the important constraint that any resulting practical probability distribution is consistent with the axioms of additive probability. In the coin-tossing case, it is permissible to choose any relative frequency within the [0.49, 0.51], provided that this does not create an incoherent practical probability distribution.

Additionally, S is *not* incorporated into the total evidence. Thus, if an Evidential Probabilist learns some new statement E, then her new evidential probability for a hypothesis H is EP(H | E ^ K), not EP(H | E ^ S ^ K). For example, in the coin-tossing example, imagine that the Evidential Probabilist learns that this particular £1 coin is weighted towards tails, so that it lands heads with a relative frequency of 40-48%. Consequently, the updated evidential probabilities are:

$$EP(H \mid E \wedge K) = [0.40, 0.48]$$
$$EP(T \mid E \wedge K) = [0.52, 0.6]$$

The calculation of a new set of expected utilities will require a new speculative statement, S´, because the earlier speculation S results in a probability that conflicts with the new evidential probabilities. She might choose S´ the midpoints of the intervals, such that:

$$EP(H \mid E \wedge S' \wedge K) = [0.44, 0.44]$$
$$EP(T \mid E \wedge S' \wedge K) = [0.56, 0.56]$$

Alternatively, she might try to minimize the distance between the practical probabilities and 0.5, and choose a statement S´´:

$$EP(H \mid E \wedge S'' \wedge K) = [0.48, 0.48]$$
$$EP(T \mid E \wedge S'' \wedge K) = [0.52, 0.52]$$

My proposal is similar to Kyburg's approach, in that the evidential probabilities put constraints on the expected utilities and there is an exogenous choice. In Kyburg's decision theory, the exogenous choice is the selection of a decision rule. In my proposal, the choice is the selection of a speculative statement. In both decision theories, the updating involves Kyburg's rules for reference class selection rather than Bayesian updating.

My proposal is also similar to some forms of Objective Bayesianism, such as the

"Evidential Probability-Calibrated Objective Bayesianism" of Williamson and Wheeler[183]. As in my proposed decision theory, they use intervals from Evidential Probability to provide constraints on the choice of probabilities for decision-making. The principal differences are that (1) their theory involves additional constraints on the choice of probability, (2) the selection in my procedure is formalised in Evidential Probability alone, whereas Williamson and Wheeler use an Objective Bayesian formalism, and (3) they also discuss degrees of belief and my decision theory is silent on degrees of belief. These differences should not be overestimated. Firstly, nothing I say is inconsistent with the additional constraints that Williamson and Wheeler propose. I merely do not endorse these requirements as part of the general theory of decision-making. Secondly, my choice of formalism is only a method to emphasise the distinction between (a) our evidence and background knowledge ($E \wedge K$) and (b) the speculative statement S which we are only pretending to know; this distinction is important for my answer to the Ellsberg Paradox. Thirdly and finally, since my decision theory is silent on degrees of belief, it does not explicitly contradict anything that Williamson and Wheeler have to say about such quantities.

I shall immediately address three possible objections. The first is whether my proposal creates the possibility of synchronic Dutch Books. A Dutch Book is a set of minimum acceptable odds such that, if a player commits to betting positive sums on them, it is possible for the bookie to have a strategy that guarantees that the player loses money. Synchronic Dutch Book Arguments are intended to prove that one should adopt odds that correspond to an additive probability distribution in order to avoid the possibility of a Dutch Book, such that:

[183] Wheeler and Williamson (2011) p. 327-329.

$$k(H) = \frac{P(H)}{1 - P(H)}$$

- where k(H) is the minimum betting odds-ratio that the player will accept on a statement H. Such a probability distribution is allegedly a necessary and sufficient condition for avoiding Dutch Books.

Rather than review the various synchronic Dutch Book Arguments and the many criticisms that have been made of them, I shall simply note that my proposal *does* involve betting on a practical probability distribution that satisfies the axioms of additive probability. (There are reasons to weaken this requirement if the Dutch Book scenario is insufficiently constrained[184].) As I described in Subsection 5.2 of Chapter 1, it can be proven that there is always a coherent additive probability function whose values are within the intervals of evidential probabilities. It is uncontroversial that having such a function is a sufficient condition for avoiding Dutch Books. Provided that a player ensure that her choice of practical probabilities are not in conflict within her Evidential Probability intervals, she will not be vulnerable to Dutch Books, *even if* the synchronic Dutch Book Arguments are sound.

The second possible objection that I shall address comes from *diachronic* Dutch Book Arguments, which are intended to establish conditionalization as the sole updating rule. Allegedly, such arguments establish that someone who does not update via conditionalization alone will be diachronically incoherent: they can become committed to a series of bets such that they will necessarily lose money overall, but updating via conditionalization alone avoids this possibility. Since evidential probabilities are not always updated via conditionalization

[184] Rowbottom (2007).

(as discussed in Chapter 1 Subsection 5.3) a Bayesian might object that my proposal will lead to diachronic incoherence.

However, as Howson and Urbach note, there are no arguments that prove conditionalization is a necessary condition of diachronic coherence[185]. For example, one of the most prominent Diachronic Dutch Book Arguments comes from Paul Teller, but it is provable that deductive consistency and a desire not to lose money is sufficient to avoid diachronic incoherence in Teller's scenario[186]. In general, there is no established argument in the literature that proves that a rational decision-maker must always update by conditionalization.

Furthermore, even if there were such an argument, it would be problematic for *all* decision theories that require coherent precise probabilities, including Bayesian decision theories. If a person's initial probability distributions are incoherent, then it can be proven that no amount of conditioning will enable them to attain coherence[187]. Let P be an incoherent probability function such that P(H | E) is inconsistent with the axioms of additive probability; let P´ be a coherent probability function; let E be the total evidence that a decision-maker using P will ever have. For P(H) to be altered to P´(H) by conditionalizing on E, it must be the case that:

$$P'(H) = P(H \mid E) = \frac{P(H \wedge E)}{P(E)}$$

[185] Howson and Urbach (1993) p. 103.

[186] Bacchus, Kyburg, and Thalos (1990) p. 486-487.

[187] Bacchus, Kyburg, and Thalos (1990) p. 494.

- but if the equation above is true, then P(H | E) is coherent, which is contrary to the initial assumption that P(H | E) is incoherent. It follows that an incoherent probability distribution can never become coherent via conditionalization.

Therefore, *if* there was a sound Diachronic Dutch Book Argument and if incoherence is irrational, then there would be a problem for both my proposal *and* normative decision theories that have conditionalization as their sole updating norm. Anyone who had incoherent initial probabilities and who unwisely chose to update using only conditionalization would be *eternally* irrational. Since, in the real world, nobody is born as a perfectly coherent Bayesian, this would be a non-trivial problem for Bayesian decision theory.

A supporter of conditionalization could object that normative decision theory concerns ideally rational decision-making and an ideally rational agent would not have incoherent initial probabilities. However, while it is theoretically interesting to imagine ideally rational agents, the primary appeal of normative decision theory is the advice that they can provide for real people, and a theory in which one can never move from incoherence to coherence (even for simple problems) is a counsel of irrationality. If I am betting on a horse race and I realise that my initial choice of odds was inconsistent with the probability calculus, and I have the option to change my choice, then intuitively I should change my odds, regardless of the fact that this requires that I do not change them via conditionalization. Fortunately for both Bayesians and Evidential Probabilists, there is no established Diachronic Dutch Book Argument. Updating using Evidential Probability does not commit us to diachronic incoherence.

Those who think that epistemic probabilities *must* be imprecise might also object to my proposal. They could argue that I am assuming away the problem of ambiguous probabilities and ignoring the difference between situations where imprecise probabilities are appropriate and where precise probabilities are appropriate. However, practical probabilities are not epistemic probabilities, in the sense that they do not describe *evidential* relations. Instead, they are simply inputs for decision-making. Furthermore, unlike standard decision theorists, I *do* make a distinction between (a) precision that comes from evidence and (b) precision that comes from speculation. In the next subsections, I shall use this distinction to provide the basis for my answer to the Ellsberg Paradox.

There are many other reasons for adopting a purely imprecise probability framework for decision-making. Since the intended target of my argument is primarily those seeking a purely *precise* decision theory akin to MEU, I shall not discuss the many interesting reasons to adopt other heterodox alternatives.

5.2 A Degree of Uncertainty Measure

For most contexts, my proposal produces a decision theory that is similar to MEU theory: an ideally rational agent has precise probabilities and precise utilities; she multiplies these to calculate her expected utilities; she acts so that she maximizes expected utility. The similarity will be especially strong when the decision problem is static or the MEU agent is conditioning on new statistical knowledge.

Before I introduce the important difference between my proposal and orthodox decision theory, I shall introduce an additional function. This function measures the 'Degree of Uncertainty' (DU) of an agent's epistemic state with respect to a hypothesis. By "uncertainty", I mean the extent to which the evidence does not mandate a particular practical probability distribution. Informally, DU function measures the distance from (a) the evidential probability of a statement given the total evidence, to (b) a practical probability obtained by speculating relative frequency statements. Put another way, the function measures the extent to which the practical probabilities involve speculation beyond the evidence. Formally, for a hypothesis H and total evidence K, the function DU is the following:

$$Degree\ of\ Uncertainty = DU(H \mid K) = (y - x)$$

- where *x* and *y* are the upper and lower bounds of the evidential probability of H given K.

This function is formally identical to the degree of imprecision measures that Walley and Kyburg proposed for quantifying the weight of argument. In Chapter 2, I argued that such measures were unsatisfactory for the concept of weight. However, I shall instead use them to measure the 'speculativeness' of the practical probabilities. Another way of characterising the DU measure is that it measures the extent to which the possible practical probabilities can be 'evidence-based'.

This function is notably simple, because it contains just two variables and a single operation. Furthermore, some other very simple functions would have significant problems. For instance, suppose that one used the ratio of the limits rather than the difference, such that the function is DU´(H | K) = ($y \div x$). Such a measure would mean that [0.01, 0.02] would provide a different uncertainty value from [0.03, 0.04], since ($0.01 \div 0.02$) = 0.5 and ($0.03 \div 0.04$) = 0.75. However, it is intuitive that the entering into a lottery in which I have 1-2% of the tickets involves the same degree of uncertainty as entering into a lottery in which I have 3-4% of the tickets. In contrast, the DU measure gives the same value, 0.01, for the uncertainty involved in either lottery. There might be more complex functions that are also satisfactory measures, but it is reasonable to use such a simple function in the absence of a reason to consider the DU function inadequate.

I shall provide some examples of the DU measure in action. Firstly, the evidential probability might be a degenerate interval, so that a practical probability is entirely determined by applying Kyburg's system to the hypothesis and total evidence in question. For instance, suppose that I know that a card will be randomly drawn from a normal playing deck and this is my best statistical information about which card will be drawn:

Key

H: The card drawn will be the Queen of Diamonds.

K: My total relevant evidence.

$$EP(Q \mid K) = [\frac{1}{52}, \frac{1}{52}]$$

$$DU(Q \mid K) = \frac{1}{52} - \frac{1}{52} = 0$$

My evidence uniquely determines a practical probability and so the DU value is zero. Generally, whenever the Evidential Probability is a degenerate interval, then the DU function gives its minimum value.

Secondly, suppose that the evidential probability is very close to a precise value. For example, when tossing a £1 coin, I know that the relative frequency of such coins landing heads is close to 50%. Suppose that my best statistical information is that £1 coins land heads in tosses with a long-run relative frequency of 49-51%.

Key

H: The coin will land heads.

K: My total relevant evidence.

$EP(H \mid K) = [0.49, 0.51]$

$DU(H \mid K) = 0.51 - 0.49 = 0.02$

The DU value in this case is higher than when choosing the card, but it is still small.

In contrast, consider tossing a gömböc. A gömböc is a homogenous three-dimensional solid that has just two equilibria on a flat surface, one stable and the other unstable. The standard construction is similar to a sphere, but with a sharp top. Given my (very limited) knowledge of gömböcs and physics, I am very ignorant regarding the relative frequency of tossed gömböcs settling on either of these points. Suppose that someone tells me that gömböcs settle on their stable point at least 10% of the time. For all that I know, it is possible that gömböcs *always* settle on their stable point. Hence:

Key

H: The coin will settle on its stable equilibrium point.

K: My total relevant evidence.

$$EP(H \mid K) = [0.1, 1]$$

$$DU(H \mid K) = 1 - 0.1 = 0.9$$

This DU value reflects the high level of uncertainty regarding my choice of a precise relative frequency, because it would have been consistent with my evidence to choose any precise relative frequency from 0.1 to 1. In contrast, in the case of tossing the £1 coin, my choice of practical probability was limited to the interval from 0.49 to 0.51.

Finally, the DU measure will have a maximum value if Evidential Probability does not provide *any* non-vacuous probability. For instance, suppose that I must bet on the hypothesis that 'Most things are gavagai'. Since I do not know what 'gavagai' means, I lack any statistical basis for the probability value. An Evidential Probabilist can represent this state of great uncertainty as:

Key

G: Most things are gavagai.

K: My total relevant evidence.

$$EP(G \mid K) = [0, 1]$$

$$DU(G \mid K) = 1 - 0 = 1$$

When the value of DU is maximal, the selection of *any* S for *any* practical probability would be consistent with the evidential probability of G given H. The only additional constraint is that any practical probability distribution created by this procedure must not violate the axioms of additive probability. Such a purely speculative choice of probability distribution is akin to the choice of a Subjective Bayesian prior.

5.3 The Principle of Lesser Uncertainty

Intuitively, the DU measure seems to be tracking something that is important for decision-making. *Ceteris paribus,* it would be preferable to make decisions where the degree of uncertainty is zero. For example, if we are gambling, we would ideally like our choice of odds to be grounded on evidence about relative frequencies, rather than probabilities based on speculations. (This desire for guidance by evidence is distinct from the desire for an extreme probability.) Intuitively, DU is also relevant because a rational agent should (when possible) avoid making decisions that are based on speculation. This normative constraint also seems to admit of degrees: a practical probability that was determined from an evidential probability of [0.74, 0.76] is preferable to one determined via [0.5, 1]. Even if the choice of practical probability (perhaps 0.75) was identical in both cases, the former provides the basis for a more evidence-based decision.

However, these preferences are comparative, rather than absolute. I am *not* claiming that one should want to stake more decisions when the practical probabilities are relatively evidence-based. Instead, I am merely claiming that a rational agent should use DU as a potential tiebreaker when expected utility requires indifference. Importantly, this means that I am not claiming that DU should *always* be part of our decision-making, as practical probabilities and utilities should be.

I shall now present a principle that formalises these intuitions. ‘Utilities’ will refer to precise utilities that satisfy all the normal axioms in MEU theory, *except* for a decision problem in which the expected utilities are identical.

Principle of Lesser Uncertainty: When the expected utilities for a set of mutually exclusive and exhaustive circumstances are equal, then choose from among the set of actions that yield the highest expected utility given the circumstance with the lowest DU value relative to your total evidence.

Thus, DU provides an additional dimension of assessment for choosing between actions with equal expected utilities. Consequently, an agent’s preferences will be modelled using two parameters: a cardinal utility *and* the DU measure.

Before returning to the Ellsberg Paradox, I shall illustrate the Principle of Lesser Uncertainty with some other examples. Firstly, imagine that you have a choice between betting on a toss of one of two coins. Coin A is a perfectly normal 50p coin. Coin B is a very unevenly shaped Ancient Greek coin from the Island of Patmos, with a face of Saint Paul on one side. Someone offers you a choice between betting heads and betting tails. You know that 50p coins land heads with a relative frequency of about 50%, whereas you are ignorant with respect to the relative frequency of Coin B landing heads. Assume that the utilities for winning with each bet are equal.

Key

H_a: Coin A will land heads.

H_b: Coin B will land heads.

K: Your total relevant evidence.

$$EP(H_a \mid K) = [0.49, 0.51]$$

$$EP(H_b \mid K) = [0, 1]$$

Suppose that you determine the expected utilities by postulating S and S´:

$$EP(H_a \mid S \wedge S' \wedge K) = [0.5, 0.5]$$

$$EP(H_b \mid S \wedge S' \wedge K) = [0.5, 0.5]$$

Given that H_a and H_b have equal payouts and equal practical probabilities, they have equal expected utilities. Using the DU measure:

$$DU(H_a \mid K) = 0.51 - 0.49 = 0.2$$

$$DU(H_b \mid K) = 1 - 0 = 1$$

Since the DU for H_a is less than the DU for H_b, it follows from the Principle of Lesser Uncertainty that you ought to have a comparative preference for betting that H_a to betting that H_b.

To summarise, I am suggesting the following constraints for the practical probabilities:

(1) The practical probabilities must be within the Evidential Probability intervals. In the special case of degenerate intervals, this fully determines the probability distribution.

(2) If you have formulated the practical probabilities using speculated relative frequency statements, then the resulting distribution for the domain must be consistent with the axioms of additive probability.

(3) An agent should maximize the product of their practical probabilities and a special type of cardinal utility function. (Special, because the outputs do not fully represent the agent's preferences.) In the special case where outcomes have equal values for this product, but one outcome (or set of outcomes) has a lesser DU value given the total evidence, the agent should prefer actions with the highest expected utility given this outcome. This modification of the standard MEU approach is motivated by the intuition that reducing the uncertainty involved in one's decisions is preferable, *ceteris paribus*.

It is worth stressing that I have said nothing about degrees of belief. Practical probabilities are not degrees of belief, but devices for decision-making; evidential probabilities are not interval-valued degrees of belief, but formal representations of evidential relations. My decision theory is compatible with a large variety of perspectives on the role of beliefs in formal epistemology and the philosophy of science: ranging from philosophers like Popper who argue that belief should not be at the heart of epistemology[188], to philosophers like Haack who consciously reject the Popperian approach and put the subjective experiences of a knowing subject at the centre of their epistemology of science[189]. In short, my decision theory is not a theory of degrees of belief. I shall now apply this theory to the Ellsberg Paradox.

SECTION 6: THE ELLSBERG PARADOX AND THE PRINCIPLE OF LESSER UNCERTAINTY

In this section, I shall examine the Ellsberg Paradox using the DU measure and a decision theory that uses the Principle of Lesser Uncertainty. I shall demonstrate how this is sufficient to rationalise the standard choices that people make. My answer is essentially that *if* people have practical probabilities such that the expected utilities are equal, then the Principle of Lesser Uncertainty can be used as a tiebreaker that mandates the standard Ellsberg Scenario preferences.

[188] Popper, 'Epistemology Without a Knowing Subject' in (1972).

[189] Haack (2003) p. 57-88.

6.1 The Ellsberg Scenario

In this subsection, I begin by presenting a general algebraic proof that my proposed decision theory permits, under particular circumstances, the standard Ellsberg Scenario preferences. I shall then discuss the classic version of the Ellsberg Scenario to illustrate this answer.

Firstly, I shall describe what is required (on my proposed decision theory) for the practical probabilities and a cardinal utility function to underdetermine the choices in an Ellsberg Paradox decision problem:

Key

Red: Choosing to bet that the ball will be red.

Black: Choosing to bet that the ball will be black.

Black or Yellow: Choosing to bet that the ball is not red.

Red or Yellow: Choosing to bet that the ball is not black.

R: The ball will be red.

B: The ball will be black.

Y: The ball will be yellow.

¬R: The ball is not red.

¬B: The ball is not black.

¬Y: The ball is not yellow.

K: The agent's background knowledge, including the information that there are 90 balls, and that 30 of the balls in the urn are red, 0-60 are black, and 0-60 are yellow, and that all the balls in the urn have *one and only one* of these colours.

Assume equal monetary gains[190] for successfully betting on R, B, Y, ¬R, ¬B, and ¬Y. I shall now calculate the evidential probabilities that an agent would need for equal expected utilities for (1) **Red** and **Black** and (2) **Black or Yellow** and **Red or Yellow**. This is very easy in the case of R. From the agent's knowledge E about the relative frequencies, they know that:

$$EP(R \mid K) = [\frac{1}{3}, \frac{1}{3}]$$

- because 30 out of the 60 balls are red. Thus, their knowledge of the relative frequencies provides a degenerate interval. Similarly, for ¬R, the evidential probability is precise:

$$EP(\neg R \mid E) = [\frac{2}{3}, \frac{2}{3}]$$

- because they know that exactly 1/3 of the balls in the urn are red, and therefore that 2/3 of the balls in the urn are not red.

For B, Y, ¬B, and ¬Y, their evidential probabilities are imprecise, because their

[190] Again, monetary values are assumed to correspond unproblematically to utilities.

knowledge of the relative frequencies is imprecise. The evidential probabilities are as follows:

$$EP(B \mid K) = [0, \frac{2}{3}]$$

$$EP(Y \mid K) = [0, \frac{2}{3}]$$

$$EP(\neg B \mid K) = [\frac{1}{3}, 1]$$

$$EP(\neg Y \mid K) = [\frac{1}{3}, 1]$$

Suppose that the agent uses the mid-points of the intervals to decide their choice of their speculative relative frequency statements. This is permissible, though not mandatory. Thus, they have the following practical probabilities:

$$EP(R \mid S_1 \wedge K) = [\frac{1}{3}, \frac{1}{3}]$$

$$EP(B \mid S_2 \wedge K) = [\frac{1}{3}, \frac{1}{3}]$$

$$EP(Y \mid S_3 \wedge K) = [\frac{1}{3}, \frac{1}{3}]$$

$$EP(\neg R \mid S_4 \wedge K) = [\frac{2}{3}, \frac{2}{3}]$$

$$EP(\neg B \mid S_5 \wedge K) = [\frac{2}{3}, \frac{2}{3}]$$

$$EP(\neg Y \mid S_6 \wedge K) = [\frac{2}{3}, \frac{2}{3}]$$

To simplify, suppose that ‘1’ is the cardinal utility for winning with each choice of bet. Thus, the expected utilities for **Red** and **Black** are both 1/3. The expected utilities for **Black or Yellow** and **Red or Yellow** are both 2/3.

The standard choices in the Second Ellsberg Scenario are a choice of **Red** over **Black** and **Black or Yellow** over **Red or Yellow**. Assuming the above practical probabilities, such choices are irrational according to MEU theory: the rule of maximizing expected utility requires that an agent must be indifferent.

Here, the difference between my proposal and MEU theory becomes apparent. The DU values are not identical for the possible outcomes:

$$DU(R \mid K) = \frac{1}{3} - \frac{1}{3} = 0$$

$$DU(B \mid K) = \frac{2}{3} - 0 = \frac{2}{3}$$

$$DU(Y \mid K) = \frac{2}{3} - 0 = \frac{2}{3}$$

$$DU(\neg R \mid K) = \frac{2}{3} - \frac{2}{3} = 0$$

$$DU(\neg B \mid K) = 1 - \frac{1}{3} = \frac{2}{3}$$

$$DU(\neg Y \mid K) = 1 - \frac{1}{3} = \frac{2}{3}$$

Using the Principle of Lesser Uncertainty, an agent has a reason to favour **Red** over **Black**, because **Red** pays out if R obtains and there is no DU for this hypothesis, whereas **Black** pays out if B obtains and there is some DU for this hypothesis. Similarly, they have a reason to favour **Black or Yellow** over **Red or Yellow**, because **Black or Yellow** pays out if ¬R obtains and there is no DU for this hypothesis, whereas **Red or Yellow** pays out if ¬B obtains, and there is some DU for this hypothesis. In short, the standard preferences minimize DU. Therefore, my proposal rationalises the standard preferences: they are permissible, but not mandatory.

I provide a general algebraic proof below:

Key

A1, A2, A3, and A4: An exhaustive and mutually exclusive set of actions for an agent.

H1, H2, and H3: An exhaustive and mutually exclusive set of possible circumstances.

U: The agent's utility function.

K: The agent's total relevant evidence.

(1) $U(A1) > U(A2) \leftrightarrow H1$

(2) $U(A2) > U(A1) \leftrightarrow H2$

(3) $U(A3) > U(A4) \leftrightarrow \neg H1$

(4) $U(A4) > U(A3) \leftrightarrow \neg H2$

(5) $DU(H1 \mid K) < DU(H2 \mid K)$

(6) $DU(\neg H1 \mid K) < DU(\neg H2 \mid K)$

We assume that the choice between A1 and A2 is otherwise undetermined. We also assume that the choice between A3 and A4 is otherwise undetermined. From (1), (2), (5), and the Principle of Lesser Uncertainty:

(7) The agent should choose A1 over A2.

From (3), (4), (6), and the Principle of Lesser Uncertainty:

(8) The agent should choose A3 over A4.

Therefore, if the agent has utilities such that (1-4) are true, their evidential probabilities are such that (5) and (6) are true, and they are otherwise indifferent in each choice, then they should have the standard preferences of (7) and (8).

The standard preferences are not obligatory, because the derivation above depended on the choice of practical probabilities. If the agent had chosen to speculate that 2/3 of the balls in the urn are black, then (*ceteris paribus*) they would be committed to **Black** rather than **Red** and **Red or Yellow** rather than **Black or Yellow**, In general, my rationalisation depends on the assumption of equal expected utilities for each of the two choices.

An additional feature of my answer to the Ellsberg Paradox is that it is an example of how Keynes's concept of the weight of argument (the quantity of relevant evidence) can have practical consequences for rational decision-making. Although there is no established quantitative measure of weight, intuitively there is more evidence (in an informal sense of 'more evidence') in the Ellsberg Paradox for the conjectures that 'The ball is red' and 'The ball is not red' than 'The ball is black' and 'The ball is not black'. This difference in the quantity of evidence produces a difference in the DU measure. As I argued in Chapter 2, this will not always occur: an increase in relevant evidence can occur without narrowing the width of evidential probability intervals. Nonetheless, the Ellsberg Paradox is an example of how differences in weight can have important consequences for rational decision-making.

6.2 Degree of Uncertainty and Sunk Costs

Now that I have demonstrated how the Principle of Lesser Uncertainty justifies the standard Ellsberg Paradox choices, I shall show how it also avoids the Sunk Cost fallacy in the scenario that I described in Subsection 3.1.

In this scenario, the set-up is the same as the Ellsberg Paradox scenario, except that the player has the option to invest a sum G to hedge against the risk of a yellow ball. Someone other than the player draws the ball behind a screen. If the ball that is drawn is yellow, then a player who invested R will receive a net gain of £10 – G. If the ball that is drawn is not yellow, then the player is offered a choice between (a) betting that the ball is red or (b) betting that the ball is black. A player who guesses incorrectly receives no net gain. If the player guesses correctly, then a player who has invested G receives £10 – G, whereas a player who has not invested G receives £10.

As discussed earlier in Subsection 3.1, an MEU player avoids the Sunk Cost fallacy in this scenario. By the same reasoning, a player in the Ellsberg Scenario who follows my proposal will not take sunk costs into account in cases where an MEU player would be not be indifferent, because the Principle of Lesser Uncertainty only applies when an MEU player would be indifferent.

Additionally, when the Principle of Lesser Uncertainty applies, a player using my decision theory will avoid taking G into account and thus avoid the Sunk Cost fallacy. Thus, their choices will depend purely on their expected utilities and the degree of uncertainty of the different options. I shall begin by considering a player who has not chosen to invest G. I shall represent the information that the ball is not yellow by ¬Y. The player faces the following decision-matrix upon learning ¬Y:

Matrix II

	R	B	Y
Red	£10	0	0
Black	0	£10	0

By assumption, the practical probabilities and the cardinal utility function underdetermine choice. Since the potential gains are the same and we assume that the monetary sums represent utilities, a player with equal expected utilities must have speculated that 1/3 of the balls are black. Therefore, they have speculated that 1/3 of the balls are black and 1/3 of the balls are yellow. (They know without speculation that 1/3 of the balls are red.)

Suppose that they have learned that ¬Y. Assume that R and B still have equal probabilities. Their new practical probabilities and degrees of uncertainty, given their knowledge and speculations, must be:

$$EP(R \mid \neg Y \wedge S \wedge K) = [\frac{1}{2}, \frac{1}{2}]$$

$$EP(B \mid \neg Y \wedge S \wedge K) = [\frac{1}{2}, \frac{1}{2}]$$

$$EP(Y \mid \neg Y \wedge S \wedge K) = [0, 0]$$

$$DU(R \mid \neg Y \wedge S \wedge K) = 0$$

$$DU(B \mid \neg Y \wedge S \wedge K) = \frac{2}{3}$$

$$DU(Y \mid \neg Y \wedge S \wedge K) = 0$$

The value of DU for B above might require some explanation. While learning that ¬Y informs the player that the ball is either red or black, they are no wiser regarding the proportion of black balls in the urn. It is consistent with ¬Y and K that the proportion of black balls is 0, 2/3, or any possible point between. By the Principle of Lesser Uncertainty, a player ought to choose to bet **Red** rather than **Black**, for the same reasons as in the Ellsberg Paradox scenario: the degree of uncertainty of Red is 0, whereas the degree of uncertainty of Black is 2/3. Such preferences are identical to the preferences of an ambiguity-averse player with the standard Ellsberg Paradox preferences.

The interesting difference from an ambiguity-averse player emerges when we consider what happens if the player has decided to invest G. Such a player faces Matrix III. Once again, we assume that **Red** and **Black** have equal expected utilities.

Matrix III

	R	B	Y
Red	£10 – G	0	£10 – G
Black	0	£10 – G	£10 – G

Since the player knows ¬Y, choosing **Red** provides a payout if (R v Y) occurs. Similarly, choosing **Black** provides a payout if (B v Y) occurs. R, B, and Y are mutually exclusive outcomes, relative to the total evidence K. From (1) the fact that practical probabilities satisfy the probability calculus and (2) the evidential probabilities described above, the practical probabilities of (R v Y) and (B v Y) must be:

$$EP(R\ v\ Y \mid \neg Y \wedge S \wedge K) = \frac{1}{2} + 0 = \frac{1}{2}$$

$$EP(B\ v\ Y \mid \neg Y \wedge S \wedge K) = \frac{1}{2} + 0 = \frac{1}{2}$$

(For ease of reading, I have used single real values rather than degenerate intervals.)

The evidence that ¬Y does not remove the equality of probability for the two disjunctions, but it does change the degrees of uncertainty. The degree of uncertainty of

(R v Y) is still zero. In contrast, the degree of uncertainty of (B v Y) is 2/3, because it can only occur if the ball is black, and the agent only knows that between 0 and 2/3 of the balls are black. Consequently, under these conditions, a player following the Principle of Lesser Uncertainty should bet **Red** once again. Therefore, their investment of G makes no difference to their choice, and the Sunk Cost fallacy is not committed.

This is surprising. For reasons given in Subsection 3.1, Matrix III is equivalent (under plausible assumptions) to Matrix IV. Surely, one might think, this is equivalent to the decision problem faced in the standard Ellsberg Paradox scenario and the player whom I am modelling prefers **Black or Yellow** in that version of the scenario. Either some trickery is involved or they should be committed to (Black or Yellow) in Matrix III.

Matrix IV

	R	B	Y
Red or Yellow	£10	0	£10
Black or Yellow	0	£10	£10

However, there is an important difference. From ¬Y, the agent knows that the ball is not yellow and this knowledge removes Y as a possible way that either action can provide its payout. Consequently, choosing **Black or Yellow** involves speculating on the unknown proportion of black balls in the urn, rather than speculating on the known proportion that are black or yellow. This creates a difference in the DU measure.

By contrast, in the original version of the scenario, choosing **Black or Yellow** involves betting on a possibility with a precisely known relative frequency: that the ball is one of the 2/3 that are black or yellow. By comparison, choosing **Red or Yellow** involves betting on a possibility with an *imprecisely* known relative frequency: that the selected ball is one of the 1/3 to 2/3 red or yellow balls. The order of the degrees of uncertainty is the reverse of Matrix IV.

My decision theory does not reject additive invariance (at least in this scenario) because the above reasoning holds even if we remove G from the scenario: what makes the difference is the knowledge that ¬Y, which alters the degrees of uncertainty. Instead, an agent using my decision theory avoids the Sunk Cost fallacy because my decision theory does not produce an expected utility function that fully represents an agent's preferences. There are probabilities and utilities in my decision theory, but they do not suffice to characterise rational preference in my model; it is the combination of probabilities, utilities, and (in some circumstances) degrees of uncertainty that determine what an agent should choose. Since the standard axioms of MEU imply that an agent's preferences *are* fully represented by an expected utility function, it follows that my decision theory violates at least one of those axioms. In this context, what is significant is that my decision theory does not satisfy the Axiom of Independence:

Axiom of Independence: If A_x and A_y have the same consequences when X is believed to be false, then an agent will prefer A_x or A_y if and only if these actions have different expected utilities given X.

A preference can exist in my theory due to a difference in degrees of uncertainty (as measured by the DU function) and so this axiom cannot hold. The advantage is that an agent can have the standard Ellsberg preferences, while also avoiding the Sunk Cost fallacy, as I have described above. I discuss whether this sacrifice of the Axiom of Independence is justified in the next subsection.

The avoidance of the Sunk Cost fallacy in this scenario is a definite strength for my proposal over the Ambiguity Aversion response to the Ellsberg Paradox. Of course, I have not proven that there are no circumstances under which an agent following the Principle of Lesser Uncertainty commits the Sunk Cost fallacy or other fallacies. All I have provided is a detailed argument that they avoid the *prima facie* irrational decision-making that the ambiguity-averse player committed in Subsection 3.1.

6.3 Objections

The most obvious objection to my use of the Principle of Lesser Uncertainty is that it is a deviation from MEU theory. However, such a deviation is a necessary condition of rationalising the standard Ellsberg Scenario preferences, because they are inconsistent with MEU theory. One important difference (aside from the updating method) is that MEU preferences satisfy the Axiom of Continuity:

Axiom of Continuity: It is possible an rational agent's preferences can be fully represented by a continuous expected utility function.

By contrast, on my proposal, rational preference is formalised using both an expected utility function *and* the degree of uncertainty measure. Thus, whereas MEU involves two factors (probability and cardinal utility) my proposal involves three factors: probability, cardinal utility, and uncertainty.

Additionally, as noted earlier, my proposal is inconsistent with the Axiom of Independence. For example, suppose that a player has equal expected utilities for two actions A_1 and A_2 and neither action minimizes DU. If they learn a statement X, such that (a) learning this information does not affect the expected utilities and (b) this information reduces the DU for A_1 below that of A_2, then the player will choose A_1. This violates the Axiom of Independence.

Rejecting Independence does not entail that it is always permissible to violate this axiom. In particular, I am not rejecting this axiom for any cases in which either (a) the expected utilities of the actions are not equal or (b) there are no differences in DU such that an agent can minimize the uncertainty of their actions by choosing one action over another. It *does* entail the rejection of a potentially appealing axiom (though this is one of the most controversial axioms of MEU theory[191]) but rejecting at least one of the axioms is a necessary condition of an intuitive answer in the Ellsberg Paradox, because the standard preferences are inconsistent with the axioms.

The general consequences of the Principle of Lesser Uncertainty are unexplored. It follows that it is entirely plausible that my theory will create new paradoxes, and these paradoxes might be more severe than the Ellsberg Paradox. To be confident in my answer, I would need to discuss a greater range of potential problems. Concomitantly, my answer is

[191] Steele and Stefánsson (2015).

very tentative.

From an MEU perspective, one problem with my proposal is that the rule of maximizing expected utility can no longer be derived from the standard axioms of decision theory. Those who place a great emphasis on an axiomatic derivation of MEU reasoning will regard this feature as a major cost. However, this form of justification of a theory of rationality is only one among many options[192]. One alternative is Carnap's justification of MEU, which uses our intuitions in particular examples as the grounds for choosing among possible decision theories. Carnap justifies the MEU principle via a case-by-case analysis of a set of plausible alternatives[193]. My strategy has been similar to this methodology: I began with an established decision rule (MEU) and considered an example where it fails to match standard intuitions about rational choice. The rule was subsequently modified to address the example. Different supporters of MEU will place different values on the alternative approaches to its justification. Thus, the significance of losing an axiomatic derivation is partly a matter of preference.

One might ask whether the Principle of Lesser Uncertainty is *ad hoc*, given that I have not derived it from subjective preference against uncertainty, nor from any theory of probability. There are least three forms of *ad hoc*-ness charges that one could make: (1) other decision theories could make use of this principle, because it is logically independent of the account of practical probabilities that I gave in Subsection 5.1[194]; (2) there could be other

[192] Hindmoor (2006) p. 182-183.

[193] Carnap (1962) p. 252-279.

[194] In the sense that one could consistently (a) view practical probabilities as exogenously determined within the bounds of the pertinent evidential probabilities and the axioms of additive probability, without also adopting the Principle of Lesser Uncertainty, or (b) adopt the Principle of Lesser Uncertainty but have an otherwise different decision theory e.g. Bayesian decision theory with conditionalization instead of Evidential Probability updating.

principles that do the same work; and (3) the Principle of Lesser Uncertainty lacks independent motivation. I shall deal with these charges separately and in turn.

With respect to (1), it is *logically* consistent to adopt the Principle and otherwise adopt a different decision theory, but mere logical consistency is a very weak criterion for a decision theory. One might also want (*ceteris paribus*) a decision theory to be conceptually parsimonious, in the sense of introducing comparatively few primitive concepts. If a decision theory does not otherwise feature Evidential Probability, then the introduction of the Principle of Lesser Uncertainty and the concomitant primitive concepts from Evidential Probability (the limits of the probability intervals, the rules of Sharpening, the basic semantics of the system etc.) will decrease the parsimony of the system. Unless the alternative system is at least as parsimonious as my proposal and able to handle the same problems in a satisfactory manner, then it will be less parsimonious[195]. For example, importing the Principle of Lesser Uncertainty into orthodox Subjective Bayesian decision theory would effectively double the number of probability systems used in the decision theory. Furthermore, parsimony is only one problem for the introduction of the Principle wholesale into a rival decision theory. (Another might be that the Principle is conceptually incoherent, e.g. because the probability/uncertainty distinction cannot be formulated in the rival decision theory.) Thus, while it is logically consistent to introduce the Principle of Lesser Uncertainty into rival decision theories, it is not necessarily without cost, and there

[195] With respect to the second condition, my proposed decision theory is identical to MEU for most static decision problems, and Kyburg's procedures for updating via Evidential Probability are at least a serious rival to Bayesian conditionalization, so it can apparently handle many problems as least as well as MEU decision theory. Furthermore, since I am more committed to Evidential Probability than any decision-theoretic principle, I am very open to modifying the decision theory on the basis of criticisms of MEU decision theory, though I do not discuss these in this thesis.

would be burdens of proof on someone who undertook such a project[196].

Of course, such a straight introduction of the Principle of Lesser Uncertainty (without any reinterpretation) is not the only option. Nothing I have said in this chapter rules out the possibility of different principles (perhaps very similar principles) being incorporated into an alternative modification of standard MEU decision theory or some other rival to my proposal. Hence charge (2) – the possibility of different principles that can do the same work. To assess all such alternatives is well beyond the scope of this chapter, in which I simply seek to formulate and motivate *one* possible novel answer to the Ellsberg Paradox; the mere existence of rivals is not a strong objection to this aim. However, I can at least put forward an argument as to why the Principle is better (*ceteris paribus*) than some possible rivals, whose comparable principles are unmotivated, which requires answering charge (3).

The distinction between speculation and evidence-based opinion has a pre-theoretical significance: while we seem to make decisions on a more or less arbitrary basis under different circumstances, there seems to be something rational, all else considered, about reasoning when the evidence guides us to a greater extent. This intuition is why, when decision theories prescribe the avoidance of free evidence, philosophers regard this feature as a severely problematic feature of those theories[197]. One motivation for the Principle of Lesser Uncertainty is that it enables a (partial) incorporation of this pre-theoretical significance of

---

[196] One way of doing so without introducing new primitives seems to be the importation of the Principle into an MEU decision theory based on the Objective Bayesianism of Williamson. (I briefly discussed his version of Objective Bayesianism in Subsection 4.3) The reason this would not require new primitives is that he already uses Evidential Probability to provide constraints on probability assignments; the difference from my proposal is that he puts forward further restrictions. A fair evaluation of such a rival would require developing it in detail and considering a wide variety of issues, such as Williamson's arguments for these constraints, and thus it is outside the scope of this thesis. I shall simple acknowledge that this seems to be a promising rival, and not an implausible one.

[197] Bradley and Steele (2016).

guidance by the evidence into formal decision theory: when the Degree of Uncertainty can serve as a tiebreaker between otherwise underdetermined courses of action, then one should choose the course of action with minimal DU. Of course, one does not share the intuition that guidance by the evidence has an independent significance for rationality or one does not agree that the DU measure is a good formalization of a lack of guidance by the evidence, then this reasoning will be unpersuasive. Nonetheless, it does mean that the Principle of Lesser Uncertainty is not *ad hoc* in the sense of lacking an independent motivation: even before Ellsberg formulated his paradox, Knight had an intuition that there was something special about uncertainty, as distinct from probability[198].

Furthermore, this intuition of the comparative rationality of avoiding uncertainty seems to have applications to other problems in formal epistemology. For example, as mentioned in my discussion of Popper's Paradox of Ideal Evidence in Chapter 1 Section 3, one aspect of the paradox that Popper raises is the problem of distinguishing two situations involving the hypothesis that the *n*th toss of a coin with heads and tails on either side will land 'heads':

(A) When you have no idea about the bias/fairness of the coin.

(B) When you have subsequently acquired evidence from a sample (presumably one that is random, large, and otherwise suitable) of coin tosses, which landed 'heads' in half of the tosses.

[198] See Subsection 5.2 for more discussion, as well as Knight (2006, first published in 1921) for Knight's views.

For many epistemic probability systems, your degree of belief might be the same in both scenarios, so that an additional parameter is needed to represent the difference between (A) and (B). One possible candidate for this parameter is the Degree of Uncertainty: given suitable background knowledge, the Evidential Probability that the *n*th toss will land 'heads' could be a very wide interval like [0, 1] in situation (A) and a narrower interval such as [0.45, 0.55] in situation (B)[199]. One might have the intuition that one could be more confident in the choice of 50/50 odds, *ceteris paribus*, in situation (B) than in (A), and the basic intuition behind the Principle of Lesser Uncertainty (i.e. the intuition that there is something comparatively rational in avoiding uncertainty in one's decisions) could motivate this distinction. I shall not develop this suggestion further[200], but the relevant point for my response to the *ad hoc* charge (3) is that the Principle is motivated by an intuition with more general applications than just the Ellsberg Paradox.

Finally, it is important to note that the Ellsberg Paradox is not merely a gambling problem: one can develop similar problems in a vast number of other possible contexts, from medicine to the design of experiments, and from the choice of nuclear strategy to the choice of shoes. For example, you might have a choice between two possible treatments for your brain cancer: the first is estimated to work for patients in your condition in 0-10% of cases, whereas the second treatment is estimated to work in 4-6% of such cases. You have no better information on which to estimate the desirability of adopting either course of treatment. (For

[199] The limits in situation (A) could be narrower if you know that the coin will land heads or tails in *some* tosses. In situation (B), the width of the interval might depend on your ability to estimate the relative frequency of the *n*th toss landing heads, given your sample and the background knowledge. Either value is very sensitive the background knowledge, and thus I stress that these values are illustrative, even though the evidential probabilities are unique once the total evidence has been specified.

[200] For example, a serious development of this suggestion would involve far more detail, a consideration of objections, and a critical comparison with the standard Bayesian answer to this aspect of the Paradox of Ideal Evidence, as well as Popper's own answer, and other alternatives. I do not pretend to have done any of these in this response.

instance, you are indifferent between their respective side-effects, neither will be more expensive for you, and so on.) Many people would have the intuition that the second treatment is preferable, and the Principle of Lesser Uncertainty can rationalise this preference: *if* one is otherwise indifferent between the treatments, then the lesser uncertainty regarding the efficacy of the second treatment can serve as a tiebreaker. Thus, even if it is true that there is no motivation for the Principle of Lesser Uncertainty beyond the Ellsberg Paradox, then the Principle still has a significant motivation, because Ellsberg's problem is not a small or recondite issue.

Beyond the charge that it is *ad hoc*, my answer rules out the possibility of rational indifference between the choices in the Ellsberg Scenario: if the expected utilities are equal, then the standard choices are mandatory. This might seem objectionable, but there does not seem to be a strong intuitive basis for thinking that indifference *must* be possible in Ellsberg's scenario. (If we adopt a strict revealed preference approach, it is impossible to be indifferent in any decision-problem[201].) Furthermore, I am not rejecting indifference in very similar but distinct scenarios. For instance, a rational agent would be indifferent for both choices of bets if she *knew* that 1/3 of the balls in the urn are black, as opposed to *speculating* this information. Additionally, sacrificing the possibility of rational indifference in one circumstance and for one choice of practical probabilities seems worth the benefits of preserving most people's intuitions.

Finally, one might wonder what my answer offers to those who have already rejected MEU theory. There are at least several useful properties of my answer for such decision theorists. Firstly, the rejection is not *ad hoc*, because it is grounded in a pre-theoretical

[201] Rothbard (2009) p. 307.

intuition about rational decision-making: *ceteris paribus*, it is better to avoid uncertainty; put differently, evidence-based decisions are superior, all else being equal, to speculation-based decisions. Secondly, I have specified exactly when we can depart from MEU theory and when we cannot. (There are other paradoxes, like the Allais Paradox, which might require additional departures from MEU theory, but these are different issues.) Thirdly, unlike the Ambiguity Aversion response, my answer does not seem to raise the problems of the Sunk Cost fallacy, and offers a path that avoids the Ellsberg Paradox while not requiring that one take a stand on that issue. Finally, my answer offers one way of doing justice to the intuition that there is something "cautious" about the standard preferences, in that these preferences are due to the avoidance of uncertainty.

CONCLUSION

In this chapter, I have offered an answer to the Ellsberg Paradox that uses the DU measure and Evidential Probability to develop an alternative normative decision theory for a computationally unbounded ideal reasoner. This theory enables one to retain something that is similar to MEU theory (especially in contexts of rich background knowledge or static decision problems) while avoiding the Ellsberg Paradox. It also provides an interesting illustration of how Keynes's concept of the weight of argument can be significant for decision theory: *sometimes*, a greater quantity of relevant evidence about some conjecture can reduce the degree of uncertainty for that conjecture, and this difference can sometimes affect what constitutes rational decision-making.

CHAPTER 4: FORMALISM, RELIABILITY, AND THE NEW RIDDLE OF INDUCTION

Many of the early confirmation theorists, including Kyburg[202], aimed to develop purely formal confirmation theories, which are those that involve no discriminations among the predicates or individuals in the evidence or hypothesis. However, most confirmation theorists now regard this project as unachievable due to Goodman's New Riddle of Induction (NRI)[203]. I shall argue against this consensus: formalists can answer the NRI by using a confirmation theory that incorporates the importance of reliable evidence.

In Section 1, I describe the formalist project and the relevance of the NRI to formalism. In Section 2, I consider the NRI as a challenge to Hempel's confirmation theory; I focus on this theory because it was Goodman's initial target and if the NRI has been rigorously established as problematic for any formalist, it is Hempel. Finally, in Section 3, I use the concept of the reliability of evidence to develop a formalist response to the NRI.

---

[202] Kyburg (1961) p. 40.

[203] Goodman (1946) and Chapter III of Goodman (1983).

SECTION 1: FORMALISM AND THE NRI

1.1 Formalist Theories of Confirmation

In the 20th century, many philosophers of science attempted to develop a purely formal definition of confirmation. The efforts of Keynes[204], Carnap[205], and Hempel[206] are the most famous. Goodman directs his attention towards the theories of Carnap, Hempel, Paul Oppenheim, and Olaf Helmer in particular[207]. These philosophers had many disagreements, but one idea they share is that confirmation can be analysed as a formal relationship between statements (or sentences or propositions or whatever analogous *relata* one chooses) without any restrictions on the non-logico-mathematical terms in the sentences related, like ‘flames’ or ‘hot’ in ‘Most flames are hot’. Thus, they aimed to analyse confirmation and disconfirmation in a way that is independent of the subject matter of the hypothesis and evidence[208]. Additionally, in formalism, no predicates receive special status (like the status of “natural kinds”) from the confirmation theory itself. Of course, a formalist might favour some predicates on grounds that are exogenous to the confirmation theory. For example, they might favour ‘oxygen’ over ‘phlogiston’ on empirical grounds, but this preference is not part of their confirmation theory as such. Hempel provides a list of the terms which he will use in his

---

[204] Keynes (1921) p. 54-55.

[205] Carnap (1962) p. 19.

[206] Hempel (1943).

[207] Goodman (1946) p. 383.

[208] Hempel (1945a) p. 9. Stove (1965) notes that this ambition is in tension with Hempel’s logical empiricism, because Hempel would both like to have a purely formal definition of confirmation *and* to use confirmability to demarcate empirical statements from other statements; so he restricts the evidence to observation-statements.

analysis, which he calls “syntactic” terms: the universal and existential quantifiers, individual constants, individual variables, predicate constants, parentheses, and commas[209]. Some formalists, like Keynes and Carnap, also incorporate a probability function and auxiliary mathematical terms into their theories.

Even some critics of formalism have acknowledged its attraction. Helen Longino notes that if confirmation is a purely formal relation, then analysing evidential relations in science would be a comparatively straightforward project, because many salient philosophical questions about science would become formally decidable[210]. If Longino is correct about formalism’s potential and if Goodman’s NRI successfully proves the impossibility of a formalist confirmation theory, then his accomplishment seems regrettable.

1.2 The New Riddle of Induction

I shall now pinpoint Goodman’s charges against Hempel’s theory, before generalising them to all formalist confirmation theories, and briefly explaining their historical significance. Goodman argues that there are an indefinite number of predicates that are unsuitable for inductive inferences (they are not “projectable”) because they generate unacceptable paradoxes, but these predicates cannot be excluded from inductive inference on purely formal grounds. Therefore, formalist theories are incomplete, so non-formal considerations must be part of any satisfactory confirmation theory. The most famous of these predicates is ‘grue’, which can be very loosely defined as ‘green if observed prior to *t*

[209] Hempel (1943) p. 123.

[210] Longino (1990) p. 23-24.

and blue if observed after *t*'. I shall discuss its definition in more detail in Section 2.

Branden Fitelson provides a helpful taxonomy of Goodman's criticisms of Hempel's theory:

(1) ***The Qualitative Claim***: According to Hempel's theory, reports of green emeralds prior to *t* confirm both 'All emeralds are grue' and 'All emeralds are green'.

(2) ***The Quantitative Claim***: According to Hempel's theory, reports of green emeralds prior to *t* confirm the hypotheses 'All emeralds are grue' and 'All emeralds are green' to an *equal extent*.

(3) ***The Triviality Claim***: According to Hempel's theory, anything confirms anything[211].

For the purposes of this chapter, I shall ignore the triviality claim (3), because unlike (1) or (2), it has not gained widespread acceptance. Goodman simply asserts, without argument, that (3) follows from (1) and (2).

The qualitative claim (1) requires some elaboration. The problem is that if 'All emeralds are green' and 'All emeralds are grue' are contraries, and some evidence confirms both hypotheses according to Hempel's theory, then contrary hypotheses can be confirmed by the same evidence according to his confirmation theory. This criticism can be generalised to all formalist theories:

[211] Fitelson (2008) p. 617-618.

(4) ***The Qualitative NRI***: In all purely formal confirmation theories, it is possible for the same evidence to confirm contrary hypotheses.

Similarly, (2) generalises into:

(5) ***The Comparative NRI***: In all purely formal confirmation theories, it is possible that the same evidence can equally support contrary hypotheses, in circumstances in which they are clearly *not* equally supported[212].

Most confirmation theorists believe that 'grue' proves that a purely formal analysis of confirmation cannot succeed. Thus, Goodman's view on formalism's viability has become the consensus. To give a small sample, Susan Haack[213], James Ladyman[214], George Couvalis[215], Michael Williams[216], David Stove[217], Timothy McGrew *et al*[218], John Vickers[219],

[212] The last clause is required, because some confirmation theories allow for the equal confirmation of contraries, but obviously there are some cases where contraries are clearly *not* equally confirmed.

[213] Haack (1996) p. 369.

[214] Ladyman (2002) p. 43.

[215] Couvalis (1997) p. 46.

[216] Williams (2001) p. 213.

[217] Stove (1986) p. 139.

[218] McGrew, Timothy *et al* (2009) p. 382.

[219] Vickers (2016).

Richard Miller[220], Ruth Weintraub[221], Alan Weir[222], and Timothy Williamson[223] all agree that the NRI is fatal to the formalist project. A rare exception in the literature is C. J. Nix and J. B. Paris[224]. Another dissenting view is developed by J. B. Paris and Alena Vencovská[225]. However, it is safe to say that most philosophers regard the formalist programme as hopeless because of Goodman's criticisms.

I shall make some clarificatory points before continuing. I am not presenting a general defence of inductive inference, so I shall set aside a number of related issues, like Hume's Problem of Induction. (Goodman also sees Hume's problem as distinct from the NRI[226].) I shall also put aside other underdetermination problems, like the issues regarding curve-fitting or the Duhem-Quine Problem.

SECTION 2: HEMPEL'S CONFIRMATION THEORY AND THE NRI

Despite the consensus regarding formalism and the NRI, there has never been a general proof that the NRI poses a problem for all formalist theories. In his most influential discussion of the Riddle, Goodman simply criticises Hempel's theory[227]. In the absence of a received general argument against formalism, I shall begin by following Goodman and

---

220 Miller (1995) p. 223.

221 Weintraub (2008) p. 147.

222 Weir (1995) p. 27.

223 Williamson (1998) p. 91.

224 Nix and Paris (2007) p. 738.

225 Paris and Vencovská (2015) p. 5-6.

226 Goodman (1983) p. 62-66.

227 Goodman (1983) p. 72-73.

examining the NRI as an objection to Hempel's theory. Even if I am mistaken that the NRI poses no direct problem for this theory, Goodman would still not have proven the unviability of formalism. However, if the NRI has been proven to be problematic for any (plausible) formalist theory of confirmation, it is Hempel's, and if the NRI poses no problem for Hempel, then this substantially increases the apparent propitiousness of formalism.

I shall argue against both Goodman's Qualitative Claim and his Quantitative Claim. To do this, I shall prove that contrary hypotheses cannot be confirmed in Hempel's theory, before examining some popular versions of the NRI, and I shall finish by noting that the NRI highlights some major limitations of Hempel's system. These limitations give a formalist some independent reasons to seek a better theory.

2.1 Hempel's Confirmation Theory

Hempel offers both a quantitative theory[228] and a qualitative theory[229]. Quantitative theories of confirmation assign degrees of confirmation, like 'H is confirmed by E to a degree of confirmation *r*.' In contrast, qualitative theories only licence classificatory claims like 'H is confirmed by E'. I shall not discuss Hempel's quantitative theory, because Goodman directs the NRI at Hempel's qualitative theory.

Hempel takes sentences of a first-order language-schema L as the basic *relata* of his qualitative theory. L is like a natural language whose non-logical terms have been removed

---

[228] Hempel and Oppenheim (1945) p. 108.

[229] Hempel (1943).

and whose sentences have been transformed into first-order logic[230]. For these *relata*, Hempel proposes the following definition:

Hempel's Qualitative Definition of Confirmation: A statement H is confirmed by E if and only if either (i) E implies the development of H for the individual constants in E or (ii) E implies H[231].

The ‘development’ of H for a class of individual constants is the sentence that results from replacing the variables in H with the individual constants in E. For example, $(Ra \rightarrow Ba)$ is the development of $\forall x(Rx \rightarrow Bx)$ for the individual constant *a*.

In Hempel’s theory, E disconfirms H when E confirms ¬H. If E neither confirms nor disconfirms H, then E is neutral with respect to H[232].

An example of the satisfaction of clause (ii) is the following: let H be ‘There is a green frog’, translated as $\exists x(Fx \wedge Gx)$. The statement E is ‘Alec is a green frog’, translated as $(Fa \wedge Ga)$. Since E implies H, it confirms H. Clause (ii) is satisfied in this example: let H be ‘All things are positively charged or negatively charged’, translated as $\forall x(Px \vee Nx)$. Let E be ‘Particle *a* is positively charged or negatively charged’, translated as $(Pa \vee Na)$. Since E is the development of H for *a*, it confirms H. Finally, I shall give an example of confirmation in this system in which the evidence implies a development. If H is $\forall x(Px \vee Nx)$ and E is Pa, then E

[230] Hempel (1943) p. 124.

[231] For simplicity, I have only provided Hempel’s definition for non-analytic statements.

[232] Hempel (1943) p. 127.

implies the development of H, because Pa implies (Pa v Na), and thus E confirms H.

A peculiar feature of Hempel's definition is that if E is contradictory, then E confirms every statement in L. Hempel uses 'implies' to mean strict implication, but a contradiction strictly implies every statement: via clause (i), this results in contradictions 'confirming' every statement in L. However, Hempel notes that this is an inessential feature of his system[233]. It is intuitive and simpler to exclude such cases, so I shall define 'Hempel-confirms' as:

Revised Hempelian Qualitative Definition of Confirmation: A statement H is *Hempel-confirmed* by E if and only if (i) E is consistent and either (ii) E implies the development of H for the individual constants in E or (iii) E implies H.

2.2 Proof That Hempel's System Satisfies the General Consistency Condition

I shall begin my discussion by providing a general proof that the Hempel-confirmation of contraries is impossible. It follows that Goodman's claims regarding Hempel's system are false: it is not the case that Goodman's examples prove that contraries can be confirmed in Hempel's theory, nor *equally* confirmed in Hempel's theory. Hempel requires that confirmation theories satisfy the following condition:

General Consistency Condition (GCC): The set of statements that are confirmed by a

[233] Hempel (1945b) p. 103. He also excludes contradictory evidence in his quantitative theory: see Hempel and Oppenheim (1945) p. 102.

consistent sentence must form a consistent conjunction[234].

Clearly, Hempel intended his definition to exclude contrary hypotheses being Hempel-confirmed. Hempel actually suggests a proof that his definition satisfies the GCC[235], but subsequently he strongly hints that his system fails to satisfy the GCC[236]. Hempel's suggestion of a proof is a reference to an earlier proof in that article; in the earlier proof, he proves that a *different* analysis of confirmation satisfies the GCC. Since a proof does not seem to exist in the literature, I shall provide an explicit proof of the GCC for his system.

A preliminary clarification of Hempel's concept of the "development" of a sentence will be useful. The development of a hypothesis H is a statement that entails H when conjoined with the premise that only the individuals in the development exist. The development is an atomic statement of L. When the development is consistent, it can only be entailed by consistent statements, like any other statement. If there is a set of hypotheses and the evidence-statement E entails the development of every member of the set, then there will be a statement that is the development of the conjunction of the hypotheses in the set.

Key

E: Any consistent evidence-statement.

SH: Any set of inconsistent hypotheses.

---

[234] Hempel (1943) p. 127.

[235] Hempel (1943) p. 142.

[236] Hempel (1965) p. 50-51.

Premises

(1) If E Hempel-confirms every member of SH, then E entails every member of SH or E entails the development of every member of SH[237].

(2) If E entails every member of SH, then it entails the conjunction of the members of SH.

(3) If E entails a development of every hypothesis in SH, then there is a consistent statement that is the development of the conjunction of the hypotheses in SH for the individuals mentioned in E.

(4) The conjunction of the members of SH is contradictory.

(5) A contradictory statement has no consistent development.

(6) No contradictory statements can be entailed by E.

Proof

Claim: There is no E such that E Hempel-confirms every member of SH.

By a categorical syllogism from (4) and (6):

[237] If E entails some members of SH and E entails the development of other members, then E entails the development of every member, since if E implies H, then it must imply the development of H for the individuals mentioned in E.

(7) E does not entail the conjunction of the members of SH.

By *modus tollens* from (2) and (7):

(8) E does not entail every member of SH.

By *modus ponens* from (4) and (5):

(9) The conjunction of all the members of SH has no development.

By *modus tollens* from (3) and (9):

(10) E does not entail the development of every member of SH.

By conjunction introduction and De Morgan's Theorem from (8) and (10):

(11) E neither entails every member in SH, nor does it entail the development of every member of in SH.

By *modus tollens* from (1) and (11):

(12) E does not Hempel-confirm every hypothesis in SH.

Since E and SH are arbitrarily selected, (12) can be generalised for any consistent statement and any set of contraries. Thus, Hempel’s confirmation theory satisfies the GCC and so Goodman’s Qualitative Claim must be false. By extension, Goodman’s Quantitative Claim must also be false, since contraries cannot be equally Hempel-confirmed if they cannot be Hempel-confirmed.

Given that the Qualitative NRI involves the logically weaker thesis, it is the natural place to look for counterexamples to my metatheoretic claim. Before looking at individual versions of the Qualitative NRI, I shall provide a complete list of the three ways that a statement E might fail to Hempel-confirm the conjunction of two putatively contrary hypotheses H_1 and H_2:

(i) E is inconsistent.

(ii) E does not Hempel-confirm both H_1 and H_2.

(iii) E does not Hempel-confirm either H_1 or H_2.

(iv) H_1 and H_2 are not contraries.

I shall argue that at least one of (i) to (iv) is true in all the standard versions of the Qualitative NRI.

2.3 The Colour-Change NRI

In this version of the Qualitative NRI, one defines ‘grue’ so that grue objects must change colour. Tom Settle[238], John Wright[239], David Armstrong[240], and Jerold Abrams[241] use this type of definition. For an emerald to satisfy the colour-change version of ‘grue’, it must be observed to be green before some future time *t* (like 3,000 AD) and subsequently change colour to blue at *t*. I shall now define this version of ‘grue’ formally:

Key

CCG: Colour-change grue.

OT: First observed before 3,000 AD’.

G: Green.

B: Blue.

EM: Emerald.

D_1: x is CCG $\leftrightarrow$ ((OTx ^ Gx) v (¬OTx ^ Bx))

(Throughout this chapter, the application of a colour predicate to an object means that the object is monochromatically that colour. I shall also ignore, in line with tradition, that if ‘emerald’ has its usual gemmological sense, then ‘All emeralds are green’ is true by definition, because ‘emeralds’ are defined by gemmologists as green beryls.)

238 Settle (1974) p. 743.

239 Wright (1991) p. 41.

240 Armstrong (1983) p. 57.

241 Abrams (2002) p. 544.

The hypothesis ‘All emeralds are colour-change grue’ is typically formalized as:

(1) $\forall x(EMx \rightarrow CCGx)$

- which is logically equivalent to -

(2) $\forall x(EMx \rightarrow ((OTx \wedge Gx) \vee (\neg OTx \wedge Bx)))$

The development of (2) for *a* is:

(3) $EMa \wedge ((OTa \wedge Ga) \wedge (\neg OTa \wedge Ba))$

- and the evidence-statement used in the colour-change NRI is typically -

(4) $(EMa \wedge OTa \wedge Ga)$

The colour-change challenge to Hempel is that, allegedly, in Hempel’s theory the evidence-statement (4) confirms the hypothesis:

(5) $\forall x(EMx \rightarrow Gx)$

- as well as the hypothesis (1). Thus, if this version of the NRI is correct, Hempel’s system allows for the confirmation of contraries. It was this version of the NRI (using 2000 AD as *t*) that inspired Kyburg's joke that Goodman’s Riddle “… is thus surely one of the most important problems to come out of recent discussions of inductive logic. It is also one of the most pressing, since there are now only thirty years left in which to solve it."[242]

However, Robert T. Pennock points out that, while it is true that (4) Hempel-confirms (5), it is not true that (4) Hempel-confirms (3)[243]. The evidence-claim (4) entails (EMa → Ga), which is the development of (5) for the individual *a*, but it does not entail (3), which is the development of (1) for *a*. Informally, the evidence (4) tells us that *a* is a green emerald, but not that it will change colour after *t*, so that there is no instance of the grue-hypothesis in (4) according to Hempel's theory.

Consequently, the Colour-Change NRI is not an example of the Hempel-confirmation of contraries. Frank Jackson makes a similar point regarding the Straight Rule of Induction[244].

2.4 The Disjunctive NRI

Goodman's version of 'grue' is not the colour-change version[245]. He defines 'grue' as:

D_2: *x* is grue ↔ ((OTx → Gx) ^ (¬OTx → Bx))

[242] Kyburg (1970) p. 174.

[243] Pennock (1998) p. 107-110

[244] Jackson (1994) p. 81. As Jackson defines it, Straight Rule of Induction is the principle that instances of a predicate F that correlate with a predicate G are *prima facie* evidence (perhaps with background information) that other instances of F correlate with G.

[245] Goodman (1983) p. 74.

This definition is also used in some other philosophers' presentations of the NRI[246]. Something will be 'grue' on the D_2 definition if it satisfies one of the following truth-conditions:

(i) It is first observed before *t* and it is green.

(ii) It is not first observed before *t* and it is blue.

The truth-conditions of D_2 are the same as the truth-conditions of D_3:

D_3: x is grue $\leftrightarrow ((OTx \wedge Gx) \vee (\neg OTx \wedge Bx))$

- which perhaps makes the logical form of Goodman's version of 'grue' more manifest and which some philosophers use when discussing the NRI[247].

Using Goodman's definition, we can formalise the alleged contraries 'All emeralds are green' and 'All emeralds are grue' as:

H_1: $\forall x(EMx \rightarrow Gx)$

H_2: $\forall x(EMx \rightarrow GRx)$

[246] Antony (2004) p. 12 and Huemer (2001) p. 379.
[247] E.g. Pennock (1998) p. 103 and Horwich (1982) p. 67.

- using ‘GR’ as an abbreviation of $((OTx \wedge Gx) \vee (\neg OTx \wedge Bx))$.

However, H_1 and H_2 are not contraries: if there were no emeralds, then both would be true, since any material conditional is true when the antecedent is false. Adding existential import to the hypotheses avoids this particular problem:

H_3: $\forall x(EMx \rightarrow Gx) \wedge \exists x(EMx)$

H_4: $\forall x(EMx \rightarrow GRx) \wedge \exists x(EMx)$

- so that both would both be false if there were no emeralds.

The evidence that Goodman considers is:

(5) $(EMa \wedge Ga \wedge OTa) \wedge (EMb \wedge Gb \wedge OTb) \wedge (EMc \wedge Gc \wedge OTc) \ldots (EM\mathit{n} \wedge G\mathit{n} \wedge OT\mathit{n})$

From (5), the developments of H_3 for any given individual *i* in the evidence-report can be deduced:

(6) $(EMi \rightarrow Gi) \wedge EMi$

- and similarly for H_4 -

(7) $(EMi \rightarrow GRi) \wedge EMi$

Hence, the statement (5) Hempel-confirms both hypotheses. However, the hypotheses are *still* not contraries, because it is logically possible that all the emeralds that will ever exist are green and first observed before *t*. If that were true, then both H_3 and H_4 would be true. For example, imagine a universe in which (a) all matter becomes grey goo forever at *t*, (b) no new matter ever comes into existence, and (c) all the emeralds that ever existed were first observed before *t*. In such a universe, all emeralds were green, but they were also all grue, because they were green and observed prior to *t*. Therefore, in Hempel's system, evidence can confirm H_3 and H_4, but these hypotheses are not contraries.

Pennock also notes that there is no incompatibility between the two hypotheses, but tries to reformulate the NRI to deal with this logical point[248]. He asserts that if we add our background knowledge that there will be emeralds that we do not observe prior to *t*, then the hypotheses will be contraries and confirmed[249]. Pennock claims that we can conjoin our background knowledge with the evidence, so that we have contraries that will be Hempel-confirmed. In particular, we can add our background knowledge that there exists one object *b* that is an emerald and it is not first observed prior to *t*. The new evidence-statement is:

(8) $EMa \wedge Ga \wedge OTa \wedge EMb \wedge \neg OTb$

We now need some contrary versions of the green and grue hypotheses. By conjoining $(EMb \wedge \neg OTb)$ with the H_3 and H_4, we obtain:

---

[248] Pennock (1998) p. 110.

[249] Pennock (1998) p. 113.

(9) $\forall x(EMx \rightarrow Gx) \wedge EMb \wedge \neg OTb$

(10) $\forall x(EMx \rightarrow ((OTx \wedge Gx) \vee (\neg OTx \wedge Bx))) \wedge EMb \wedge \neg OTb$

- which are contraries, because (9) implies that *b* is green and (10) implies that *b* is blue.

However, while Pennock is correct that (9) and (10) are contraries, he is incorrect in claiming that they are Hempel-confirmed by (8). Firstly, it is obvious that (8) does not entail (9) or (10). Secondly, (8) does not entail the developments of (9) and (10) for the individuals mentioned in (8). These developments are the following:

(11) $(EMa \rightarrow Ga) \wedge (EMb \rightarrow Gb)$

(12) $(EMa \rightarrow ((OTa \wedge Ga) \vee (\neg OTa \wedge Ba))) \wedge (EMb \rightarrow ((OTb \wedge Gb) \vee (\neg OTb \wedge Bb)))$

(8) does not entail (11) or (12), because it is consistent with the falsity of either statement. For example, consider the logically possible world in which *b* is red. In this world, (8) might be true, but (11) and (12) are false. Therefore, (8) does not entail (11) or (12). Since (8) does not entail (9) or (10) or their developments, it follows that (8) does not Hempel-confirm these hypotheses. Pennock has not provided an example of the confirmation of contraries in Hempel's system.

One might object that (8) entails $(EMa \wedge Ga \wedge OTa)$, and this Hempel-confirms the green and grue hypotheses. It is true that (8) entails this report about *a* and that this report confirms both of the hypotheses. However, the fact that (8) entails a report that Hempel-confirms these hypotheses does not imply that (8) also Hempel-confirms the hypotheses. In Hempel's system, confirmation does not 'flow up' the entailment relation. Formally, it is possible in Hempel's system that Φ entails χ and χ confirms Ψ, but nonetheless Φ does not confirm Ψ[250]. For example, in Hempel's system, the hypothesis:

(13) EMa ^ Ga ^ OTa ^ EMb ^ ¬Gb ^ ¬Bb

- also entails (EMa ^ Ga ^ OTa), but it *disconfirms* (9) and (10). Thus, the fact that (8) entails (EMa ^ Ga ^ OTa) does not imply that (8) Hempel-confirms (9) or (10).

However, Pennock's arguments do highlight a genuine limitation of Hempel's system. A Hempelian will presumably want to make predictions based on the formal fact that the green hypothesis H_3 is confirmed in Hempel's system. Yet, Carnap[251] and Hooker[252] independently discovered that Hempel's system provides no help in selecting among contrary predictions. The good news for Hempel is that he does not propose any method by which his system can be used to make contrary predictions. The bad news is that Hempel does not provide *any* method for making predictions using formal facts about the confirmation relations that hold in his system. Thus, a Hempelian could not directly use his confirmation theory to predict that an emerald that has not been observed prior to *t* will be blue, but nor

[250] This feature of Hempel's system can be missed due to the occasional (and misleading) description of his system as a "hypothetico-deductive" confirmation theory, which suggests that confirmation 'flows up' the deductive consequence relation in his system. Instead, Hempel-confirmation 'flows down', in that E confirms H_1 if E confirms H_2 and H_2 implies H_1. Hempel also proved the problems of systems in which confirmation flows both ways: see Hempel (1945b) p. 104.

[251] Carnap (1962) p. 480.

[252] Hooker (1968) p. 244-245.

could she directly use the Hempelian theory to predict that the emerald is green. This fact about the application of Hempel's system is a major limitation, but it is distinct from the NRI, which is a problem that contraries can be confirmed *within* Hempel's system. I shall discuss more limitations in Subsection 2.6.

In my discussion above, I have interpreted Pennock's claim as an assertion that we can confirm contrary hypotheses in Hempel's system. However, one could also interpret his revision of the Riddle in the following way: suppose that M is a claim about Hempel's theory that 'H_3 and H_4 are both confirmed by E according to Hempel's theory'. In other words, M is the formal fact that in Hempel's system, it is possible for reports of green emeralds to confirm both 'There are emeralds and that they are all green' and 'There are emeralds and that they are all grue'. If we combine M with our background knowledge, then we know that these hypotheses cannot both be true, yet both are confirmed by the same evidence in Hempel's system. Therefore, if M is combined with our background knowledge, Hempel's system would underdetermine a theory-choice between two hypotheses that are consistent but (in a non-logical sense) rivals.

Pennock's problem seems to be different from the problem above, because he talks about a "background premise" that is presumably being conjoined with the evidence *in* Hempel's system, as I have interpreted his position above[253]. However, this is still an interesting problem concerning the application of the formal facts of Hempel's system (like M) and their usage in areas like theory-choice. In contrast to formalists like Carnap[254], Hempel does not provide a discussion of how to use his confirmation theory for resolving

---

[253] Pennock (1998) p. 113.

[254] Carnap (1947) and Carnap (1962) p. 202-215.

such methodological issues. As Fitelson argues, it is possible for a defender of Hempel's theory to take a number of views on how M should be used in such examples[255]. Unless it is proven that there is no way that Hempel's system can be integrated into a methodology that addresses theory-choice, then this problem has not proven to be fatal for Hempel's system.

One can make a comparison to certain features in the formal parts of other methodologies. For instance, according to Rudolf Carnap, Imre Lakatos, and Karl Popper, the prior probability of a contingent universal generalisation over an infinite number of individuals is zero[256]. However, this formal fact does not commit them to the claim that one *should disbelieve* all contingent universal generalisations over an infinite number of individuals. Similarly, Hempel could reasonably argue that that there are grounds in our background knowledge (like the fact that Goodmanian hypotheses would not have been successful in the history of science) to overcome any underdetermination problems caused by combing M with our background knowledge.

One might think that a formalist who made such a response is begging the question. Yet Pennock's version of the NRI requires taking our background knowledge for granted and conjoining it with M. Otherwise, there is no reason to regard H_3 and H_4 as leading to contrary predictions about the emeralds that are unobserved prior to *t*. Put another way, the hypothesis:

H_3: $\forall x(EMx \rightarrow Gx) \wedge \exists x(EMx)$

[255] Fitelson (2008) p. 613.

[256] Carnap (1962) p. 570-571, Lakatos (1980) p. 3, and Popper (1980) p. 363.

- implies that there exists some *x* such that (EMx ^ Gx), but it does not imply that *x* or any other emerald exists after *t*. The most that we can derive from H_3 about an object *b* that is not observed prior to *t* is the conditional prediction:

(14) ((EMb ^ ¬OTb) → Gb)

Thus, if *b* is an emerald that is unobserved prior to *t*, then it will be green. Analogously, the grue hypothesis H_4 implies that there is some *x* such that (EMx ^ GRx), but it is silent on the existence of emeralds after *t*. For *b*, the grue hypothesis ∀x(EMx → GRx) ^ ∃x(EMx) only implies that:

(15) ((EMb ^ ¬OTb) → GRb)

- which in turn implies-

(16) ((EMb ^ ¬OTb) → Bb)

By themselves, (14) and (16) are compatible, since they will both be true if their antecedent is false. A sceptical problem only exists given our background knowledge that the antecedents of both conditionals are true, so background knowledge is available from the outset of Pennock's version of the paradox, so Hempel would not beg the question by using background knowledge to decide how to use M for theory-choice.

Therefore, even if we introduce background knowledge into the analysis, it is still not the case that we can use the disjunctive version of ‘grue’ to confirm contraries with the same evidence in Hempel’s system. However, the debate over the Disjunctive NRI does highlight some serious limitations of that system. In particular, Hempel has a large number of unanswered questions regarding the broader application of his system. Such questions are important, but they are distinct from Goodman’s Riddle.

2.5 The Predictive NRI

Goodman phrases much of his most influential discussion about the NRI as a problem about confirming predictions[257]. His assertion is that it is possible to Hempel-confirm:

P_1: EMb ^ Gb

- and -

P_2: EMb ^ Bb

- with the statement -

E: EMa ^ Ga ^ OTa

Goodman’s implicit argument seems to be that, in Hempel’s system, E confirms the hypotheses:

[257] Goodman (1983) p. 74.

H_1: $\forall x(EMx \rightarrow Gx)$

H_2: $\forall x(EMx \rightarrow GRx)$

- and if (i) H_1 entails P_1 and (ii) H_2 entails P_2, then E confirms both P_1 and P_2 because, in Hempel's system, the Special Consequence Condition is satisfied:

Special Consequence Condition: If A confirms B, then A confirms every consequence of B[258].

Put another way, confirmation 'flows down' the consequence relation in Hempel's system: if a prediction is entailed by a hypothesis, then that prediction is confirmed by all the evidence that confirms that hypothesis: if E confirms H_1 and H_1 entails P_1, then E confirms P_1. Similarly, if E confirms H_1 and H_1 entails P_2, then E confirms P_2. Consequently, it would be possible to confirm a set of contraries in Hempel's system.

The mistaken premise in this argument is that H_1 does *not* entail P_1 and H_2 does *not* entail P_2. For example, imagine a universe in which *a* is the only emerald that ever exists and *a* is a green emerald observed prior to *t*. In that universe, H_1 and H_2 are true, but P_1 and P_2 are both false, because they imply the existence of an emerald *b* that never exists. Therefore, the hypotheses do not entail the predictions.

One might try to salvage Goodman's implicit argument by adding background

[258] Hempel (1943) p. 142.

knowledge into the scenario. For instance, we can conjoin our knowledge that *b* is observed after *t* with the hypotheses, and then they will entail their respective predictions. However, the resulting conjunctions are not confirmed by E. The hypotheses:

H_5: $\forall x(EMx \rightarrow Gx) \wedge EMb \wedge \neg OTb$

H_6: $\forall x(EMx \rightarrow GRx) \wedge EMb \wedge \neg OTb$

- are not confirmed by E, since E does not entail the development of the last two conjuncts in H_5 and H_6, because Hempel defines the development of atomic statements like EMb as the atomic statements themselves and E does not entail EMb, as it only describes *a*.

One might try to avoid this problem by adding ($EMb \wedge \neg OTb$) to E. However, the resulting evidence-statement:

E´: $EMa \wedge Ga \wedge OTa \wedge EMb \wedge \neg OTb$

- does not confirm H_5 or H_6, for reasons that are analogous to those given for (8), (9) and (10) in the previous subsection: E´ is silent on the colour of *b*, and this content is needed to deduce the developments of H_5 and H_6.

One could try to keep modifying the scenarios and definitions to Hempel-confirm contrary hypotheses with the same evidence. However, unless my argument that Hempel's system satisfies the GCC is unsound, this is a hopeless exercise. The predicate 'grue' poses no direct problem to Hempel's system. However, my discussion has already touched on some of the limitations of the theory, and I shall now discuss these in more detail.

2.6 The Limitations of Hempel's System

A familiar complaint regarding Hempel's theory is that it is extremely limited. Carnap presents a wide variety of types of cases in which E intuitively confirms H, but it does not in Hempel's system[259]. For instance, suppose that E is '95% of observed swans are white' and H is 'Most swans are white'. E does not entail H and E is not a development of H, so E does not Hempel-confirm H. In general, his definition of confirmation is clearly too narrow.

My discussion of the NRI raises two further problems with Hempel's system. I shall discuss each in turn in this subsection.

2.6.1 Comparative Judgements

My arguments imply that Goodman's Quantitative Claim (that the green and the grue hypotheses can be *equally* confirmed by the same evidence) is false in Hempel's system, but there is an all-too-quick argument that will do the same work. As Stove notes[260], Hempel's definition of confirmation is purely classificatory, in the sense that it only enables classificatory judgements like:

[259] Carnap (1962) p. 480-481.

[260] Stove (1986) p. 138.

J_1: $(EMa \wedge Ga \wedge OTa)$ confirms both $\forall x(EMx \rightarrow Gx)$ and $\forall x(EMx \rightarrow GRx)$.

- but not -

J_2: $(EMa \wedge Ga \wedge OTa)$ confirms $\forall x(EMx \rightarrow Gx)$ *to the same degree as* $\forall x(EMx \rightarrow GRx)$

Aside from distinguishing cases of confirmation from cases of disconfirmation and neutrality, Hempel's theory provides no comparative guidance. Consequently, Goodman's Quantitative Claim must be false. However, it is false because of a *limitation* of Hempel's system.

This limitation is very significant, because it entails that Hempel's system is silent about many important questions. For instance, do 3,000 developments of a hypothesis confirm it more strongly than one development? If so, how much more strongly? Does this difference depend on the available background knowledge? Hempel's definition is silent on these issues. In the context of the NRI, this limitation entails that Hempel's theory cannot tell us whether the green hypothesis is better confirmed than the grue hypothesis, which is an issue that most philosophers would expect a confirmation theory to resolve.

2.6.2 The General Consistency Condition

That the set of statements that are confirmed by a consistent statement must itself be consistent (Hempel's GCC) can seem like an intuitive requirement. However, as Carnap argues, it is much too strong[261]. For example, suppose that the sample mean in our data for the height of British adult males is 183 cm. Assume that our data confirms the statistical generalization that average British adult male height is 183 cm. It is still perfectly plausible that there is a slight difference between the sample mean and the population mean, such that actual British adult male average height is close to 183 cm, but not *exactly* 183 cm. The sample report intuitively confirms each of the hypotheses stating that the population mean is 182 cm, 184 cm, 182.5 cm, and so on.

Similarly, Cavendish's measurements of the weight of the Earth in 1797-1798 confirmed his estimate that density of the Earth was approximately 5.48 times the density of water, but his measurements also confirm our best modern estimates that the density was approximately 5.52 times the density of water. In such situations, it is natural to say that there are multiple confirmed contraries, even if one believes that some of the confirmed contraries are better confirmed than others.

Standard Bayesian confirmation theory is firmly committed to Carnap's side on this dispute. According to the standard Bayesian definition, a hypothesis H is confirmed by some evidence E if $P(H \mid E) > P(H)$. Obviously, it is possible for E to have favourable probabilistic relevance towards contraries. In addition to the above examples, consider the simple case in which we are making a random draw from a normal playing deck of cards:

Key

H_1: The card will be a Spade.

H_2: The card is a Club.

---

[261] Carnap (1962) p. 476-477.

E: The card is black.

Since the draw is random and the deck is normal, one can easily calculate the relevant parts of the probability distribution:

$$P(H_1) = 0.25$$

$$P(H_2) = 0.25$$

$$P(E) = 0.5$$

$$P(H_1 \mid H_2) = 0$$

$$P(H_2 \mid H_1) = 0$$

$$P(H_1 \mid E) = \frac{P(E \mid H_1)P(H_1)}{P(E)} = 0.5 > P(H_1)$$

$$P(H_2 \mid E) = \frac{P(E \mid H_2)P(H_2)}{P(E)} = 0.5 > P(H_2)$$

It follows that E confirms both H_1 and H_2 in standard Bayesian confirmation theory. In this particular case, Bayesian confirmation theory seems to match our ordinary intuitions: E supports both hypotheses. Of course, it would be disastrous if *all* contraries were *equally* confirmed by *all* evidence in a confirmation theory, but that is very different from *some* contraries being *merely* (or even *equally*) confirmed by *some* evidence.

In response to Carnap's arguments, Hempel suggests dropping the GCC, but notes that this would require further fundamental changes to his theory[262]. If an alternative confirmation theory CT satisfies the GCC and another of the features of his system, which he

[262] Hempel (1965) p. 49.

calls the “General Consequence Condition”:

General Consequence Condition: If a set of statements SH is confirmed by an evidence-statement E, then E also confirms every statement H that can be implied by some combination of the statements in SH.

- then, in CT, every evidence-statement that confirms a set of contrary hypotheses would also confirm every hypothesis in the language-schema L. Hempel does not provide the proof for this claim, but it is simple. As in Hempel’s theory, we assume classical logic.

Key

E: Any consistent statement.

SH: Any inconsistent set of hypotheses.

CH: The conjunction of the members of SH.

Premises

(1) CH implies every statement in L.

(2) The conjunction of SH can be deduced from a combination of the members of SH.

Proof

Claim: If E confirms every member of SH, then E confirms every statement in L.

From the General Consequence Condition and (2):

(3) If E confirms every member of SH, then E confirms CH.

From the General Consequence Condition and (1):

(4) If E confirms CH, then E confirms every statement in L.

By a Hypothetical Syllogism from (3) and (4):

(5) If E confirms every member of SH, then E confirms every statement in L.

Therefore, if Hempel modified his theory to CT, then he is committed to the trivialization result that E confirms every statement in L. To avoid this problem, Hempel could take the further step of abandoning the General Consequence Condition, but this step would involve rejecting one of the distinctive features of his system, which is that confirmation always 'flows down' the deductive consequence relation. It is not clear how such a confirmation theory would be 'Hempelian' in any meaningful sense.

However, despite Hempel's own concerns, he could retain something resembling his theory despite abandoning the GCC. It is possible to avoid the trivialization result above *and* retain the General Consistency Condition. If Hempel were to modify his definition to exclude contradictory evidence (a step that I made as a simplification of his definition in Subsection 2.2) then he could avoid the inference of sub-premise (4) in my proof above. Consequently, the GCC is neither an essential nor an attractive feature of Hempel's underlying theory of confirmation. Ironically, Hempel's theory meets his desideratum of the GCC, but this is a weakness rather than a strength of his system. The NRI is not fatal to Hempel's system, but his system is objectionable on other grounds. Hempel's definition of confirmation is too narrow, which is a point that he later grants to Carnap[263].

If Hempel's system does not offer an adequate option for a formalist, then she might consider Bayesianism. However, the standard Bayesian answers to the NRI are not available to formalists, as I have defined 'formalist'. The most common response is to discriminate against 'grue' in the prior distribution[264]. For example, a Bayesian can obtain unequal antecedent probabilities by giving lower priors to 'grue' or obtain unequal degrees of confirmation (at least on standard definitions of 'degree of confirmation') by giving a higher likelihood for the evidence to green hypotheses rather than grue hypotheses. Such options are unavailable to formalists, because they involve discriminating against 'grue', so I shall not explore them.

[263] Hempel (1965) p. 50.

[264] Howson and Urbach (1993) p. 163.

SECTION 3: RELIABILITY OF THE EVIDENCE

In this section, I shall propose a strategy that a formalist can use to answer the NRI. My idea is essentially that formalists can avoid the various incarnations of the NRI by appealing to an aspect of evidential support that is distinct from confirmation: the reliability of evidence. Concomitantly, even supposing that the formalist must say that Goodmanian and non-Goodmanian hypotheses are equally confirmed or probable (when the hypotheses are clearly not equally supported) she can appeal to this additional aspect of evidential support to answer the Riddle.

I shall begin by arguing that the Qualitative NRI is not inherently problematic for formalists. My next step will be to provide a general formalist response to the Quantitative NRI, using the concept of the reliability of evidence. I shall also consider some possible objections to my basic strategy. Having answered these objections, I shall apply my strategy to the underlying worry in the NRI, which is the possibility of a global underdetermination of equiprobable hypotheses. Finally, I shall generalise my answer beyond Goodman's version of 'grue', which will also illustrate that my answer does not depend on the particular details of greenness, grueness, or emeralds.

3.1 The Qualitative NRI

In the Qualitative NRI, the challenge for a purely formal confirmation theory is that one might know:

E: All observed emeralds are green and observed prior to *t*.

- and this confirms both -

H_1: All emeralds are green.

H_2: All emeralds are grue.

I shall assume that H_1 and H_2 are contraries. (Perhaps we are interpreting them with a logical form such that they have contrary counterfactual consequences.) In addition, I shall assume that E confirms both H_1 and H_2 in a formalist confirmation theory. Even granting these assumptions, there is no genuine problem for the formalist, because the fact that H_1 and H_2 are confirmed does not imply that they are *equally* confirmed. An increase in probability for two contrary hypotheses is not a Riddle: this is a normal part of science, as I argued in Subsection 2.6.2.

I stress that this scenario is based on a number of assumptions that need not always hold. A formalist can argue that we must always consider our background knowledge when assessing confirmation. One way to incorporate this information into confirmation theory is to conjoin our relevant background knowledge K with the evidence E. It might be the case that (E ^ K) does not confirm the grue hypothesis, even though the conjunction of E and some other hypothetical background information does confirm the grue hypothesis. For example, K might contain the knowledge that emeralds are uniform in their colours, so that if one emerald is green, then all emeralds are green. Suppose that K also contains the knowledge that there are emeralds that are not observed prior to *t*. In that case, (E ^ K) falsifies the grue hypothesis H_2. Of course, this is not a good response to Goodman's version of the NRI (he could simply raise the question of how we know that emeralds are uniform in their colours) but it does prove that learning E need not always confirm both H_1 and H_2 in a formalist theory. Therefore, the mere possibility of confirming contraries like H_1 and H_2 is not inherently problematic. I shall now consider the other versions of the Riddle.

3.2 The Quantitative NRI

In the Quantitative NRI, the alleged problem for the formalist is that the evidence that Goodman considers would (in a purely formal confirmation theory) confirm the grue hypothesis just as well as the green hypothesis. This equivalence would be contrary to most confirmation theorists' intuitions.

The absence of a received measure of degrees of confirmation complicates a discussion of this issue. There is not even a consensus within specific confirmation theories such as Bayesianism[265]. However, my answer to the Riddle will not depend on any particular definition of degrees of confirmation, because I shall simply assume that the degrees of confirmation are equal for both hypotheses. This enables me to obviate the voluminous debate and simply to assume that a formalist is confronted by an apparent underdetermination.

Assuming that the degrees of confirmation are equal, there is still a relevant asymmetry. Suppose that we learn:

[265] Fitelson (2001) p. 124.

E: All observed emeralds are green and observed prior to *t*.

- and (relative to our background knowledge) this confirms both -

H_1: All emeralds are green.

H_2: All emeralds are grue.

- to an equal degree. In practice, we derive our knowledge that observed emeralds are grue from (a) our knowledge that they are green and (b) that they have been observed prior to *t*. Thus, the knowledge that the emeralds are green depends only on the veracity of our colour-perception, whereas the knowledge that they are grue depends on both (a) the veracity of our colour-perception and (b) the veracity of our measurement of time. This creates an asymmetry of reliability that breaks the underdetermination of H_1 and H_2 given E and our background knowledge.

By 'reliability', I mean *apparent* reliability, rather than objective reliability. Our perception might be entirely unreliable, in some objective sense, for detecting either green or grue. However, if our beliefs about our sensory apparatus are correct, then we can observe that an object is green by observing it under ordinary lighting conditions, whereas we would need the additional information that it was observed prior to *t* in order to determine that it is grue. Thus, there is a (possibly minute) difference in the reliability: we can reasonably be more confident that the emeralds are green than that they are grue, even though we are obviously confident in both evidence statements.

This difference intuitively makes a difference: we would like to have hypotheses that are plausible given the evidence *and* to have reliable evidence. In the NRI, a formalist who incorporates the reliability of the evidence into her model of evidential support can use this asymmetry to say that the green hypothesis H_1 is better supported by its evidence, even though the degrees of confirmation are equal, because H_1's evidence is more reliable than H_2's evidence. In other words, when the degrees of confirmation are equal, the asymmetry in the reliability of the evidence statements can function as a tiebreaker, so that the underdetermination between the two hypotheses is resolved. Even assuming that they are equally *confirmed*, they are not equally *supported*. By this reasoning, the formalist can avoid the Quantitative NRI.

I am not proposing an empirical explanation of why we regard H_1 as better supported by our actual evidence than H_2. Relative to our actual background knowledge, most confirmation theorists agree that the degrees of confirmation are unequal. (Strictly speaking, the green hypothesis is necessarily true, because emeralds are green by definition.) Obviously, if we consider a hypothesis like 'Most grass is green', then we do not have to appeal to the greater reliability of 'Most observed grass is green' in comparison to 'Most observed grass is grue'. For instance, we can refer to our background knowledge that types of plants, such as grasses, tend to be mostly uniform in their greenness.

Instead of explaining why we do not believe hypotheses like 'Most grass is grue', my argument is that the formalist does not *have* to refer to such background knowledge to resolve Goodman's Riddle. She can use relevant background knowledge if it is available, but she can also point to the asymmetry in the reliabilities of the evidence, and thereby reason that 'Most grass is green' would be better supported by 'Most observed grass is green' in comparison to

'Most grass is grue', *even* assuming equal degrees of confirmation.

In the case of emeralds and grue, the asymmetry of reliability is very minute and perhaps not immediately clear, but the difference in reliability can be greater. Imagine that you are working down an extremely deep mineshaft. All your devices for measuring time have stopped working. You have been counting the days, so that you are quite confident that the month is December 2999 AD. However, you are not sure of the exact date: you think it might be 10/12/2999, 22/12/2999, or even 31/12/2999. Suppose that we define 'grue' in the following way:

D_7: *x* is 'grue' ↔ (1) *x* is first observed prior to 3000 AD and green or (2) *x* is first observed from 3,000 AD onwards and *x* is blue.

In this imaginary world, emeralds only exist deep below the Earth's surface and they have never been observed before. You become the first person to observe emeralds. You observe a large number and you see that they are green, but you also note that your observations are prior to 3000 AD and the emeralds are green, so you infer that they are also grue. It is clearly rational to be more confident that the observed emeralds are green rather than grue (you could have easily made a mistake in counting the days while underground) and based on this asymmetry, the formalist can say that your total evidence better confirms 'All emeralds are green' rather than 'All emeralds are grue'. The distinction between observing emeralds in 2017 AD and 2999 AD is a matter of degrees; in either version of grue, there is still an asymmetry of reliability between (a) the evidence that the emeralds are green and (b) the evidence that they are grue.

Another way to elicit this intuition is by considering the following alternative definition of ‘grue’:

D_8: *x* is grue ↔ *x* is green and first observed in a virgin forest or *x* is blue and not first observed in a virgin forest.

When Alexander von Humboldt was exploring the Amazon rainforest, he encountered many new species of plants. Suppose that, when observing an apparently virgin area, he encountered a species of plant *S*, whose leaves were almost always green. If von Humboldt was using ‘grue’ and ‘green’ in his language, then he could infer both that a particular leaf of *S* is green and that it is grue. However, his evidence that the leaves of *S* are almost always green is more reliable than his evidence that they are almost always grue, because he knew that he might be mistaken in his belief that this part of the Amazon was virginal[266]. The formalist can point to an asymmetry in the reliability of von Humboldt’s evidence in this scenario to conclude that (assuming that the hypotheses were otherwise underdetermined) the hypothesis ‘Most leaves of *S* are green’ was better supported by von Humboldt’s evidence than ‘Most leaves of *S* are grue’.

I shall stress that my answer does not involve any appeal to simplicity[267]. I am not claiming that the green hypothesis or the green evidence is simpler than the grue hypothesis or the grue evidence. I am also neither affirming nor denying that any such simplicity would

[266] Historically, von Humboldt greatly underestimated the impact of humans on the Amazon region and he was incorrect about the virginity of much of the rainforest.

[267] See Slote (1967) and Friedman (1973) for some clear appeals to simplicity in the NRI literature.

be epistemically significant. Instead, in my answer to the NRI, I am only appealing to asymmetries in the reliability of the evidence.

There are several methods for modelling this asymmetry of reliability. Kyburg's model of levels of corpora is one method. Kyburg aims to develop a theory of evidence that allows for more fine-grained distinctions of reliability than a straightforward distinction between 'accepted evidence' and 'probable hypotheses'. He developed this framework in books like *The Logical Foundations of Statistical Inference* (1974), *Science and Reason* (1990), and *Uncertain Inference* (2001). Kyburg formalises scientific knowledge as a set of statements at different levels in a hierarchy. The upper parts of the hierarchy are the most certain statements in the model, while admission to the lower levels is dependent on having a sufficiently high probability given the upper levels.

Suppose that 'All observed emeralds are green' is at a high level in the hierarchy, whereas the conjunction of this statement with 'All observed emeralds are observed prior to *t*' is only at lower levels. This formal asymmetry in the position of the statements in the hierarchy corresponds to the difference in reliability between 'All observed emeralds are green' and 'All observed emeralds are grue'. Kyburg's approach to modelling scientific knowledge is more complex than the standard Bayesian idealizations, but it offers one means of modelling science in which variations in the reliability of the evidence can be modelled.

Kyburg's approach is not the only possibility for a formalist. There is also Jeffrey Conditionalization, which I discussed in Chapter 2 Subsection 2.5. I shall not advocate any particular approach. My aim has simply been to argue that, by modelling the reliability of evidence and including it within their theory of evidential support, the formalist can avoid the NRI.

3.3 Objections

Before discussing the most fundamental worry that the NRI presents to a purely formal confirmation theory, I shall consider some objections to my answer. The NRI has been a severe problem for formalist theories of confirmation for 70 years, and so it is *prima facie* unlikely that any particular answer will be satisfactory. Since it is probable that my answer is somehow defective, each of the objections I shall consider is worth taking seriously.

Firstly, one might question whether a purely formal theory of confirmation can involve the reliability of evidence. After all, reliability is presumably not a purely formal matter. However, a confirmation theory is a theory of evidential relations, rather than evidence in general. Reliability can function as an exogenous factor in the theory, just as the evidence is an exogenous factor. Different formalists might adopt different theories of reliability, just as they might adopt different theories of observation, measurement, testimony, *a priori* knowledge, and so forth. The formalist's appeal to the reliability of evidence is no more objectionable than their appeal to the claim that the evidence-statements are known, and no-one considers this latter appeal a problem for formalism *per se*.

Secondly, one might think that I have simply changed the question: I am introducing the issue of the *uncertainty* and *corrigibility* of evidence to address Goodman's problem, whereas the NRI and typically confirmation theorists make the idealization that the evidence is certain and incorrigible. However, this idealization is not essential to the study of confirmation. Jeffrey (1992) offers a sophisticated form of Bayesianism in which evidence is not modelled as certain and incorrigible. Even Carnap, who modelled the evidence as certain, regarded this idealization as a formal artefact of standard confirmation theories, rather than a basic commitment of the discipline[268]. Similarly, Hempel models evidence as certain, but he also notes that, in practice, the acceptability of scientific hypotheses depends on both (1) their confirmation by the evidence and (2) the reliability of that evidence[269]. There is nothing fundamental in the classic confirmation theories that requires the idealization of certain evidence when analysing paradoxes like the NRI. Even if all historical confirmation theories involved this idealization, then my answer would be simply an instance of a normal strategy in formal modelling: when one encounters a problem and one can avoid the problem by relaxing an idealization, then one might relax the assumption.

Furthermore, the assumption that our evidence is incorrigible is not attractive for its own sake, because it is perfectly common scientific practice to dismiss previously accepted data. Critics of a scientific theory can legitimately pursue either presenting problematic data *or* critically examining the existing supporting data as their critical strategy. The assumption of incorrigible evidence in confirmation theory has (at best) pragmatic justification. If this idealization proves to be problematic when considering a paradox like the NRI, then it is reasonable to relax the assumption, so that the model becomes both more realistic and less problematic.

A third objection is that we might learn ‘All observed emeralds are green’ and ‘All

[268] Carnap (1968). p. 146. Hempel and Oppenheim (1945) p. 114-115 regard distinguishing between different degrees of reliability as an important step in the development of a confirmation theory, so my step is also consistent with Hempel’s broader project for formalism.

[269] Hempel (1945a). p. 25.

observed emeralds are grue' by independent and equally reliable means, so that we did not derive the latter using the former. In such circumstances, my response to the NRI will not apply. If we assume further that there is no background knowledge available that can break the symmetry, then both the green and the grue hypotheses would be underdetermined. However, I am not claiming that my answer can (or should) resolve all underdetermination problems involving 'grue'. Formalists can comfortably believe that evidence *sometimes* underdetermines hypotheses, especially when we assume that our evidence is far more austere than in reality. In contrast, the NRI is a problem of underdetermination in an indefinitely large number of scenarios in which it is strongly intuitive that our evidence is unequivocal. My answer to the Riddle is unscathed by this objection.

One might also object that non-Goodmanian hypotheses are still intuitively better confirmed than Goodmanian hypotheses *even when* there is no asymmetry in the reliability of the evidence and nothing in our background knowledge to resolve the underdetermination. However, it is doubtful that we have any intuitions about such alien scenarios. It is reasonable to put the burden of proof on a philosopher who claims to have a strong feeling about the evidential relations of a sample report of emeralds towards 'All emeralds are grue' in the absence of any background knowledge and when they are no more certain that the observed emeralds are green than that they are grue.

Fourthly, my answer might seem *ad hoc*, because I am introducing the importance of the reliability of the evidence to avoid a particular problem for formalist theories. Even if this was true, it is not a powerful objection: the NRI allegedly proves that a purely formal theory is untenable, rather than the mere claim that formalism must be somewhat *ad hoc*. However, philosophers have sought to model the reliability of evidence for reasons that are independent of the NRI[270]. Furthermore, the importance of obtaining reliable evidence is a basic point in applied methodology; incorporating it into formal models of scientific reasoning can be seen as doing justice to its importance in scientific practice, with the additional benefit (from a formalist's perspective) of answering the NRI.

Fifthly, I might appear to be discriminating in some way against 'grue', because I am saying that our evidence that emeralds are grue is less reliable than our evidence that they are green. There is a discrimination here, but it is not an internal discrimination in the confirmation theories I am recommending. The reliability of evidence is an external factor, just as the content of the evidence is an external factor.

This absence of discrimination can be seen from the fact that it is consistent with my account that some grue evidence is more reliable than some green evidence. For example, imagine that scientists observed a new type of star in distant galaxies using an instrument that could only detect if the star was grue or not grue. Suppose they discover that all stars of this type are grue. Assume that they have no ancillary reason in their background knowledge to favour 'green' over 'grue'. My response to the NRI means that their evidence favours 'Most stars of this type are grue' more than 'Most stars of this type are green', because the evidence for the former hypothesis is more reliable.

Additionally, on my answer, it is possible that 'All emeralds are grue' could be better supported than 'All emeralds are green'. Imagine a universe in which emeralds were no different to our own, but our eyes could directly observe whether things were grue or bleen, but not green or blue. (The predicate 'bleen' is defined as 'Blue and observed prior to *t* or

[270] See Roush (2005), as well as Fennell and Cartwright (2010).

green and not observed prior to *t*.') In such a universe, we could look at an emerald and see that it is green *if* we are observing it prior to *t*. To know that it is green, we shall have to use some additional instruments for measuring time. My answer would require that *if* 'All emeralds are green' and 'All emeralds are grue' were otherwise equally supported given our evidence, then we should opt for the grue hypothesis. We might be mistaken, but the possibility of error is not a sound objection to a confirmation theory. The important point is that hypotheses like 'All emeralds are grue' can be favoured over hypotheses like 'All emeralds are green'. There is no fundamental discrimination in favour of 'green' over 'grue' in my answer to the Riddle.

Sixthly, a critic might claim that my answer depends on how we describe the observations. They might argue that, if we used a language with 'grue' and 'bleen' rather than 'green' and 'blue', then we could have described our observations of green emeralds as "This emerald is grue, that emerald is grue…" Consequently, my answer would be relative to a particular choice of language, and thus it would not be a general formalist answer to the NRI. However, there is a false premise in this argument: simply observing that an emerald *e* is green does not prove that '*e* is grue'. If we spoke such a language, but our senses were unchanged, then we could observe that:

O_1: '*e* is grue and observed prior to *t* or bleen and not observed prior to *t*.'

This statement O_1 is simply the way of saying, in such a language, that *e* is green. It is logically distinct from:

O_2: '*e* is grue.'

- because O_1 will be true and O_2 false if the emerald is green and first observed after *t*. My answer to the Riddle would still be correct if we spoke such a language, because we would still have to infer statements like O_2 from O_1 *and* our background knowledge that the emerald was first observed prior to *t*. In general, a difference in description does not create a difference in reliability and I am using a difference in the reliability of evidence in my answer, as opposed to a difference in the language of evidence. Since my answer is independent of any linguistic properties of 'green' and 'blue' over 'grue' or 'bleen', but instead involves an epistemic asymmetry in the case of emeralds, it follows that my answer is not language-relative.

Seventhly, one might worry that my answer only establishes the comparative claim that the green hypothesis is better supported than the grue hypothesis, when it should really establish the claim that it is *much* better supported. From the outset, it must be noted that this is not any version of the classic NRI: proving such a claim from a formalist perspective would be contrary to what Goodman asserts, but he only argues that purely formal theories cannot identify an asymmetry of support between the two hypotheses, rather than that they cannot appropriately assess the strength of this asymmetry. More importantly, when our actual background knowledge is available, a formalist can use it to establish asymmetries of strength of confirmation: for example, we might empirically know that 'green' refers to a natural kind and 'grue' does not, and that hypotheses that refer only to natural kinds are much more likely to be true given inductive evidence than those that make no such reference. The precise details will involve attention to the relevant scientific and historical background knowledge in the context under analysis, but the point is that the formalist *can* do justice to the claim that the green hypothesis is much better supported given the evidence and our background knowledge.

This point is important, because a simple appeal to background knowledge on the part of a formalist is likely to leave critics unsatisfied: if a critic assumes that the relevant background knowledge has been obtained by inductive reasoning, then a mere appeal to this information simply pushes Goodman's Riddle back to the underdetermination of the *background knowledge* in contrast to gruesome alternatives. For example, if we learned that hypotheses non-gruesome properties tend to produce better science than their gruesome alternatives, but this information was inferred inductively from the past experience of scientists, then a critic of formalism could simply adapt the Riddle for this inductive inference. An appeal to the reliability of the evidence offers a formalist a means of breaking the symmetry of support without recourse to such background knowledge.

This consideration raises the issue of the origin of the beliefs about the reliability of evidence. Are they not part of our inductively-acquired background knowledge? Certainly, experience can lead us to modify beliefs about the reliability of some type of evidence. For instance, almost everyone today rejects the notion that clairvoyance is a reliable source of evidence, but this scepticism has not always been so widespread. Since formalism is compatible with a wide range of epistemologies, there is no single response that a formalist must make and defend to this query, but a fallibilist foundationalist (i.e. one who holds that our knowledge is ultimately justified by a bedrock of basic beliefs, but these beliefs are open to revision) could hold that some beliefs about the reliability of different sources of evidence are part of the foundations of our knowledge: I might believe that my faculties of perception or logical reasoning are reliable but fallible sources of knowledge under some circumstances; that my hands are a better means of determining solidity than my eyes; and so on. However, there is nothing essentially foundationalist about my answer: that we have knowledge of the reliability of *some* forms of evidence, prior to acquiring inductive knowledge, is consistent with multiple epistemologies – even those that dispense with justification. Thus, even if the formalist is driven back to very exiguous background knowledge that cannot break the asymmetry between the gruesome and non-gruesome hypotheses, they might still be able to appeal to the reliability of evidence, and thus avoids the danger of a vicious regress in their appeals to background knowledge.

One might reply that, intuitively, the two hypotheses are clearly very asymmetrically supported even when our relevant background knowledge is unavailable. However, appeals to intuition about such scenarios are very suspect. Indeed, the case of green and grue emeralds is itself a reason to be sceptical about such arguments: in reality, 'All emeralds are green' is true by definition (if 'emerald' has its usual gemmological acceptation) whereas 'All emeralds are grue' can only be contingently true. A critic could further argue that, even for claims like 'Most grass is green' and 'Most grass is grue', the former is intuitively much better supported when our relevant background knowledge (or any other background knowledge that might provide a strong asymmetry) is unavailable. Here, I shall simply demand more than an appeal to intuition: there needs to be an argument for why such a classificatory claim must be honoured by a formalist, given that the relevant comparative claim that 'Most grass is green' is better confirmed than 'Most grass is grue' can be provided by a formalist theory.

Eighthly, some philosophers might claim that the NRI is unsolved unless one can prove that the grue hypothesis is not confirmed at all. However, intuitions are divided on this question: the standard Bayesian responses of providing a higher prior probability to the green hypothesis or a higher likelihood for the evidence given the green hypothesis do not satisfy this intuition either. Few philosophers, if any, regard this feature of the standard Bayesian answer to the NRI as a fatal flaw in Bayesianism. Yet, if it is not a severe problem for Bayesianism, then it does not seem to be a severe problem for formalism either, *provided* that the formalist can preserve the intuition that the green hypothesis is better confirmed than the grue hypothesis, so that the evidence does not underdetermine them. That is precisely what my answer enables the formalist to do.

Finally, one might doubt that my answer is a formalist answer, because I am suggesting the use of probabilities as well as the terms from first-order logic that Hempel used. There is certainly an important sense of 'formalist' in which Hempel's theory is formalist and which is incompatible with my proposal. However, Carnap also used probabilities and he is just as paradigmatically a 'formalist' as Hempel. (Indeed, Goodman targets *both* Carnap and Hempel, alongside Oppenheim and Helmer, in his first presentation of the NRI.) Perhaps 'syntacticist' would be a useful label for Hempel's approach, 'probabilist' for Carnap's approach, and 'formalist' for their shared ambition of a subject-matter independent theory of confirmation. For my purposes, it is unnecessary to insist on terms: the important point is that fundamental discriminations among predicates are superfluous in confirmation theory. If one wants to reserve 'formalist' for some other position, then I shall simply need another appellation for what I mean by 'formalist'.

3.4 Equal Probability

At the heart of the NRI is the worry that our total evidence might underdetermine theory-choice in an indefinite number of absurd cases, because it is possible to formulate arbitrarily many predicates like ‘grue’ to create rivals to any given hypothesis. A formalist confirmation theorist forswears any discrimination among the rivals in the prior distribution. However, if there is no prior discrimination against any of the rival hypotheses, such that the rivals have both (1) equal prior probabilities and (2) equal posterior probabilities given the total evidence, then the formalist might confront a very general underdetermination problem. It is this general underdetermination problem that is the greatest single challenge to the formalist, and which seems to be the central reason why the NRI is regarded such an intolerable defect in formalist confirmation theories.

However, even granting the assumption of equiprobability for all the rivals, there can still be an asymmetry in the reliability of the evidence. Provided that a formalist’s definition of support incorporates the reliability of the evidence as a potential tiebreaker between rival hypotheses, then such a formalist can say that the Goodmanian hypotheses like ‘All emeralds are grue’ are not as well-supported as their non-Goodmanian rivals. A typical formalist definition of comparative evidential support might be:

Comparative Evidential Support (1): H_1 is better supported by its total evidence E_1 than H_2 is supported by its total evidence E_2 if and only if $P(H_1 \mid E_1) > P(H_2 \mid E_2)$.

- whereas a formalist following my proposal could make a simple and intuitive modification to this definition -

Comparative Evidential Support (2): H_1 is better supported by its total evidence E_1 than H_2 is supported by its total evidence E_2 if and only if either:

(I) $P(H_1 \mid E_1) > P(H_2 \mid E_2)$

\- or -

(II) $P(H_1 \mid E_1) = P(H_2 \mid E_2)$ but E_1 is more reliable than E_2.

(This second definition leaves open how one models the difference in reliability.)

However, my answer to the Riddle does not depend on the exact details of this second definition, provided that a clause that is analogous to Clause (II) is in the definition or can be derived from the definition. Firstly, a non-probabilistic definition of evidential support can be consistent with my proposal: one might use a Popperian notion of corroboration, a hypothetico-deductive notion of confirmation, or some other approach to evidential support. Secondly, my answer is consistent with a more general incorporation of the reliability of evidence into the analysis of evidential support. One might think that a large difference in the reliability of evidence should be able to offset a small difference in the conditional probability. Provided that the reliability of evidence can perform the function of a tiebreaker in cases like the NRI, such theories are consistent with my answer.

One might think that this modification is unnecessary. It might seem that, if one defines ‘reliability’ in terms of marginal probability in a Bayesian probability distribution, then the probability calculus will guarantee that the asymmetry in the reliabilities of the evidence will require an asymmetry in the probabilities of the hypotheses. Thus, clause (II) would be redundant.

Formally, the objection is the claim that (1) and (2) imply (3):

(1) $P(H1 \mid E1) \ = \ P(H2 \mid E2)$

(2) $P(E1) > \ P(E2)$

(3) $P(H1) \ > \ P(H2)$

While there is a plausible line of thinking that *almost* leads from (1) and (2) to (3), they are actually logically independent of (3), since they are also consistent with $P(H_1) = P(H_2)$. Therefore, it is possible that (1) and (2) are true, but nonetheless the prior probabilities underdetermine a choice between the hypotheses.

To prove this, I shall first examine the line of thought that suggests that (1) and (2) imply (3). I shall prove that an additional assumption is required. In my second proof, I shall demonstrate that when the extra assumption is removed and an alternative postulate is introduced, then it follows that $P(H_1) = P(H_2)$. I shall further prove that (1) and (2) are even consistent with $P(H_2) > P(H_1)$, such that H_2 has a greater probability in the marginal distribution. These first three proofs will be algebraic, but to make the point clearer I shall finish with a proof that $P(H_1) = P(H_2)$ in a joint distribution using numerical values.

Proof 1

Claim: (1) and (2) do not imply (3), but imply (3) with an additional assumption regarding the relative values of $P(E_1 \mid H_1)$ and $P(E_2 \mid H_2)$.

Premise (1) implies that $P(E_1)$ and $P(E_2)$ are both greater than zero, because the conditional probabilities must be defined in order to be equal. (I am using the simple definition of conditional probability; there would be nothing essentially different in my argument if I were using a more general measure-theoretic definition that did not have this property.) Using Bayes's theorem:

(4) $P(H1 \mid E1) = \frac{P(E1 \wedge H1)}{P(E1)} = \frac{P(E1 \mid H1)P(H1)}{P(E1)}$

(5) $P(H2 \mid E2) = \frac{P(E2 \wedge H2)}{P(E2)} = \frac{P(E2 \mid H2)P(H2)}{P(E2)}$

By substitution from (1), (4), and (5):

(6) $\frac{P(E1 \mid H1)P(H1)}{P(E1)} = \frac{P(E2 \mid H2)P(H2)}{P(E2)}$

The equation above puts no constraints on the values of the terms in the numerators. Furthermore, (2) and (6) only require that the denominators are unequal and they have a non-zero value. Otherwise, the values of the denominators are unconstrained. Therefore, $P(H_1)$ can have any value r such that $0 \leq r \leq 1$. The same is true, *mutatis mutandis*, for $P(H_2)$. Consequently, the prior probabilities are unconstrained, such that (1) and (2) are logically independent of (3).

The line of thought that leads from (1) and (2) to (3) seems to involve the implicit assumption that the likelihoods in the numerators of (6) are equal:

(7) $P(E1 \mid H1) = P(E2 \mid H2)$

This equation does not follow from (1) or (2), because they put no constraints on either term, though it is consistent with those assumptions. Once (7) has been assumed, the derivation of (3) is possible. The equality in (6) can be rewritten by removing the likelihoods from the numerators:

(8) $P(E1 \mid H1)\frac{P(H1)}{P(E1)} = P(E2 \mid H2)\frac{P(H2)}{P(E2)}$

From (7) and (8), it follows that the likelihoods can be removed from both sides of the equality, such that:

(9) $\frac{P(H1)}{P(E1)} = \frac{P(H2)}{P(E2)}$

Rewriting (9) to position the priors and the expectedness terms together provides the useful equation (11):

(10) $\frac{P(H1)P(E2)}{P(E1)} = P(H2)$

(11) $\frac{P(E2)}{P(E1)} = \frac{P(H2)}{P(H1)}$

Thus, the probabilities of the hypotheses and the probabilities of the evidence vary in proportion, so if $P(E_1) > P(E_2)$ and the likelihoods are equal, it must be the case that $P(H_1) > P(H_2)$ to a proportionate degree. Hence, (1), (2), and (7) imply that (3) is true. It is this sort of reasoning that seems to suggest that (1) and (2) imply (3), so that relatively improbable evidence for the grue hypothesis seems to imply a relatively lower prior probability for that hypothesis given the underdetermination in the probability distribution.

Proof 2

Claim: There are probability distributions in which (1) and (2) are true, but $P(H_1) = P(H_2)$, such that (3) is false.

As previously noted, (1) and (2) place no constraints on the likelihoods, so it is possible that the likelihoods vary with $P(E_1)$ and $P(E_2)$ such that:

(12) $\frac{P(E1 \mid H1)}{P(E2 \mid H2)} = \frac{P(E1)}{P(E2)}$

Since the likelihoods are unconstrained, this is consistent with the assumption (2) that

$P(E_1) > P(E_2)$. Dividing both sides of (12) by $\frac{P(E1)}{P(E2)}$ provides:

(13) $\frac{P(E1 \mid H1)P(E2)}{P(E2 \mid H2)P(E1)} = 1$

Reiterating and reformulating (6):

(6) $\frac{P(E1 \mid H1)P(H1)}{P(E1)} = \frac{P(E2 \mid H2)P(H2)}{P(E2)}$

(14) $\frac{P(E1 \mid H1)P(H1)}{P(E2 \mid H2)P(E1)} = \frac{P(H2)}{P(E2)}$

(15) $\frac{P(E1 \mid H1)}{P(E2 \mid H2)P(E1)} = \frac{P(H2)}{P(E2)P(H1)}$

(16) $\frac{P(E1 \mid H1)P(E2)}{P(E2 \mid H2)P(E1)} = \frac{P(H2)}{P(H1)}$

From (13) and (16):

(17) $1 = \frac{P(H2)}{P(H1)}$

Finally, from (17), it follows that H_1 and H_2 must be equiprobable, since if both sides of the equation are multiplied by $P(H_1)$, then we obtain:

(18) $P(H1) = P(H2)$

Therefore, (1) and (2) are consistent with the falsity of (3). Furthermore, (1) and (2) are possible in a probability distribution with equal conditional probabilities for the hypotheses given their evidence and greater probability for one hypothesis, so that the hypotheses can be underdetermined in spite of an asymmetry in the reliability of evidence.

Interestingly, since a difference in the proportions of the likelihoods can offset a difference in $P(E_1)$ and $P(E_2)$, it is even possible that $P(H_2)$ exceeds $P(H_1)$ given the assumptions (1) and (2), as I shall now prove.

Proof 3

Claim: There are possible distributions in which (1) and (2) are true, but $P(H_2) > P(H_1)$.

Since the likelihoods are unconstrained by (1) and (2), it is consistent with those assumptions that (12) is replaced by the following inequality:

(19) $\frac{P(E1 \mid H1)}{P(E2 \mid H2)} > \frac{P(E1)}{P(E2)}$

Dividing both sides by $\frac{P(E1)}{P(E2)}$:

(20) $\frac{P(E1 \mid H1)P(E2)}{P(E2 \mid H2)P(E1)} > 1$

Recalling (16), which was derived from (1) and (2):

(16) $\frac{P(E1 \mid H1)P(E2)}{P(E2 \mid H2)P(E1)} = \frac{P(H2)}{P(H1)}$

From (16) and (20):

(21) $\frac{P(H2)}{P(H1)} > 1$

(22) $P(H2) > P(H1)$

Thus, (1) and (2) are consistent with $P(H_2) > P(H_1)$ as well as $P(H_1) > P(H_2)$ and $P(H_1) = P(H_2)$.

Proofs 1 to 3 are purely algebraic. It might help if I give a specific numerical example of a distribution in which, despite (1) and (2), H_1 and H_2 are equiprobable.

Proof 4

Claim: In the particular probability distribution that I discuss below, $P(H_1) = P(H_2)$ despite the fact that (1) and (2) are also true in this distribution.

It is consistent with (1), (2), and Bayes's Theorem that:

(23) $P(H1 \mid E1) = P(H2 \mid E2) = 0.3$

(2) implies that $P(E_1) > P(E_2)$ and (23) implies that $P(E_1)$ and $P(E_2)$ are both greater than zero, but they do not provide any further constraints on their values. It is consistent with (1), (2), and (23) that:

(24) $P(E1) = 0.6$

(25) $P(E2) = 0.5$

Assume that (12) is true in this distribution, so that:

(12) $\frac{P(E1 \mid H1)}{P(E2 \mid H2)} = \frac{P(E1)}{P(E2)}$

(24) and (25) can be combined with (12) to calculate the quotient of the likelihoods:

(26) $\frac{P(E1 \mid H1)}{P(E2 \mid H2)} = \frac{P(E1)}{P(E2)} = \frac{0.6}{0.5} = 1.2$

The likelihoods are still otherwise unconstrained. Additionally, their quotient does not imply their individual values. It is consistent with (26), as well as (1) and (2), that:

(27) $\frac{P(E1 \mid H1)}{P(E2 \mid H2)} = \frac{0.9}{0.75} = \frac{P(E1)}{P(E2)} = \frac{0.6}{0.5} = 1.2$

It is now possible to derive $P(H_1)$ and $P(H_2)$. Substituting the values from (27) into (23):

(28) $\frac{0.9P(H1)}{0.6} = \frac{0.75P(H2)}{0.5} = 0.3$

I shall first derive $P(H_1)$ from (28):

(29) $\frac{0.9}{0.6}P(H1) = 0.3$

(30) $0.9P(H1) = 0.18$

(31) $P(H1) = 0.2$

Secondly, by the same reasoning, I shall derive $P(H_2)$ from (28):

(32) $\frac{0.75}{0.5}P(H2) = 0.3$

(33) $0.75P(H2) = 0.15$

(34) $P(H2) = 0.2$

From (31) and (34):

(35) $P(H1) = P(H2) = 0.2$

Proofs 2 and 4 imply that a difference in the reliability of two hypotheses' evidence (defined as a difference in marginal probabilities) is consistent with the underdetermination of those hypotheses within a probability distribution that conforms to the probability calculus. Such an underdetermination poses a sceptical problem, because (1) and (2) do not set any constraints on $P(H_1 \mid H_2)$ and $P(H_2 \mid H_1)$. Since the hypotheses' probabilistic relations are unconstrained by the initial assumptions, it is possible that these hypotheses are logically inconsistent. Consequently, the underdetermination of contraries with (1) equal conditional probabilities and (2) asymmetrically probable total evidence is possible in a coherent probability distribution.

However, a formalist can avoid this possibility once she incorporates the reliability of the evidence into her confirmation theory. Suppose that the green H_1 and grue hypotheses H_2 are contrary hypotheses, such that:

(a) $P(H_1) = P(H_2)$.

(b) The total evidence in favour of the green hypothesis is more reliable than the total evidence in favour of the grue hypothesis.

The formalist can use a definition of comparative support in which the reliabilities of the hypotheses' evidence can function as a tiebreaker. Thus, there will be no underdetermination. By considering the reliabilities of the evidence as well as the differences in conditional probability, the formalist can avoid this version of the NRI.

Furthermore, as noted earlier this development of formalist theories is desirable for independent reasons. Scientists ideally want both (i) that their hypotheses are strongly confirmed by their total evidence and (ii) that their total evidence is reliable. Just as we want our houses to be both strongly connected to their foundations *and* based on solid foundations, so we ideally want our scientific hypotheses to be both well-connected to the evidence *and* based on solid evidence. If one regards this simile as objectionably foundationalist, then I can make the same point using Otto Neurath's raft metaphor[271]. Just as we want the parts of the raft to be strongly connected *and* individually sturdy, so we ideally want our scientific hypotheses and evidence to be strongly confirmed *and* individually reliable.

Even in a scenario in which the grue and green hypotheses are equally probable given the total evidence, one can answer the NRI using a purely formal confirmation theory. For formalists, this strategy offers a strong incentive to consider the reliability of the evidence, as well as conditional probabilities, in their theories of evidential support.

3.5 Generalisation

[271] Neurath (1973) p. 199.

Throughout my discussion of the NRI, I have mostly followed Goodman's own choice of predicates for his Riddle: 'emerald', 'green', 'grue', and 'observed prior to *t*'. However, the NRI can be adapted using a wide variety of different predicates. Goodman could have used a spatial predicate such as 'observed within the known universe' rather than 'observed prior to *t*' and still generated the essence of his paradox. Furthermore, the use of two colour-predicates rather than one is inessential: Goodman could have used 'not-green', so that 'grue' means 'green and first observed before *t* or non-green and not first observed before *t*'. This definition would also be somewhat simpler, since we would not need to use the potentially controversial claim that we know *a priori* that it is impossible for an object to be monochromatically green and monochromatically blue.

The general recipe for a Goodmanian predicate is:

<u>Key</u>

F: Some uncontroversially genuine and unproblematic predicate.

A: Some auxiliary predicate that is known to be satisfied by all observed instances of F.

D_9: *x* is GP $\leftrightarrow$ ((Ax ^ Fx) v (¬Ax ^ ¬Fx))

For example, all observed emeralds weigh less than 400 kg. Plausibly, somewhere in the universe, there is an emerald that weighs more than 400 kg. Using the above formulas for constructing Goodmanian predicates, one could define 'GP_1' as:

D_{10}: *x* is $GP_1 \leftrightarrow$ (1) *x* weighs more than 400 kg in diameter and green or (2) *x* weighs more than 400 kg and *x* is not green.

Such variations do not pose a problem for my answer. In NRI scenarios, the predicate A is always some additional predicate we have separately inferred about each known *x*. Our knowledge that the instances in the Riddle satisfy A is combined with our knowledge that they satisfy F, in order to infer that they are GP. Thus, when some Goodmanian and non-Goodmanian hypotheses are equally probable given the evidence, we can typically appeal to the asymmetry of reliability. Similarly, we can typically appeal to the asymmetry in reliability of the evidence for equiprobable Goodmanian and non-Goodmanian hypotheses to avoid underdetermination versions of the NRI.

In the example of GP_1 it is *possible* that we have somehow incorrectly measure the weight of some emeralds (the largest observed emerald is approximately 380 kg) such that they are green but also over 400 kg. Of course, it is extremely likely that all observed emeralds are grue (in the D_{10} sense) but it is still slightly less likely than that they are green.

To give another example of a Goodmanian predicate that does not use a temporal predicate:

D_{11}: *x* is $GP_2 \leftrightarrow$ (1) *x* dissolves in water under laboratory conditions and *x* is within the known universe or (2) *x* does not dissolve in water under laboratory conditions and *x* is *not* within the known universe.

All of our observed samples of sodium are GP_2. Even if we assume a body of background knowledge that equally confirms:

H: All sodium dissolves in water under laboratory conditions.

H´: All sodium is GP_2.

- using my strategy to answering the NRI, we could still reasonably predict that if there is sodium outside of our region of the cosmos, then it would also dissolve in water under laboratory conditions, because the evidence for H is more reliable than the evidence for H´. The difference might be very slight, but it is still enough to rank H as the better-supported hypothesis.

I shall now give a scenario in which the ascription of the additional predicate A to our samples is more obviously corrigible than the above example. (This example is not intended to be strongly analogous with the sort of scenario that Goodman discusses: for instance, in Goodman's original NRI scenario, he is assuming that our total evidence solely consists of observation reports of green emeralds, whereas I am not assuming such austere total evidence in this scenario.) I shall define a 'geodesic' as:

D_{12}: *x* follows a 'geodesic' path ↔ (1) *x* moves along a Euclidean straight line and *x* is in a Euclidean space or (2) *x* does not move along a straight line and *x* is not in a Euclidean space.

Imagine that we are confident that all observed moving objects will travel along a Euclidean straight line in the absence of an external force. Thus, they conform to Newton's

First Law, as he presumably intended it. However, imagine that we believe that we exist in a Euclidean space, but we are less confident about this conjecture than we are that all objects conform to Newton's First Law. We consider the two hypotheses:

S: In the absence of an external force, all objects move along a Euclidean straight line.

G: In the absence of an external force, all objects move along geodesics.

Hypotheses S and G, when conjoined with the auxiliary hypothesis that an object *a* is travelling in a non-Euclidean space, make contrary predictions about the paths that *a* will take in the absence of an external force. Relative to our evidence in this hypothetical scenario, my answer to the NRI entails that the prediction that *a* will follow a Euclidean straight line in the absence of an external force is better supported than the prediction that *a* will not travel along such a path, because the evidence for S is more reliable.

One interesting aspect of this scenario is that my strategy could lead to a false prediction, according to modern physics. While the overall geometry of the *total* universe is still a subject of ongoing debate (a Euclidean "flat" universe is one of many possibilities) the *local* geometry of the universe is non-Euclidean, given Einstein's General Theory of Relativity. Thus, S is better supported than G by our actual *total* relevant evidence. However, this total evidence is far greater than the simple generalisation that objects would travel along Euclidean straight lines in the absence of external forces. Of course, this possibility is not surprising: philosophers of science have long agreed that even the best-supported scientific hypothesis can be false. As Joseph Butler noted, it is probability, not certainty, which must be "the very guide of life"[272].

Finally, one might wonder about the significance of this answer to the NRI for practicing scientists, as gruesome hypotheses are very far from their consideration. For my purposes, this answer to the NRI is important because it answers a philosophical objection to formalist approaches to confirmation theory and thus helps legitimate the adoption of such theories, but that response simply raises a further query: what use are formalist theories of evidence for practicing scientists? There are many answers to this question, but I shall focus on three points.

Firstly, in some cases, a formalist theory can answer some debates about evidence that can result from philosophical controversies about evidence, *without* requiring that the scientist adopt logically contingent *a priori* claims. For instance, a scientist might think that it is impossible that quantum mechanics could be false, given our evidence that there are many well-functioning technologies that were designed using quantum mechanics; another scientist might disagree. One might think that the issue ultimately is one of metaphysical presuppositions: perhaps the first scientist is presupposing the uniformity of nature, while the other is not. Some philosophers have argued that such questions about evidential import are fundamentally metaphysical: Arthur W. Burks developed a "Presuppositional Theory of Induction" on these grounds[273]. More broadly, one might think that the adoption of one set of prior probabilities that are favourable towards some predicates rather than another is a choice of metaphysics. Thus, the first scientist might claim that she is entitled to her metaphysics

[272] Butler (1736) p. iv.

[273] Burks (1953).

and therefore to believing that the technologies' success proves the truth of quantum mechanics. The formalist offers a different view: such a dispute is ultimately a formal dispute, and its resolution depends solely on the formal relations between the total evidence and quantum mechanics. A more systematic controversy along these lines is the debate in statistics over the use of Bayesian methods or classical methods. Such disputes frequently raise what are philosophical issues about evidence: what role do prior probabilities have to play in statistics? Should one reject a hypothesis that is logically consistent with the evidence? If one adopts a pluralistic and contextually-dependent attitude towards the choice of statistical methodology, then how does one choose which method to use and when? Kyburg took an interest in these questions; in fact, Evidential Probability was partly designed to help resolve such controversies[274]. Essentially, his answer is that Bayesian methods should be used in areas where we have rich statistical background knowledge that either provides precise knowledge of frequencies or enables the use of approximation techniques to generate such precision (so that Sharpening by Richness plays the central role in updating) *but* classical methods should be used when such rich statistical background knowledge is not available[275]. As a result, an Evidential Probabilist would tend to be favourable towards the use of Bayesian methods in many areas of sciences such as solid-state physics, where we can use probability distributions that are empirically-based, but be very suspicious of their use in areas like the social sciences, where precise probability distributions would require speculating far beyond the available evidence. Consequently, for practicing scientists who are averse to grounding scientific methodology on contingent *a priori* claims, formalism offers an alternative route to the resolution of such disputes.

---

[274] Kyburg (1974) Chapter VI.

[275] Kyburg and Teng (2001) p. 263-265.

Secondly, formalist theories can offer relatively neutral (in the sense of minimizing *a priori* presuppositions) answers to important methodological questions in science. The crucial element in having a *formalist* answer to such questions is that it reduces the role of presuppositions: formalists will tend to agree with the somewhat widespread view (held by more than a few scientists but one that is anathema to some academic philosophers) that the scientific method is about reasoning logically from the evidence; it does not depend leaps of faith or *a priori* intuitions about the structure of the universe. For many scientists, there is something special about science in contrast to other sources of beliefs about the world, and the perspective that I have briefly characterised is one way of cashing-out this ‘specialness’. Therefore, while the NRI does not seem to be directly relevant to the practice of scientists, it is indirectly relevant to the defence of a certain theory of the scientific method that guides and motivates many scientists.

CONCLUSION

I have argued that the NRI does not present a direct problem for Hempel’s theory. There are major problems with Hempel’s theory (such as its very limited scope of application) but these are logically independent of the NRI. Additionally, I have proposed and illustrated a very generally applicable strategy that formalists can use to answer the NRI, which is that formalists can incorporate the reliability of evidence as well as confirmation into their model of evidential support. I have not ruled out the possibility of finding a predicate that presents serious problems for formalism. It might be possible to formulate a predicate that (1) can be used to construct a severe NRI-type problem and that (2) the formalist cannot handle using my answer to the NRI (or via any other strategy) but this would not be Goodman’s Riddle. It is the NRI and its simple variations that are supposed to be fatal for formalism, and those are the Riddles that I have sought to answer. The burden of proof rests on the anti-formalist to present such a New New Riddle of Induction; the Old New Riddle of Induction is answerable by a formalist.

For my thesis *in toto*, there are two main lessons. Firstly, the principal objection to purely formal confirmation theories, like Evidential Probability, does not hit its mark. There might be other reasons to reject formalism, but the NRI is not such a reason. Secondly, when faced with underdetermination problems like the NRI, it can sometimes be advantageous to attend to other aspects of evidential support, such as the reliability of the evidence. I have not provided a detailed discussion of the alternative methods for formalising reliability of evidence. However, I have provided an additional reason for formalists to undertake the project of incorporating this aspect of scientific reasoning into their theories of evidential support.

CHAPTER 5: PROBABILITY AND BALANCE OF EVIDENCE

In the preceding chapters, I have applied Evidential Probability to several problems in confirmation theory and decision theory. In this chapter, I shall apply this approach to analysing the balance of evidence, which has been the traditional focus of confirmation theory. In particular, I shall argue that supporters of probabilistic analyses of the balance of evidence can address some of Norton's criticisms of such analyses of inductive reasoning by adopting Evidential Probability as their epistemic probability theory.

In Section 1, I narrow the scope of my discussion by drawing some important distinctions. In Section 2, I provide an overview of Kyburg's theory of inductive inference. In Section 3, I argue that this theory can respond to Norton's criticisms.

SECTION 1: AMPLIATIVE INFERENCE AND INDUCTIVE INFERENCE

Before discussing Kyburg's analysis, I shall clarify the *analysandum*. There is an important distinction between (i) ampliative inference and (ii) inductive inference. As I shall use these terms, 'ampliative inference' refers to any non-deductively valid inference. One can model a deductively valid inference as a deductively valid argument in which the inferred proposition is the conclusion, while the implicit and explicit premises of the inference are the premises of the argument. In the sense that I am using the term, a 'deductively valid argument' is one in which the premises are inconsistent with the negation of the conclusion. (There are many alternative ways of characterising deductive validity; in this chapter, I use a classical definition.) Aside from induction, there are many other types of ampliative inference, including abductive inference, analogical inference, and direct inferences[276]. I shall use 'inductive inferences' to mean inferences from observational evidence to logically contingent non-observational conclusions. For example, the inference from 'All observed metals conduct electricity' to 'All metals conduct electricity' is an inductive inference. Inductive arguments *per se* are invalid, although some inductive inferences can be modelled as deductive arguments, as I shall discuss later in this section. These definitions follow Peirce's definition of 'ampliative'[277] and Stove's definition of 'induction'[278]. They are stipulative rather than descriptive definitions, because some philosophers take these two terms to be synonymous; for instance, Carnap defines 'induction' as "nondeductive" inference[279]. As I am using the terms, neither category fully includes the other.

There is also an important distinction between inductive inference and scientific inference. As Larry Laudan notes, many philosophers confuse these two categories by assuming that the scientific method *in toto* could be justified by the justification of induction alone, whereas it is consistent to think that (a) induction is a *part* of the scientific method and (b) induction is not the *whole* of science or the *crux* of science[280]. Instead of endorsing any

[276] When the proportion cited in the direct inference is not 0% or 100%. When the proportion takes these values, a direct inference is a deductively valid categorical syllogism.

[277] Peirce (1932) 2.680.

[278] Stove (1982) p. 56.

[279] Carnap (1962) p. 580.

[280] Laudan (1981) p. 240-241.

precise thesis about the scope of induction in science, I shall simply assume that it is an interesting topic within confirmation theory.

A further useful distinction is between (1) those probabilists who believe that it is sometimes rationally permissible (or even obligatory) to accept a highly probable theory and (2) those probabilists who think that scientific theories can be assigned new probabilities in the light of inductive evidence, but never accepted as true, such that even very high probability is insufficient for the rational acceptance of a theory. There are not standard labels for (1) and (2), but I shall use ‘acceptance-theories’ for probabilist theories of type (1) and ‘ascription-theories’ for probabilist theories of type (2). Henry Kyburg was an acceptance theorist[281], whereas Carnap was an ascriptionist[282].

Finally, there are some inductive arguments that have peculiar and interesting logical forms that distinguish them from other forms of induction. These are demonstrative inductions, which are inductive arguments whose conclusions are intended to be deductively implied by the *total* available evidence. For example:

<u>A1</u>

(1) All observed pure samples of the alkali metals dissolve in water.

---

[281] Kyburg (1990) p. 60-61.

[282] Carnap (1968). p. 146.

Therefore, with deductive certainty, (C) All pure samples of alkali metals dissolve in water.

On the surface, A1 is a poor argument, because premise (1) does not imply (C), yet it is supposed to be deductively valid. The argument seems to overstate the support that the premise (1) provides to C. However, in good demonstrative inductions, there is at least one known implicit premise (what Stove calls a "validator"[283]) that results in a deductively valid argument when one makes it explicit:

<u>A2</u>

(1) All observed pure samples of the alkaline metals dissolve in water.

(2) The alkaline metals are elements and if some pure samples of an element dissolve in water, then all pure samples of that element will dissolve in water.

Therefore, (C) All pure samples of alkaline metals dissolve in water.

Demonstrative inductions are inferences from the observed to the unobserved, and thus they are 'inductive' in the sense I am using that term. However, demonstrative induction is distinct from standard inductive reasoning, because it does not have to be formalised as an ampliative argument: articulating the validator as well as the explicit premises creates a deductively valid argument. Ideally, a theory of induction should provide a thorough, intuitive, and useful analysis of both demonstrative and non-demonstrative inductions.

[283] Stove (1982) p. 66.

SECTION 2: EVIDENTIAL PROBABILITY AND INDUCTION

I shall now outline Kyburg's theory of induction. I shall begin by providing a description of the formal framework he uses. Using this framework, I shall propose a probabilistic definition of confirmation that is essentially very similar to the standard Bayesian definition. To close the section, I shall illustrate this definition via a discussion of some types of inductive inference.

2.1 The Formal Framework

Kyburg's formal framework is the application of a metalinguistic function EP to a domain Ω. A metalinguistic function takes sentences of a language as its domain. The function EP itself is not part of that language. The domain Ω is a body of statements, which are an idealized model of our actual total knowledge. Ω includes set theory and any other requisite mathematics for the analysis. Ω is weakly deductively closed, but not strongly deductively closed:

Weak Deductive Closure: If a statement Φ is in Ω and Φ implies Ψ, then Ψ is in Ω. For example, if Ω contains 'Ruthenium is a transitional metal and an element', then Ω contains 'Ruthenium is a transitional metal'.

Strong Deductive Closure: If Ω contains a finite conjunction of statements Γ and Γ implies χ, then Ω contains χ. For example, if Ω contains 'Ruthenium is a transitional metal' and 'Ruthenium is an element', then Ω contains 'Ruthenium is a transitional metal and an element'.

Strong deductive closure implies weak deductive closure, but not vice versa. Both requirements are unattainable ideals, given our cognitive limitations: no person has explicitly derived all the deductive consequences of every statement in which she believes. However, for each form of closure, there are some philosophers who argue that it captures part of how an ideally rational being would reason, as well as providing some important guidance to imperfect reasoners like ourselves.

There are several reasons why one might require weak deductive closure rather than strong deductive closure. Kyburg's primary motivation for endorsing weak closure but not strong closure was his Lottery Paradox[284]. Suppose that there is a fair lottery of 100 tickets in which you will make a random selection. Your best statistical information about whether a particular ticket will be selected is that the ticket will be randomly selected from the 100 tickets. It is assumed that your acceptance standard is such that you will accept any statement H such that EP(H | K) = [x, 1] and $x \geq 0.9$[285].

The evidential probability that the ticket *i* will *not* be selected is [0.99, 0.99] and so you must accept that *i* will not be selected. Since the draw is random, the probability that any other ticket is selected is equal to the probability for *i* and thus there is a [0.99, 0.99] probability that any given ticket will be selected, so you must accept that each ticket will not

[284] Kyburg (1990) p. 64-66.

[285] The value of x is constrained using '≥' rather than '=' in order to allow acceptance when x is greater than the minimum standard.

be selected. This information is incorporated into Ω. If Ω is strongly deductively closed, then you must conjoin all these individual statements regarding each ticket. However, this conjunction implies that none of the tickets will be selected. Yet this is absurd: you know that you will select *one* of the tickets; you just know that any *particular* ticket is very unlikely to be selected. In contrast, if Ω is merely weakly closed, then Ω can contain 'Ticket *i* will not be selected', for each ticket *i*, but nevertheless Ω does not include the claim that no ticket will be selected and a contradiction is avoided.

It is important for Kyburg's paradox that you only know that ticket *i* is a randomly selected lottery ticket. Suppose that *w* is a name whose reference is identified via the definite description 'The winning ticket'. The statement 'The ticket *w* will be selected' has an evidential probability of [1, 1], because if a ticket is selected, then it will be *w*, and your background knowledge implies that a ticket will be selected. Kyburg aims to develop a formal framework in which a body of statements can include both 'Ticket *w* will be selected' and 'Ticket *i* will not be selected', for any given numerical value of *i*, while also avoiding any explicit contradictions.

The Lottery Paradox is not simply a paradox of gambling, because there are similar problems with mensuration. (Mensuration is the process of measurement.) Imagine a group of scientists at Chernobyl, who know that each of a very large set of measurements using their Geiger counter is probably within the standard margin of error ε for that instrument. They also believe that at least one measurement had an error outside of ε, because there will almost certainly be such errors in a very large set of measurements. If the scientists accept both claims into Ω, then they would believe the contradictory claims that 'All of the measurements were accurate within ε' and 'At least one of the measurements was not accurate within ε'.

Requiring only weak deductive closure allows the scientists to believe that each particular measurement was within ε, but not that all of them were.

Another reason for rejecting strong deductive closure is that it creates problems when discussing sets of mutually inconsistent scientific theories. Kyburg does not discuss this motivation, but it is an added benefit of his model of scientific knowledge. Assume that the Standard Model of Particle Physics and the Standard Model of Cosmology are inconsistent. A formal analysis of a model of scientific knowledge that contained both models could be interesting, but if the conjunction of these models is contradictory and we are using classical logic, then there will be problems due to logical explosion. In Kyburg's framework, Ω will contain the consequences of a conjunction $(\Phi \wedge \Psi)$ only if either (1) it is learned directly or (2) it is acceptable given the evidence. If $(\Phi \wedge \Psi)$ is a contradiction, then it cannot be learned directly, nor can it be highly probable given Ω, and thus neither condition (1) nor condition (2) can be satisfied. Therefore, even if the domain Ω contains both Φ and Ψ, the domain will not contain $(\Phi \wedge \Psi)$, so that Kyburg's framework enables us to apply the EP function to a domain Ω that includes mutually inconsistent scientific theories.

There are some statements in Ω which are our most certain evidence. Kyburg calls such statements the "Ur-Corpus". Other statements can be acceptable as evidence if they are sufficiently probable relative to the Ur-Corpus. He calls these statements the "Evidential Corpus". Finally, some statements are acceptably probable relative to the Evidential Corpus. This last category are the "Practical Certainties". I shall not discuss, in this section, Kyburg's views on which statements should be included in the different corpora, because such a discussion would raise questions of general epistemology, not confirmation theory or decision theory, which are my focus in this thesis. However, it is significant that Evidential

Probability is consistent with a multiplicity of different theories of evidence. As in Bayesianism, the acquisition of evidence that can be included in Ω is an exogenous matter and different Evidential Probabilists can have different views on such questions.

I shall use the following definition of ‘inductive argument’:

Inductive Argument: An argument A_i is inductive if and only if it is a formalisation of an inductive inference such that the explicitly cited premises in A_i are known via observation.

Finally, one can measure the strength of such an argument by determining the evidential probability of the conclusion relative to the premises and any other relevant statements in Ω[286].

2.2 Confirmation

One theme of my thesis is that there are many important aspects of scientific reasoning that should be formalised when modelling confirmation. In this chapter, I am interested in the balance of evidence. To simplify, in this chapter I shall use ‘confirmation’ to mean a positive change in the balance of evidence, as opposed to a broader notion of evidential support. The standard Bayesian definition of this sense of confirmation is very straightforward:

Bayesian Definition of Confirmation: E confirms H if and only if $P(H \mid E) > P(H)$.

[286] Kyburg (1990) p. 69.

- and disconfirmation occurs when P(H | E) < P(H).

Kyburg never offered an explicit definition like this, but the Bayesian definition can be adapted into the Evidential Probabilist framework. I shall analyse confirmation as a three-place relation between a hypothesis H, an evidence-statement E, and the relevant background knowledge K:

Evidential Probabilist Definition of Confirmation: E confirms H relative to K if and only if EP(H | E ^ K) = [x, y], EP(H | K) =[z, w], and $1/2(x + y) > 1/2(z + w)$.

- and disconfirmation will occur when $1/2(x + y) < 1/2(z + w)$[287].

Clearly, the Evidential Probabilist definition of confirmation is very similar to the Bayesian definition. The salient differences are (a) Kyburg allows the acceptance of hypotheses and (b) the probabilities are interval-valued. I shall not discuss (a) and the surrounding controversies about acceptance theories of induction versus ascription theories. Although (b) shall be important for much of the discussion below, it does not significantly affect the basic definition of confirmation, since using the mean of an interval's limits provides a number that is formally very similar to precise probabilities[288].

---

[287] The use of $1/2(x + y)$ for converting an imprecise probability into a single number was proposed by Walley (1991) p. 522.

[288] There are some differences, e.g. there can be a change in the evidential probability without a change in the mean of the interval. For example, 1/2[0.2, 0.4] is identical to 1/2[0.23, 0.27].

2.3 Statistical Induction

In his theory of induction, Kyburg's primary focus is statistical induction. An induction is a "statistical induction" in his theory if H is a hypothesis about a population, E is a sample-report about that population, and K is the relevant background knowledge.

On Kyburg's analysis, learning E confirms H relative to K when it is probable that E describes a representative sample of the population described by H[289]. Like Bayesians, Evidential Probabilists do not require that this sample is randomly selected. Instead, what matters is whether it is probable (relative to K) that the sample is representative. Furthermore, the uniformity of nature as a whole is not important for a particular statistical inductive inference: what matters in Kyburg's theory is the probability that *this* particular sample is representative of the particular population under investigation. Methodologically, Kyburg's analysis of statistical induction directs our attention to examining issues of probable representativeness given our total evidence, in contrast to a fundamental focus on randomization or postulates about uniformity.

Background knowledge plays an important role: what we know about the parameters of the population can affect the evidential probability of the hypothesis given the sample-report. For example, if K includes the claim that the population is approximately normally

[289] We can define 'representative sample' for different contexts using margins of error. For example, in political opinion polling, psephologists (statisticians who study voting patterns) generally aim for margins of error of about 3% at a confidence of level of 95%, so that at least 95% of samples with this size will match the actual population frequency within ± 3%. (The precise margin of error increases as the sample mean for a party becomes further from 50%.) In other contexts, we might be willing to use much wider or narrower margins of error. The evidential probabilities will vary (*ceteris paribus*) in proportion to the width of the margins of error: the larger margin of error, the more probable the statistical generalization will be given the evidence, because there will be a greater proportion of possible samples with means within this margin of error of the population mean. (Additional issues are raised when the population in a statistical induction is infinite, but I shall not discuss them here.)

distributed, then someone who accepts K can use E for parametric statistical tests about H. (Parametric tests require some information about the parameters of the population like its distribution; non-parametric tests can be used even if this information is unavailable.) For example, suppose that we know that the weighting of a sample is such that the sample is probably representative, in the sense that (1) it will produce results within a margin of error of ± 3% in at least 95% of a long-run series of trials and (2) this long-run tendency is our best statistical information regarding the representativeness of the sample. We assume that the population is normally distributed. Under these assumptions, the evidential probability of a claim like '30% ± 3% of voters intend to vote Labour at the next UK general election' would be EP(H | E ^ K) = [0.95, 1].

There is a clear similarity between this approach and confidence interval methods. Indeed, the choice of 95% corresponds to the standard choice of a confidence level in classical statistics. If 1/2[0.95 + 1] = 0.975 > 1/2(z + w), where w and z are the limits of the interval EP(H | K), then learning E has confirmed H relative to K by statistical induction.

In contrast, if K includes reasons to believe that the sample described in E is probably *not* a representative sample, then E might not confirm H relative to K, because EP(H | E ^ K) = EP(H | K). For instance, K could include the information that a bitter employee has selected the sample from a larger and unknown sample and that the employee intends to mislead the company. Thus, we know that the sample is probably unrepresentative and E does not confirm H relative to K.

Under some circumstances and for a fixed margin of error, the probability of representativeness will increase in a linear fashion as the sample size increases, but there is no simple universal relationship between sample size and the probability that the sample is representative. (Except for trivial facts, like that the sample must have more than zero members and that a sample size that is equal to the population size will always be representative.) Relative to some background knowledge, a small sample might be more probably representative than a large sample. For example, we might know that the large sample will contain a large and unspecified number of voters who have been selected from a highly unrepresentative region of the country, so that the sample mean of the large sample is comparatively less likely to be representative of the population mean. More generally, there is no mechanical relationship in Evidential Probability between sample size and probability; the data from *both* E and K are important. Scientists can sometimes use small samples for powerful statistical inductions, as in well-controlled laboratory experiments. Relative to some background knowledge, one can infer (within a margin of error) the acidity of all pure instances of a new compound from a few tests of its acidity. Consequently, in Kyburg's theory of statistical induction, it is the probability that a sample is representative that is always the salient issue, and sample size (such as whether the selection procedure was random) is only important insofar as it is relevant to representativeness.

2.4 Eduction

Eduction is a form of induction with the following pattern:

(1) A particular set of individuals a_1-a_n are F in some proportion p.

Therefore, probably, (C) An individual a_{n+1} is F.

Imagine that an alien is eating some green apples for the first time. The labels a_1-a_n refer to each of the green apples that the alien tastes. Suppose that the alien's taste buds are very similar to human taste buds and she experiences a sour taste for each of the apples. The alien eductively infers that the a_{n+1} apple will also be sour. I take the term 'eduction' from W. E. Johnson[290]. Other names for this form of inference are "singular predictive induction"[291] or "instantial induction"[292]. Some philosophers of science, like Rudolf Carnap[293] and John Stuart Mill[294], have regarded this form of induction as particularly important.

Kyburg models eduction as two separate inferences, in contrast to Carnap and Mill who model it as a single inference[295]. Firstly, there is the statistical inference from (1) the conjunction that ($Fa_1 \wedge Fa_2 \wedge \ldots \wedge Fa_n$), which I shall abbreviate to E, to (2) a statistical generalisation H about a population, whose members are a_1-a_n and a_{n+1}, where H states that this population's proportion of F matches the proportion in E. If the evidential probability of H given E and the background knowledge is greater than the acceptance threshold, then she can accept H into K.

Secondly, there is the direct inference from (H ^ K) to Fa_{n+1}. Assume that the evidential probability of this direct inference is $EP(Fa_{n+1} \mid H \wedge K) = [0.99, 1]$. Provided that the mean, 0.995, exceeds the mean for Fa_{n+1} relative to K alone, learning that E has confirmed the prediction Fa_{n+1} via the indirect route of H. When 0.99 exceeds the agent's

[290] Johnson (1924) Chapter IV.

[291] Carnap (1962) p. 569.

[292] Kyburg (1990) p. 59.

[293] Carnap (1962) p. 574-575.

[294] Mill (1882) p. 142.

[295] Kyburg (1990) p. 68-69.

acceptance threshold, then the agent must add Fa_{n+1} to the corpus of Practical Certainties.

Background knowledge can bar either inferential step in eduction. The statistical induction might be blocked due to concerns about the sample's representativeness, as I have already discussed in Subsection 2.3. The direct inference could also be blocked because the population described in H is not the proper reference class for the probability of Fa_{n+1}. In the apple example, the alien might know that the apple a_{n+1} has been injected with a substance that will give it a sweet flavour, such that she should use a statistical generalisation about sweetness in the reference class of injected green apples.

A more complex example can be taken from psephology (the study of voting patterns) in order to illustrate the subtle manner in which background knowledge of relative frequencies can affect inductive inference. Imagine that we are campaigners in a political party and we are wondering whether to send our literature to Ms. Smith. We have two recent opinion polls: one poll for voters in Great Britain as a whole and another poll for the voting intentions of voters in Scotland[296]. From these polls, we inferred statistical generalisations for Great British voters and Scottish voters. Suppose that part of our decision regarding whether we should send her campaign literature depends on whether we expect Ms. Smith to vote for the Scottish National Party (SNP). From our statistical database, we know that Ms. Smith is eligible to vote in the next UK general election. If we use the information from the Great Britain poll, then the probability that she will vote SNP might be a very low value, like [0.03, 0.04]. However, if we know that Ms. Smith lives in Scotland, then Evidential Probability means that we should use the Scottish poll, for which the probability of her intending to vote

[296] Usually, psephologists in the UK only study Great Britain (England, Scotland, and Wales) and exclude Northern Ireland due to the extreme differences in voting patterns.

SNP might be something like [0.36, 0.38].

Summarising, Kyburg models eduction as two inferences rather than one. It can be blocked due to problems with either inference: (1) relative to K, the sample cannot be used for the statistical induction and/or (2) relative to K, the population consisting of the sample and the individual under inquiry does not provide the appropriate reference class for the direct inference.

2.5 Eliminative Induction

Kyburg does not provide an extensive account of eliminative induction. However, he regards it as appropriate in contexts of relatively rich background knowledge[297]. A strong version of eliminative induction occurs when K contains the knowledge that one of a finite partition of hypotheses (H_1 v … v H_n) is true, while E implies that all of the hypotheses other than H_1 are false. If (E ^ K) is acceptable, then learning E implies that H_1 is true. For instance, if Miss Marple accepts the conjunction of clues that (1) one member of a set of suspects committed the murder and (2) all the suspects except the Under Secretary have an acceptable alibi, then she can deduce that the suspect without an acceptable alibi committed the murder.

A weaker version of eliminative induction can occur when there is a known Bayesian-style joint distribution for the partition of hypotheses. Under such circumstances, learning that one particular hypothesis is false will lead to the reallocation of the probability among the surviving hypotheses. For example, imagine that I have randomly drawn a card from a

[297] Kyburg (1980) p. 628.

normal deck behind a screen. Initially, the probability that the card is a particular type *i* is 1/52, but if you accept my claim that the card is not the Ace of Diamonds, then the probabilities for every other type *i* will increase to 1/51.

This theory of eliminative induction is similar to Bayesian analyses. One difference is that Bayesians typically require strong deductive closure, whereas Evidential Probabilists do not. For Kyburg, there is an important question of the acceptability of the *conjunction* of the claims that (1) (H_1 v … v H_n) is a finite partition and (2) some number *x* of the members of this partition are false. If the conjunction of (1) and (2) is insufficiently probable relative to the Ur-Corpus, then it cannot be accepted as evidence (i.e. into the Evidential Corpus) and thus the probability might be unchanged after learning (2).

2.6 Demonstrative Induction

Kyburg also discusses demonstrative induction[298]. In some cases, K might include the claim that a sample is representative of a population, such that the conjunction of our evidence E with K might *deductively* imply a hypothesis H about that population. If (E ^ K) is acceptable, then H will also be acceptable. For instance, if we know that the melting point of an element A is a physical constant and that a single sample of A has a melting point of $r \pm \varepsilon$, then we can infer that the melting point of the element A is $r \pm \varepsilon$. I shall discuss this form of induction in more depth in Subsection 3.3.

[298] Kyburg (1976).

Summary

In Kyburg's theory of induction, the interaction of evidence and background knowledge is crucial. Additionally, the use of probability enables a quantitative analysis: an Evidential Probabilist is not limited to qualitative claims like 'E confirms H relative to K', but she can also say that 'H has a probability of [x, y] relative to E and K'. In contrast to Bayesianism, there is no prior distribution: all the probabilities are ultimately derived from information about relative frequencies. Indeed, perhaps Kyburg's most interesting contribution to the theory of induction is an epistemic theory of probability that can analyse inductive reasoning and which involves no prior probabilities. The absence of a prior distribution will be crucial to many of my arguments in the next section.

SECTION 3: NORTON'S CRITICISMS OF PROBABILISTIC THEORIES OF INDUCTION

Norton makes several criticisms of Bayesianism, which is currently the predominant probabilistic theory of induction. I shall argue that Evidential Probability enables a probabilist to avoid the criticisms he develops for Bayesian theories of induction. I shall also argue that a general problem that he raises for what he calls "formal" theories of induction does not apply to an Evidential Probabilist theory. Finally, I shall explain how Evidential Probabilists can analyse a puzzle that Norton raises for all theories of induction.

3.1 Norton's Criticisms of Bayesianism

Norton has been one of the most notable critics of probabilistic theories of induction in recent years. Since Bayesianism is the most common probabilistic theory of induction, it has been his principal target. In this section, I shall not take a position on whether his problems are insurmountable for Bayesianism. Instead, I shall merely argue that they are not problematic for an Evidential Probabilist.

3.1.1 Bayesianism and Precision

Norton claims that Bayesianism involves "spurious precision", in the sense that a Bayesian theory enables inductive reasoning that involves a level of exactitude that exceeds what the evidence can intuitively provide[299]. Kyburg makes a similar criticism of Bayesianism: he even uses an almost identical phrase, "implausible precision", to describe many Bayesian probability statements[300]. He gives the example of the statement 'A ball is purple'. For an agent who is a Subjectivist, Objectivist, or Logicist Bayesian, the prior probability of this statement might be 0.01. The conditional probability of 'A ball is purple' given 'This other ball is purple' might be 0.02. Intuitively, according to both Norton and Kyburg, the evidence does not warrant this degree of precision.

Naturally, Kyburg would disagree that this charge applies to his own system. As I explained in Chapter 1 Subsection 5.3, Evidential Probabilists regard the precision of Bayesian analyses of induction as appropriate only when there is sufficient relative frequency

[299] Norton (2011) p. 402.

[300] Kyburg (1974) p. 282.

data. Thus, Kyburg's examples of Bayesian reasoning in Evidential Probability tend to involve simple and well-understood gambling apparatuses[301]. In contexts like the evaluation of interesting hypotheses, he regards imprecise probabilities as better representations of the evidential relations. Thus, Evidential Probabilists agree with Norton that Bayesian probabilities can be inappropriately precise. In contrast to the Bayesian, the Evidential Probabilist restricts the precision of Bayesianism to circumstances where our statistical information is precise, and under such circumstances, such precision is not counterintuitive.

Of course, there is always some degree of idealization in formal epistemology, and the limits of an Evidential Probability interval will typically be precise. However, even idealized precision in Kyburg's system cannot exceed the idealized precision of the evidence, because all intervals must be based on relative frequency data. Consequently, any precision in Evidential Probabilism does not appear to be spurious: if the intervals seem too precise, then one can always reduce the precision by weakening the relevant statistical evidence. For instance, if our initial evidence includes the claim that 49-51% of rabbits in the laboratory warren are female and the class of 'rabbits in the laboratory warren' is the appropriate reference class to determine the evidential probability that a particular rabbit from the warren will be female, yet [0.49, 0.51] seems an implausibly precise interval in this context, then we can reduce the precision by accepting only the weaker claim that a wider interval-valued proportion of rabbits in the warren are female, such as 48-52%, 38-62%, and so on. Even if *any* precision seems spurious, the maximally imprecise [0, 1] interval can be used, as discussed in Chapter 1 Subsection 5.3. By such adjustments, an Evidential Probabilist can always avoid any spurious precision via weakening the body of accepted evidence in context. Such adjustments do not make Evidential Probability an arbitrary system: they are

301 E.g. Kyburg and Teng (2001) p. 216-217.

appropriate only if we have incorrectly specified our relative frequency evidence in a particular context by accepting excessively precise statistical claims. Another option is to use inequalities like > 0.5 or < 1 as the limits of the intervals when this is the best representation of the available relative frequency data.

3.1.2 Prior Probabilities

Norton also directs a number of criticisms against Bayesian prior distributions. Opponents of Bayesianism have targeted this part of Bayesian epistemology since its earliest days: George Boole criticises the use of Bayes's Theorem (except when the terms are based on known relative frequencies) in his critique of proto-Bayesians like Pierre-Simon Laplace and Augustus De Morgan[302]. Norton's arguments are particularly interesting, because he develops this criticism beyond the mere charge of arbitrariness. Firstly, he criticises Bayesian priors for the "imaginary" nature of a complete prior distribution. Secondly, he notes the possibility that scientific theories will not always determine a real-valued prior probability for possible evidence, so that a Bayesian cannot always avoid his first objection in local contexts by deriving priors from accepted scientific theories. I shall describe each of Norton's points in turn.

Firstly, Norton criticises what he calls the "curious" fact that, in Bayesianism, the evidential relations are determined by a prior probability distribution that is entirely imaginary, because no-one actually has a Bayesian prior probability distribution over all the

[302] Boole, G. (1958) p. 286-292.

statements that they know[303]. Norton does not precisely state why this unreality is problematic, but I shall develop one challenge that it raises for a Bayesian.

As a preliminary, one must clarify what a “prior probability distribution” means in this context. I shall assume, initially, that Norton is interested in Subjective Bayesian priors. As described in Chapter 1 Subsection 2.1.3, Subjective Bayesians hold that epistemic probabilities are rational degrees of belief, where “rationality” is satisfied by satisfying the constraints of the axioms of additive probability and perhaps some further restrictions. For a Subjective Bayesian, the challenge that Norton presents is to justify the normative significance of rational Bayesian prior probability distributions for our actual inductive practices, on the assumption that we do not actually possess such prior distributions.

In conditionalization, when E does not entail or contradict H, the extent to which E confirms a hypothesis H will depend on an infinite number of unconditional probabilities. Conditioning H on E involves replacing P(H) by P(H | E) in the new distribution. Assuming that P(E) > 0, Bayes’s Theorem defines P(H | E) as:

$$P(H \mid E) = \frac{P(H \wedge E)}{P(E)} = \frac{P(E \mid H)P(H)}{P(E)}$$

P(H) is one term in the conditional probability. In turn, this term will depend on the complete probability distribution. For example, if (H, H′… H_n) is an exhaustive and mutually exclusive set of hypotheses, then:

[303] Norton (2003) p. 662.

$$P(H \ v \ H' \ldots \ v \ H^n) = 1$$

Thus, prior probabilities of a vast number of rival hypotheses put restrictions on the extent to which learning E increases the probability of H. Since it is possible that $n = \infty$ (it is possible to mechanically construct contrary hypotheses for H in any domain that is rich enough to represent scientific knowledge) the value of P(H | E) depends on the value of an infinite number of probability values. The prior distribution is a vital part of a Bayesian analysis induction, but Norton correctly notes that no human being has such a distribution for any algebra of statements that could plausibly represent the corpus of scientific knowledge.

A Subjective Bayesian could respond that, when 0 < P(H) < 1 and 0 < P(E) < 1, one can have a good model of the evidential relationship between H and E via simply examining P(E | H) and P(E | ¬H), as E will confirm H given those assumptions if P(E | H) > P(E | ¬H). Yet, under those assumptions:

$$P(E \mid H) = \frac{P(E \wedge H)}{P(H)} = \frac{P(E)P(H \mid E)}{P(H)}$$

- so that P(E | H) is only defined when P(H) and P(E) are also defined and the necessity of the full probability distribution is once again apparent.

Objective Bayesians and Logical Bayesians might object that their prior distributions are not imaginary, because the probabilities exist independently of their correspondence to any actual person's degrees of beliefs. In other words, in such interpretations of Bayesianism, the correct value for P(H | E) exists for any suitable domain, and we can use it for analysing inductive reasoning. However, Norton's point still stands if we are analysing a real instance of inductive reasoning, because it is mysterious how a prior distribution that need not remotely correspond to the actual person's overall degrees of belief can provide a serious normative constraint and how updating such a prior distribution can constitute the method of inductive reasoning, any more than updating an imaginary person's subjective credences can constitute the inductive method. Additionally, the justification of these idealized priors raises yet another horde of problems for a Bayesian.

Another possible response could be to deny that their epistemology requires a simple correspondence between coherent degrees of belief and our actual beliefs, so that the incoherence of our actual degrees of belief is compatible with the relevance of Bayesian norms for our inductive practices. Put another way, there are a variety of *bridge principles* between Bayesian priors and human reasoning[304]. For instance, a Subjectivist could argue that the guidance from their theory of induction is not that one should attempt to formulate a coherent prior distribution and conditionalize using those priors; instead, they could argue that one should try to approximate a coherent agent conditionalizing upon the evidence one encounters. A similar approach that Howson and Urbach develop is to interpret Bayesian norms such as coherence and conditionalization as standards to which reasonable people commit themselves, rather than descriptions of the reasoning that they do or can, in practice, perform[305]. Furthermore, the Subjectivist (and indeed Objectivist) interpretation of degrees of belief has at least as much flexibility as the interpretation of "belief". Norton has certainly not shown that there is a problem here for every variety of analysis of belief: one might have a

[304] A bridge principle connects some matter of fact – such as the fact that an action will cause an increase in expected human suffering or the logical inconsistency of a person's belief – with some normative claim, such as a prohibition against that action or that the person should change their beliefs.

[305] Howson and Urbach p. 422-423.

theory of beliefs in terms of overt behaviour, dispositions, a *sui generis* psychological state, a form of near-knowledge, or one of a panoply of other possible analyses. In short, while there is a potential problem for Bayesians that Norton raises, Bayesian epistemology is very diverse and no short argument is likely to be problematic for every type of Bayesian.

There are further defences that a Bayesian might make in response to Norton's criticism, but since my goal is merely to motivate Norton's concern, I shall proceed onto Norton's next argument regarding prior probabilities on the assumption that imaginary priors are at least *prima facie* problematic for Bayesians. One can sometimes use scientific theories to derive Bayesian priors. As a result, it might seem as though Bayesians can avoid the problem of imaginary priors in some local contexts. For instance, the probability of a particle undergoing radioactive decay in a particular interval of time might be determined by the accepted theory of its half-life. By founding their priors on values from scientific theories, Bayesians might attempt to escape Norton's charge, at least in the context of applying their methodology to a particular case of inductive reasoning: once priors for the hypotheses have been derived in this way, a Bayesian analysis of induction can proceed in terms of updating these priors via conditionalization.

However, Norton argues that a Bayesian cannot always obtain a real-valued probability using this method[306].

[306] Norton (2003) p. 660-661 and (2007) p. 166-169.

Key

X: A random variable.

H_1: The hypothesis that X will have a value *r*.

H_2: A nondeterministic scientific theory that merely says that *r* is a member of an infinite set of equiprobable possible values of X.

Clearly, H_2 is insufficiently informative to provide a precise value for the probability of H_1. However, such theories might be the best that is available in a particular case of inductive reasoning. Norton gives the example of the Steady State Theory of Cosmology: according to this theory, there is an infinite space in which hydrogen atoms can materialize, but this materialization is both random and uniformly probable for any point in space-time[307]. Since the Steady State Theory says that hydrogen atoms can materialize randomly and uniformly across an *infinite* region, there is no possible positive real-valued uniform distribution[308].

Norton's point is most obvious if the probability distribution satisfies the Principle of Countable Additivity. Suppose that the domain of P consists of sets and UA_i is a mutually disjoint union of sets[309]. For P to satisfy the Principle of Countable Additivity, it must be the cast that $P(UA_i) = \sum P(A_i)$ for every A_i in the union. Let UA_i be an infinite set consisting of possible values for a random variable X. The distribution of P over the members of UA_i is

[307] Norton (2003) p. 660-661.

[308] A uniform distribution is a distribution which has constant probability. In the simple case of a discrete distribution, this means that each statement is equiprobable. Uniform continuous uniform distributions are those in which equally-large intervals in the continuum of possible values of a random variable are all equiprobable.

[309] A disjoint union of sets is a set that contains the union of some sets that share no members.

uniform if and only if $P(A_j) = x$ for every arbitrarily selected set A_j. Since UA_i is countably infinite disjoint union and P satisfies the Principle of Countable additivity, $P(UA_i) = \sum P(A_i) = 1$. If there is a uniform distribution, then $x \geq 0$. However, if $x > 0$, then $\sum P(A_i) = \infty$, contradicting the assumptions; whereas if $x = 0$, then $\sum P(A_i) = 0$, also contradicting the assumptions. Therefore, if P satisfies the Principle of Countable Additivity, there is no possible uniform distribution for an infinite set of possible values for a random variable X[310].

However, the Principle of Countable Additivity is not an essential feature of a Bayesian probability function. If P does not satisfy this principle, then we can consistently assign a probability of zero to the materialization of the hydrogen atom for each particular region of space. Yet, in such a distribution, the probability of a hydrogen atom materializing at any particular point is the same as the probability of a contradiction. That does not seem to reflect the actual epistemic situation: the Steady State Theory states that it is *possible* that such a materialization could occur, which is intuitively different from stating that it is as *unlikely as a contradiction*. Furthermore, in standard Bayesianism, such an assignment would entail that it was impossible to increase the probability of a hydrogen atom materializing at a particular point in space-time by acquiring additional evidence, because a hypothesis with an extreme value cannot have its probability increased via conditionalization. This is a counterintuitive result: one would think that the mere acceptance of the Steady State Theory does not warrant certainty that ‘A hydrogen atom will materialize in front of my nose’ is unconfirmable by any possible evidence.

A Bayesian might try to avoid this problem by assigning an infinitesimal probability

[310] Howson and Urbach (1993) p. 34.

to each possible coordinate position. An infinitesimal number is a non-zero number that is so small that is cannot be measured. Unlike real values, infinitesimals can have a product that is greater than 0, such that a set of mutually exclusive and exhaustive hypotheses could have a joint product of 1, in accordance with axioms (i) and (iii) in Chapter 1 Subsection 2.1.2. However, this is not a possible output for a standard Bayesian probability function, because infinitesimals are not members of the set of reals, which is the co-domain of Bayesian functions.

Norton gives another example of this problem that is especially exciting, because it occurs in Newtonian mechanics, and this theory has traditionally been thought to be deterministic[311]. Suppose there is an object with a point mass, on top of a perfectly symmetric dome, and the object can slide in any direction without any friction whatsoever across the Dome. According to Newtonian mechanics, it is possible that the point mass will stay on top of that dome forever. However, it is also possible that the point mass will slide in any one of an infinite number of vectors across the Dome. It is even possible that the point mass will be motionless for any of an infinite number of possible periods of time and *then* slide (without any external force) along any of an infinite number of possible vectors across the Dome. If we attempt to use this theory to determine a real-valued prior probability that 'The point mass will move along a particular vector at a particular time period *t*', then we shall have to arbitrarily impose a non-uniform distribution or assign a value of zero for all the possible vectors, but this assignment is unsatisfactory for the same reason it was unsatisfactory in the example of the Steady State Theory.

[311] Norton (2007) p. 166-169.

John Worrall objects that a Subjective Bayesian can happily accept arbitrariness in the Dome[312]. (Worrall's point also applies for the Steady State Theory example.) Subjective Bayesianism is famously consistent with very arbitrary distributions. However, this does not elude Norton's criticism, because the values of such a distribution are not derived from accepted scientific theories and so the problem of imaginary priors returns[313]. Norton confronts Bayesians with a dilemma: priors that are not taken from scientific theories are objectionable because they form an imaginary distribution, but scientific theories will not always supply priors, even when they provide some potentially relevant information. It is possible that the Bayesian can address this dilemma, but Norton has certainly presented a live challenge to Bayesian theories of induction.

Evidential Probabilists avoid both of these problems by abandoning the ambition of a prior distribution. In Evidential Probability, something resembling a Bayesian prior distribution can only be obtained in epistemically rich contexts in which there is precise relative frequency data that can be used to determine a full distribution, so that there is no oddity of an imaginary prior distribution. Since evidential probabilities come from the available relative frequency data, there is no need to worry about a complete real-valued distribution over *all* statements in the domain of the Evidential Probability function. What matters are the formal relations between (1) the evidence plus the background knowledge and (2) the hypothesis just as all that matters for a deductive logic claim like '$\Gamma \vdash \Phi$' are the derivation rules in the deductive system in use and the formal relations between Γ and Φ.

[312] Worrall (2010) p. 752.

[313] If the distribution only applies to the statements under consideration, then there will be the danger the agent's total credences are inconsistent with the axioms of the probability calculus, which is an unacceptable possibility in conventional Bayesian epistemology.

If a scientific theory only provides an imprecise value for the relative frequency of some statement (like ‘A hydrogen atom will materialize at point *x* in space-time’) and that scientific theory is our best available basis for the probability of that statement, then Evidential Probabilists have exactly the formal tools of imprecise probability to represent that statement’s probability given the scientific theory. At the extreme, they can use the [0, 1] interval, as explained in Chapter 1 Subsection 5.3. Overall, the precision of the intervals need not (and must not) exceed the precision of the relative frequency data provided by the scientific theory and the available background knowledge.

Norton might object that the domain of Evidential Probability is weakly deductively closed, such that every statement’s consequences are also in the domain, and this requirement is beyond what any human being can possibly achieve. This weak deductive closure is certainly an idealization, but it involves real deductive logical relations, rather than imaginary prior probability values. That a particular hypothesis entails another hypothesis is (presumably) a true or false question. In contrast, outside of special cases, whether $P(H \mid E) = r$ is true is a question that has literally no answer in Bayesianism unless P is fully defined over the complete domain to which H and E belong. In addition, one might regard weak deductive closure as placing a norm on an agent’s commitments: if I believe H_1 and H_1 implies H_2, but I have never explicitly considered H_2, then it is arguable that I am committed to believing H_2, and must either accept it or reject H_1 when I consider this hypothesis. That is quite a different matter from being committed to an imaginary prior distribution(s?). In short, there is no *prima facie* case for a strong “ought implies can” assumption in formal epistemology. Furthermore, Evidential Probability still has much to say about inductive inference even in domains that are not fully deductively closed. Therefore, even assuming that Norton's criticism of Bayesianism is sound, it does not apply to Evidential Probability.

I shall now explain how an Evidential Probabilist can analyse Norton's examples without the use of imaginary priors. In the case of Steady State Cosmology, suppose that we initially have a set of statements that includes the Steady State theory and this set is our best statistical basis for the evidential probability that a hydrogen atom will materialize at *x*, where *x* is a particular point in space-time. From the Steady State theory and the properties of hydrogen atoms, we can deduce that it is possible that a hydrogen atom will materialize at *x*. However, the theory tells us nothing about the relative frequency of this event, and hence the evidential probability is [0, 1]. Of course, depending on the available background knowledge, it is possible that the lower or upper limits are different from 0 or 1. For instance, suppose that our total evidence contains some reasons to think that it is extremely unlikely that a hydrogen atom will materialize at *x*. An Evidential Probabilist can assign a very low probability to the hypothesis, like $[0, 1^{-67}]$, with the limits depending on the relative frequency data in the total evidence.

The analysis for the Dome example is essentially identical. Suppose that Newtonian physics and the available background knowledge jointly provide absolutely no guidance regarding the relative frequency of the object moving along any given vector $\overrightarrow{AB}$. It does not even provide us with a relative frequency for how often (and for how long) the object will be motionless on top of the dome. Once again, an Evidential Probabilist must use the [0, 1] interval: for all we know in Norton's scenario, point-massed objects on a frictionless dome will never move along $\overrightarrow{AB}$, or sometimes move along $\overrightarrow{AB}$, or even *always* move along $\overrightarrow{AB}$. If we have richer background knowledge, then a narrower interval (or even a precise interval) might be applicable.

In general, the [0, 1] interval enables Evidential Probabilists to have probabilities that are (a) positive and (b) uniform across an infinite set of contrary hypotheses. It is mathematically impossible to satisfy both (a) and (b) with real-valued probability functions. However, it is possible to satisfy both *desiderata* with interval values, as a function can consistently assign [0, 1] to each member of the set.

3.1.3 The Problem of Old Evidence

Another problem for Bayesianism that Norton uses is Clark Glymour's "Problem of Old Evidence"[314]. This is the alleged problem that the comparison of P(H) with P(H | E) (or P(H | B) with P(H | E ^ B)) is insufficient to determine whether E confirms H, because E might be "old evidence" that has already been used for conditionalization, such that P(H) = P(H | E). Norton uses this problem to object against the very idea of using a degree of belief or degree of support function.

I shall give an example of the Problem of Old Evidence. Suppose that we learn the evolutionary history of some newly discovered species of deep-sea fish and this history is what we would expect if that species evolved by natural selection. We suppose that we have a probability distribution P over a set of statements {E, H} with values defined for P(H), P(H | E), and P(E), such that P(E) < 1. Upon learning E, our distribution shifts to P´, in which P´(E) = 1. To determine P´(H), we use the conditional probability P(H | E). Representing Darwinian evolutionary theory by H and our new evidence by E, we formulate a new probability

[314] Norton (2011a) p. 402-403.

distribution P′, in which P′ is calculated by Bayes' Theorem and the relevant parts of the old prior distribution:

$$P'(H) = \frac{P(H)P(E\,|H)}{P(E)}$$

Under the conditions that:

(a) $0 < P(H) < 1$

(b) $0 < P(E) < 1$

(c) $P(E \mid H) > P(E)$

- then P′(H) > P(H). Just as we would expect, E confirms H.

The Problem of Old Evidence looms when P(E) = 1. For example, in *The Origin of Species*, most of Darwin's evidence consists of pre-existing biological knowledge. Let E′ represent this old evidence. As E′ was already known, it will be assigned the value P(E′) = 1. For any hypothesis H_i such that $P(E' \mid H_i)$ is defined, then $P(E' \mid H_i) = 1$. Plugging these values into Bayes' Theorem gives:

$$P(H \mid E') = \frac{P(H)1}{1} = P(H)$$

- and hence P′(H) = P(H). Updating a model of Darwin's reasoning using Bayesian conditionalization gives the absurd result that the accepted evidence in 1859 did not confirm

Darwinism.

I shall not discuss the Bayesian responses to the Problem of Old Evidence, because my focus is whether Norton's criticisms are problems for Evidential Probability, as opposed to whether they are problematic for Bayesianism. Kyburg's system avoids the Problem of Old Evidence because updating in Evidential Probability is invariant to the time at which the evidence is learned: if we can use (E ^ K) to derive new relative frequency information about H and this information is the proper basis for H's evidential probability (according to the rules in Chapter 1 Subsection 5.3) then H can have a new probability, regardless of whether E was learned before or after we considered the probability of H. The key formal difference with Bayesianism is that, in Evidential Probability, updating is not a matter of using a distribution of unconditional probabilities, because there are only conditional probabilities in Evidential Probability. Changes in the probability are purely a result of changes in the accepted evidence, rather than an ongoing series of shifts from a prior distribution to new probability distributions.

In the case of Darwinism, we can say (greatly simplifying the subtleties of Darwin's arguments) that natural selection was an extremely good explanation of the evolution of the species that Darwin discussed. I shall suppose that the intended scope of Darwinism was the development of the *overwhelming majority* of distinct species on Earth. Thus, the theory is a statistical generalisation asserting that almost all species evolved via natural selection[315]. Darwin's reasoning can be (greatly) simplified as follows: given that natural selection is the best explanation for the species that he discussed, it was probably the best explanation for

[315] There is a solid historical justification for this interpretation: in *The Origin of Species*, Darwin was notably taciturn regarding whether humans had evolved by natural selection. This rhetorical strategy would have been pointless if his theory of evolution was a universal generalisation over all species.

most species' evolutions. The central facts that he needed to establish were (a) that natural selection was the best explanation of this sample of species and (b) that the sample he discussed was representative of species in general.

This statistical induction could be formally modelled using the strategy I discussed in Subsection 2.3. Some salient factors in such an induction are (1) whether natural selection was the best explanation of the evolutions that Darwin discussed, (2) the probability that Darwin's sample was representative, and (3) the probability of such a statistical inference leading to error. Historically, all three factors were controversial. Samuel Wilberforce, in his 1860 review of *The Origin of Species*, argued that Darwin's explanations were lacking, as Darwin had no evidence of favourable mutations, yet such mutations are a crucial part of the natural selection mechanism[316]; this debate between Wilberforce and the Darwinians is an instance of a controversy regarding (1). He also argued that Darwin's reasoning involved leaps of fancy – that it was incompatible with "the true Baconian Philosophy" – which can be modelled as challenging (2) and (3)[317]. Insofar as the defenders of Darwin were successful in responding to Wilberforce and other critics, the evidential probability would be a higher interval value for Darwinism's probability. (The exact numbers would depend on the confidence level for such a statistical inference.) Since the value of $1/2[x, y]$ for Darwinism would be increased by adding the old evidence to its total evidence in any plausible formal model of the debate, the old evidence that Darwin cited would confirm his theory. Consequently, even if old evidence poses a problem for Bayesians, it is unproblematic for Evidential Probabilists.

[316] Wilberforce (1860) p. 238.

[317] Wilberforce (1860) p. 249.

Not all philosophers would agree that the Problem of Old Evidence is a problem. For example, Lakatos argues that only novel predictions can confirm a hypothesis; specifically, they can confirm what he calls a "research programme" that includes that hypothesis[318]. Alan Musgrave labels this position the "Strictly Temporal" view of evidence: if the evidence was a part of the background knowledge of science (as a whole, rather than an individual scientist's background knowledge) then the evidence cannot confirm any hypothesis[319]. More broadly, one might think that some types of old evidence will not be relevant for confirmation[320]. However, Bayesians can agree with philosophers such as Lakatos, and argue that the Problem of Old Evidence is a positive feature of their theory, rather than a bug.

This raises the issue of the logical relations between such theories, which I shall label using the term 'predicitivist', and (1) my proposed definition of confirmation and (2) Evidential Probabilism more broadly. Since evidential probabilists are not committed to any single formal definition of confirmation (or even the existence of a formal definition of this concept) these logical issues must be treated separately.

Firstly, predictivist theories are incompatible with the definition I offered in Subsection 2.2, since it is possible (as in the example of Darwin's theory of evolution given above) that old evidence can confirm a hypothesis on my definition. Secondly, predictivist theories and Evidential Probability are compatible. For instance, one could add an additional

[318] Lakatos (1978) p. 6.

[319] Musgrave (1974) p. 8.

[320] For example, Zahar (1973a) and Zahar (1973b). Such theories of evidence must be distinguished from the position of philosophers such as William Whewell (1857, p. 464) who argue that (in a particular sense of 'new') new evidence confirms more strongly than old evidence.

clause to my definition stating that 'E was not a generally accepted part of the body of scientific data prior to the formulation of H'. Such a definition would be significantly alter my definition, since the current form is ahistorical and this additional clause would make confirmation relative to a particular moment in the history of science. There are also a variety of alternative versions of such a clause; Martin Curd and J. A. Cover distinguish four distinct senses in which one might require that the evidence is 'new':

(1) Temporal Novelty: The evidence was not known to anyone prior to the proposal of the theory that it confirms.

(2) Epistemic Novelty: The evidence was not known to the person proposing the theory, nor was it generally known by scientists prior to proposal of the theory that it confirms.

(3) Design-Novelty: The evidence was not a factor in the scientist's development of the theory.

(4) Use-Novelty: The evidence was not used to decide the value of a parameter in the theory and the evidence is not built into the theory[321].

Since I accept both my definition of confirmation *and* Evidential Probability, my position is incompatible with predictivism. Given sufficient reasons to accept predictivism, one simple option for me would be to maintain Evidential Probability, but to adopt a

---

[321] Curd and Cover (1998) p. 512-513.

historically relative definition with a clause(s) that excludes old evidence. Additionally, the general spirit of predictivism (though not the letter) is compatible with my definition of confirmation, because my definition of confirmation is *qualitative*: it only states whether the evidence confirms a hypothesis relative to particular background knowledge, and that it is compatible with a range of theories about the *comparative* strength of confirmation. For example, I could consistently hold that new evidence confirms more strongly than old evidence, *ceteris paribus*. That would still not satisfy predictivists, who insist that only new evidence can confirm a hypothesis, but it *would* do some justice to some people's intuitions that there is something special about new evidence. However, since my fundamental aim in this chapter is to advocate an Evidential Probabilist *qualitative* definition of confirmation, I shall not explore these comparative issues further.

To summarise this subsection, my definition accommodates the view that old evidence can confirm a hypothesis; as a consequence, it is incompatible with predictivists' contention that only new evidence can confirm a hypothesis. However, Evidential Probability is compatible with this view. Thus, Evidential Probability as such does not imply that old evidence can confirm hypotheses, but it is compatible with this position.

3.1.4 The Representational Limits of Bayesian Probability

Norton also raises problems regarding the ability of the Bayesian formalism to represent certain types of epistemic states. There are two types of epistemic state that Norton discusses: (1) the problem of distinguishing between "disbelief" and "ignorance"[322] and (2)

[322] Norton (2011a) p. 407.

the problem of representing ignorance over a countable infinity of outcomes[323]. Regardless of whether Bayesians can formalise these distinctions, I shall argue that they pose no problem for an Evidential Probabilist.

I shall begin with (1). In a rational agent, 'disbelief' might be modelled as a low Evidential Probability. To analyse 'ignorance', there are actually two different senses that an Evidential Probabilist can distinguish. Firstly, we might have relevant evidence for a statement, but this relevant evidence cannot be used to form a precise probability. Adam Elga has a good example of such a situation[324]. Imagine that a stranger on the street walks in front of you and suddenly takes three objects from his bag: a travel-sized tube of toothpaste, an ordinary tube of toothpaste, and a live jellyfish. Let H be the hypothesis that the next object that is drawn will be a tube of toothpaste. Although you seem to have relevant evidence regarding H, it is not intuitive that your evidence provides any basis for a precise probability.

An Evidential Probabilist can represent this first sort of ignorance via the width of the evidential probabilities. In Elga's example, the probability that the stranger will take out a tube of toothpaste will be a very wide interval, which represents a state of deep ignorance regarding H. In contrast, if we are making a random selection from a bag with a known proportion of tubes of toothpaste, then the evidential probability will be precisely valued. Thus, the degree of imprecision of Evidential Probability intervals can be used to represent this form of ignorance. Such a formal analysis of ignorance is distinct from the Evidential Probabilist analysis of disbelief, because it involves the *precision* of the evidential probabilities rather than their *values*.

[323] Norton (2011a) p. 412-415.

[324] Elga (2010) p. 1.

Another kind of ignorance occurs when there is no relevant evidence for a statement. For example, suppose you know that X and Y correspond to genuine predicates in the language of a previously isolated Amazonian tribe, but you do not know the denotation of X and Y. Assume that you have no relevant evidence for the truth or falsehood of the statement 'Some X is Y'. (If this sort of occasion never occurs, then an Evidential Probabilist does not face a problem with representing this type of ignorance, because no-one does.) Using Evidential Probability, one can represent this second type of ignorance as the satisfaction of the following conditions:

(1) EP(H | K) = [0, 1]

(2) There is no reference class statement in K regarding H[325].

The first claim says that K does not provide the basis for any precision in your probability for H. The second says K has no potential reference class data that you might have been able to use for an interval other than [0, 1]. Collectively, they imply there is no relevant evidence in K regarding the probability of H.

Norton also raises the problem of representing ignorance over a countable infinity of outcomes. If there is a random variable that can take an infinite number of values given our total evidence, then an epistemic probability theory would ideally be able to (i) assign positive values to each hypothesis that the random variable will take one of these values and also (ii) assign a uniform distribution, because we have no reason to regard any particular

[325] See Chapter 1 Subsection 5.3 for a definition of a reference class statement.

value of the random variable as more likely than any other. No real-valued probability distribution can satisfy these *desiderata*. However, as I argued in the examples of the Steady State Cosmology and the Dome, it is possible to satisfy both (i) and (ii) by assigning [0, 1] as the probability for each of the outcomes. An Evidential Probabilist does not seem to face any special problems in representing ignorance over an infinite number of possibilities.

3.1.5 Summary

In most of my responses to these criticisms, there is a common pattern: challenges to Bayesianism have been developed; these challenges present particular sorts of contexts where Bayesianism seems to lead to counterintuitive results; and these challenges vanish once a probabilist abandons the Bayesian ambition of having a precise probability in all situations. If one regards Evidential Probabilism as a generalisation of Bayesianism to cases in which the relative frequency data does not warrant precise probabilities, then this pattern is not surprising. From an Evidential Probabilist perspective, Norton is correctly pointing out that precise probabilities are not always legitimate. Nonetheless, an Evidential Probabilist can reason like a Bayesian when they have suitably rich data, while obviating the problems that Norton raises.

3.2 The Reliability of Inductive Schemas

Norton uses the term “formal” to describe theories of induction that aim to use only formal relations that hold between the evidence and the hypothesis to analyse inductive reasoning[326]. In formal theories of induction, one only sues the formal relations between H and E when analysing an inductive argument. A very simple example of a 'formal' theory would be a naïve form hypothetico-deductivism: 'If H implies E, then E confirms H'. A much more sophisticated example of a formalist theory of induction (in Norton's sense) is Bayesianism. Kyburg's theory of induction is also formalist, because in this theory the rationality of an induction is purely a matter of the Evidential Probability relations.

Norton[327] raises the following question for formalists: how do we know that these formal schemas are reliable? Norton does not define what he means by "reliable", but the Oxford English Dictionary provides two potentially relevant definitions:

(1) "[Something] That may be relied upon."

(2) (In modern statistics) "[a method] that yields consistent results when repeated under identical conditions." [328]

Definition (1) is not helpful in this context, because 'reliable' in this sense applies if it is reasonable to rely on the method in question and this would simply be an open challenge for a justification of the formal schema. Naturally, it would not be a very interesting criticism. However, if Norton means something like (2), then he has raised an interesting problem. We

[326] Norton (2003) p. 649. His usage is different from my use of 'formal' in Chapter 4: a theory of induction, as I have defined 'induction' and as Norton seems to use the term, must be informal in the sense used in that chapter, because restrictions are placed on the predicate terms (they must be observational) in the evidence before an argument can be inductive. Stove (1965) makes this point regarding the informality of Hempel's theory of confirmation, because Hempel requires that the premises are observation statements.

[327] Norton (2003) p. 667.

[328] OED Third Edition, updated 2009, from www.oed.com accessed 29/08/2016.

would ideally like to know that there is a high relative frequency (or at least greater than 50%) of true results using our inductive rules, assuming our premises and background knowledge. Similarly, if we knew that using an inductive rule in a particular type of context would almost always lead us to falsity (even when the premises and our background knowledge are true) then there would be a strong *prima facie* case against using the rule. If Norton's use of 'reliability' refers to an inductive rule leading us to truth in a high frequency of cases, then he is raising an important challenge for a theory of induction.

For example, we know that enumerative induction is an unreliable rule when applied to the colour of swans. (Enumerative inductions are inferences from uniform samples to generalisations like 'All *x* are *y*' or 'Most *x* are *y*'.) Consider this enumerative inductive argument:

<u>A3</u>

(1) All of the swans in a 1,000-fold sample of members of a newly discovered species of Antarctic swan are blue.

Therefore, probably, (C) All swans of this species are blue.

Such inferences are famously unreliable, even when they involve an extrapolation from large sample sizes. Suppose that, as far as we know, A3 is no more reliable than past enumerative inductions about swans. Intuitively, we should be very reticent to infer from (1) to C. Such reticence can be justified by direct inference from a known statement about enumerative inductions in general (that they are unreliable when applied to swans) to *the particular* inductive inference that I have formalised in A3.

At its heart, Evidential Probability is a theory about direct inference, since evidential probabilities are ultimately based on direct inference; the logic of direct inference is crucial to Evidential Probabilism. One can use this logic to evaluate the reliability of inferences like A3. Since we know that swans tend to be non-uniform in their colours from the failure of previous enumerative inductions, the Evidential Probability of the conclusion of A3 will be very low given (1) and our background knowledge. This means that the Evidential Probabilist has a method for distinguishing inductions that are reliable given what we know from unreliable inductions.

Norton has a follow-up problem for those formal theories of induction that appeal to some *other* formal theory of induction to prove their validity: how do we know that this second formal theory of induction is reliable? As Norton points out, this question begins a regress problem for some formalists[329]. Consider a formalist who argues that her theory of *eliminative* induction is a reliable method by providing an *enumerative* inductive argument in favour of her theory. She might argue that the theory would have been successful in the past and accordingly it is likely to be successful in the future as well. If she justifies this enumerative induction by appealing to a formal theory of enumerative induction, then Norton correctly notes he can simply reformulate his question: what is her justification for thinking that *this* second formal system is reliable?

[329] Norton (2003) p. 667.

However, Evidential Probability is not just another formal system of induction: fundamentally, it is a system of deriving epistemic probabilities from statistical information by direct inference. Unlike the inductivist who does not rest her theory of induction on a non-inductive form of reasoning like direct inference, an Evidential Probabilist can answer to the question 'How do we know that this formal inductive schema is a reliable method in this context?' without an appeal to *another* formal inductive schema.

Nevertheless, Norton could raise a new question: how do Evidential Probabilists know that their system of direct inference is reliable? Yet, in contrast to 'Is this inductive rule reliable', it is not clear what this second question means. A logic of direct inference is how one can determine whether something is reliable given some information, but it is not clear what it means to say that such a logic is 'reliable'. Norton himself does not provide a justification of his own theory of direct inference, which he employs in his "material" theory of induction[330].

Until the question is clarified, it is not clear that either Norton or an Evidential Probabilist is obliged to answer it. According to the Evidential Probabilist, one can evaluate an inductive inference by using the relevant evidential probabilities, which are ultimately derived from statistical information by direct inference. According to Norton, one can evaluate an inductive inference by applying his theory of direct inference to the relevant background information, in order to assess the reliability of the inductive inference. There is a clear dialectical parallel, and if Norton can answer the question of the reliability of induction in this way, then he must allow the same use of direct inference by the Evidential Probabilist. The salient difference between both answers and inductive justifications of induction (which,

[330] Norton (2013) p. 673-674.

as Norton argues, suffer from a vicious regress) is that they appeal to direct inference, rather than induction.

Therefore, an Evidential Probabilist can answer Norton's first question without generating an obvious regress: to assess the reliability of a particular inductive argument, we need to examine the probability of its conclusion given its premises and our background information. It is worth noting that my discussion does not constitute an answer to problems like Hume's Problem of Induction. The question Norton raises is simply whether a particular inductive argument is reliable *given our background knowledge*. The broader epistemological question of how that background knowledge is acquired is beyond the scope of Norton's challenge. The salient points are that (a) Evidential Probabilists have a method for evaluating the reliability of an inductive argument and (b) this method does not seem to face a genuine regress problem.

3.3 Mill's Muddle

Norton argues that formalist theories of induction struggle to analyse demonstrative inductions, by presenting a particular example in which two (apparently) formally identical inductive arguments are clearly qualitatively different in the support that their premises provide to their conclusions[331]. I shall call this example 'Mill's Muddle', as Norton develops it from a passage in Mill[332]. Consider A4 and A5:

---

[331] Norton (2003) p. 649.

[332] Mill (1882) p. 228.

A4

(1) All tested pure samples of the element bismuth melt at 271° C.

Therefore, (C_1) All pure samples of the element bismuth melt at 271° C.

A5

(1) All tested pure samples of wax melt at 91° C.

Therefore, (C_2) All pure samples of wax melt at 91° C.

Norton points out that A4 is clearly a much better argument, relative to our background knowledge. Mill's Muddle is the following puzzle: *why* is A3 a better argument than A4? On the surface, the arguments are formally identical. A theory of induction must involve looking below the surface to justify our intuitive comparative judgement.

Kyburg uses Evidential Probability to analyse a similar example. Imagine that an organic chemist has developed a new compound NC[333]. Under normal laboratory conditions, NC is a crystalline compound. She conducts a laboratory test to discover NC's melting point by measuring a single sample a_1. On the surface, the reasoning is:

[333] Kyburg (1976) p. 194-196.

<u>A5</u>

(1) a_1 is a pure sample of NC and a_1 melted at x degrees.

Therefore, (C_3) All pure samples of NC melt at x degrees.

In fact, her reasoning is more sophisticated. Firstly, there are the implicit premises:

(IP1) NC is a crystalline compound.

(IP2) If y is a crystalline compound, then it has a uniform melting point.

Additionally, any actual measurement procedure in science has a margin of error. Her actual premise is:

(1´) a_1 is a pure sample of NC and a_1 melted at $x \pm \varepsilon$ degrees.

- where ε is the estimated margin of error for her measurement method.

Kyburg analyses this reasoning using evidential probabilities. If E is the conjunction (IP1 ^ IP2 ^ 1´), then E deductively implies C_3, so $EP(C_3 \mid E) = [1, 1]$. The scientist can rationally follow this chain of reasoning if the conjunction E is acceptable in this context: each conjunct must be acceptable and the conjunction as a whole must also be acceptable. Under those circumstances, she can accept that NC melts at $x \pm \varepsilon$ degrees. Accordingly, once her background information is fully articulated and incorporated into the formal model, an Evidential Probabilist can explain how she may reasonably make a general induction from a single sample.

Of course, Kyburg's analysis assumes that the scientist has fully accepted E. In practice, statements like 'If *y* is a crystalline compound, then it has a uniform melting point' are not typically accepted with certainty. Instead, they can become increasingly probable. In Subsection 2.1, I described how Kyburg models scientific knowledge as a hirearchy of sets of statements: at the top of the hirearchy are our most certain knowledge, whereas at the lower levels are the hypotheses that are acceptably probable given our evidence. What constitutes 'evidence' in this model will depend on our standards of acceptance in a particular context: a statement like 'If *y* is a crystalline compound, then it has a uniform melting point' might not be acceptable relative to an exacting standard that we might use to establish (C_3) as part of received scientific knowledge, but it might be sufficiently probable to be acceptable as a working assumption in a laboratory test of (C_3). As E implies H, if it has a high probability, then H will have at least as high a level of probability within the hirearchy of statements.

I shall now apply a similar Evidential Probabilist analysis to Mill's Muddle. In the case of A4, Norton points out that chemistry provides us with the implicit premises:

(IP3) Most known elements are non-allotropic.

(IP4) Non-allotropic elements have uniform melting points under laboratory conditions.

(IP5) Bismuth is an element[334].

We need to know if we can infer that bismuth has a uniform melting point. Suppose that the reference class that we have selected by the rules of Sharpening is the set of known elements and we formalise 'Most' as '> 50%'. In that context, relative to our background knowledge K, the evidential probability of the claim-

(IP6) Bismuth is non-allotropic.

- is EP(IP6 | K) = [>0.5, 1].

Suppose that, in this context, we are willing to accept a statement whose probability, relative to K, has a lower limit that is > 0.5, so that we can add (IP6) to the Evidential Corpus. An additional implicit premise is the following:

(IP7) Bismuth has a uniform melting point under laboratory conditions.

We might be able to obtain this premise (that is the validator of A5) from (IP4) and (IP6). In other words, we need to see if we can accept the conjunction that 'Non-allotropic elements have uniform melting points under laboratory conditions and Bismuth is non-allotropic.' If the probability of (IP4 ^ IP6) is sufficiently high, then it can be added to the Evidential Corpus, and therefore so can IP7, since (IP4 ^ IP6) implies (IP7).

---

[334] Norton (2003) p. 651.

Suppose that we have found that our implicit and explicit premises are acceptable in this context. We can now formulate A4 in an Evidential Probabilist representation of our reasoning in the bismuth case:

A6

(1) All tested samples of the element bismuth melt at 271° C [335].

(IP7) Bismuth is uniform with respect to its melting point under laboratory conditions

Therefore, (C_3) All samples of the element bismuth melt at 271° C under laboratory conditions.

- and since our accepted premises (1) and validator (IP7) jointly imply (C_3), we can accept (C_3) when using any standard of acceptance at which we accept both premises. Via such a representation, an Evidential Probabilist can explain why A4 is such a strong argument.

In contrast, if we try to repeat the same reasoning with A5, we know that wax is *not* uniform with respect to its melting point because we know that there have been different pieces of wax that have melted at different temperatures under controlled conditions. Consequently, the probability of A5's conclusion given our total evidence is [0, 0], since we

[335] Like Norton, I shall ignore the fact that actual measurement involves a margin of error.

know that wax does not have a uniform melting point of 91 degrees. An Evidential Probabilist can also explain our intuitions in this example, because Kyburg's theory always requires that we consider relevant background knowledge.

In this way, Evidential Probability can formalise the asymmetry between the bismuth induction and the wax induction. The former is akin to the crystalline compound example that Kyburg discusses, whereas the latter is an irrational inference, because our background knowledge contains the statistical information that not all wax melts at 91 degrees. This formal analysis requires the articulation of our relevant background knowledge, but Kyburg's system can express this information.

This response to Norton is not a bare appeal to background knowledge: my point is not *just* that we have asymmetric background knowledge between wax and bismuth. Evidential Probabilism *also* enables one to distinguish between (a) background knowledge that affects the evidential relations and (b) background knowledge that does not affect the evidential relations. Specifically, in order to alter the probability of the hypothesis given the total evidence, a piece of background knowledge K_x must exclude the probability without K_x via the rules of Sharpening, as described in Chapter 1 Subsection 5.3. Thus, my answer goes beyond merely saying that there is background knowledge that creates an asymmetry: an Evidential Probabilist analysis can identify and justify the significance of some particular piece of background knowledge. This analytic potential is an advantage over theories that do not provide for the detailed articulation and analysis of the import background knowledge.

However, there are theories (such as Bayesianism) which *do* allow for background knowledge to feature in the analysis of inductive reasoning. This raises the question as to whether Evidential Probabilism has any special advantages over such theories. There are two unusual (though not unique) features of Evidential Probability that are especially useful in this context. Firstly, as explained in Chapter 1 Subsection 5.3, evidential probabilities are unique, in the sense that relative to the same evidence and the same background knowledge, evidential probabilist scientists will agree on the probability of a hypothesis. Not all forms of Bayesianism have this uniqueness for all of their probabilities. This feature makes the theory attractive for those who think that scientific confirmation is independent of extra-evidential opinion: while Evidential Probability does not remove all scope for reasonable disagreement, it does narrow this scope. In a particular sense of 'objectivity', it allows for analyses of demonstrative induction (and induction more generally) that are more objective than alternatives.

Secondly, as I also discussed in Chapter 1 Subsection 5.3, Evidential Probabilists have an answer to the Problem of the Reference Class. When background knowledge enters into the formal modelling of inductive reasoning, it is crucial that this problem is answered: for instance, in the bismuth case, we must determine whether the relevant reference class data for the probability that a sample of bismuth is representative is (a) our data about elements, (b) our data about chemicals, (c) our data about laboratory samples, (d) our data about samples of bismuth, or (e) some other reference class. The Evidential Probabilist answer might be incorrect (it is doubtful that it is the last word on the Problem of the Reference Class) but it is still a systematic answer that fits many of intuitions about reference class selection. Such an answer is not easily available: Alan Hájek argues that the Problem of the Reference Class is a severe problem for many probability theories, though he does not discuss Kyburg's theory in particular[336]. I do not claim that either this answer, nor the objectivity feature, are unique features of Evidential Probability. Nonetheless, they are not features that every formalist theory can satisfactorily provide: for instance, Subjective Bayesians have no pretensions to such a strong sense of objectivity in their analyses, and if Hájek is correct, then a good answer to the Problem of the Reference Class is a rare treasure.

Norton argues that probabilistic readings of the direct inferences in the bismuth case are "contrived", because there are only approximately 100 known chemical elements[337]. (Currently, there are 118 discovered chemical elements.) However, he does not explain why the small size of the population in a direct inference is problematic for a probabilistic interpretation of the inference. For example, it seems uncontrived to offer a probabilistic reading of an argument that:

<u>A7</u>

(1) 99 out of 100 tickets in the lottery are red.

Therefore, probably, (C) A randomly drawn ticket in the lottery will be red.

Indeed, probabilistic reasoning is often illustrated using normal decks of cards and such a deck contains only 52 cards. A probabilistic analysis of card-game reasoning does not seem "contrived" at all. It is possible that Norton means that it would be contrived to have a probabilistic analysis that used *precise* probabilities to analyse 'most'. He is correct that

[336] Hájek (2007).

[337] Norton (2013) p. 674.

interpreting a particular use of ‘most’ as a precise value like 75% or 90% would be yet another case of the “spurious precision” of Bayesian theories of induction that both he and Kyburg consider objectionable. However, an *imprecise* probabilist analysis of terms like ‘most’, as I have used above, does not seem contrived: ‘most’ and ‘more than half’ seem roughly synonymous. Furthermore, if ‘most’ is being used in a stronger way in a particular context (like ‘more than 75%’) then this can also be formalised using imprecise probabilities. The Evidential Probabilist has the tools to help clarify puzzles like Mill’s Muddle.

CONCLUSION

Epistemic probabilities offer one way of developing a theory of the balance of evidence in science. In this chapter, I have argued that Evidential Probability can answer Norton’s criticism of such theories in the context of inductive reasoning. It can also analyse cases of demonstrative induction, which Norton regards as an important criterion of adequacy for theories of inductive reasoning. One important consequence of my arguments is that a probabilist can avoid some of the contemporary criticisms of probabilistic theories of induction by adopting Kyburg’s system. In particular, Bayesians who are troubled by the criticisms that I have considered might be attracted to Kyburg’s alternative (but still closely related) probabilistic theory of induction.

CONCLUSION AND FURTHER CONSIDERATIONS

I have applied Evidential Probability to several important problems involving the concepts of confirmation and decision. As I argued in Chapters 1, 2, and 3, Kyburg's theory offers the tools for a fresh look at some classic problems. In Chapter 4, I argued that a formalist can use his Evidential Probabilist model of scientific knowledge to answer Goodman's otherwise fearsome New Riddle of Induction. Evidential Probability also avoids some of the most prominent objections to probabilistic theories of confirmation and decision, as was seen in Chapter 5. It seems to be a particularly powerful system for the analysis of neglected aspects of evidence and rational choice.

Clearly, I think that Kyburg's theory offers much of value. However, there are differences between my own views and those of Kyburg. For example, the decision theory that I develop in Chapter 3 is relatively close to MEU theory. Furthermore, while we agree on most of the issues I have discussed in this thesis, we have disagreements about some other parts of the philosophy of science. For instance, I do not endorse Kyburg's conventionalist views on theory choice and the nature of scientific laws. My thesis offers grounds to think that Kyburg's theory of probability is useful, but I do not intend to encourage a complete acceptance of Kyburg's philosophy of science.

Although I have given reasons to be interested in the formalisation of the quantity of evidence, I have not provided or endorsed a full quantitative formalisation of this concept. Perhaps no satisfactory general formalisation is possible, but it is a natural area for further investigation. An auspicious approach might be to combine my qualitative account in Chapter

1 with a suitable information theory. Such an inquiry would require a detailed discussion of information theory, which I have not attempted in this thesis.

Evidential Probability and Bayesianism are rivals. Criticism is one way that rival philosophical theories can confront each other. Suppose that there is a topic C, for which there are rival theories A and B. Adherents to A can develop problems for B; adherents to B can return the favour by developing problems for A and seek to answer the challenges from adherents to A; this cycle can be repeated indefinitely, or at least until both A and B fall out of fashion. Such a dialectic can produce a great quantity of brilliant philosophical work: both the problems and solutions can be stunningly ingenious. However, there is a danger that such a dialectic can lead to a neglect of the original topic, C.

Rather than focus on criticising Bayesianism, it seems more promising for Evidential Probabilists to seek to develop new answers to topics of common interest. Like Bayesians, I am interested in developing better formal theories of confirmation and rational decision. Like Bayesians, I think that probability is a crucial concept for understanding evidential relevance and inductive reasoning. Like the early Carnap (a founding father of Bayesianism) and perhaps some modern Bayesians, I think that a purely formal theory of confirmation is an interesting and attainable goal. My thesis is a part of all these enterprises. It seems that Evidential Probabilists can be the best sort of rivals to Bayesianism by making positive contributions to the pursuit of common goals. I hope that this thesis is such a contribution.

There are a vast number of promising unexplored areas in confirmation theory and decision theory where an Evidential Probabilist can make useful investigations. For example, I have not stressed the fact that Evidential Probability is a member of the family of *logical* *interpretations of probability*. It is also very different from the other members of this family. Perhaps these differences offer the means by which a logicist can address some of the criticisms of this interpretation.

Similarly, the application of Evidential Probability to decision theory is almost entirely unexplored. (The principal exception is Chapter 8 of Kyburg's *Science and Reason*.) For instance, I argued in Chapter 3 that an increase in the quantity of relevant evidence can reduce the degree of uncertainty of our decisions. It would be interesting to investigate the significance of this formalism for the longstanding debate on the *value of evidence*.

Finally, my strategy to answering the New Riddle of Induction suggests that the reliability of evidence can be useful for addressing at least some types of general underdetermination problem. The reliability of evidence offers a dimension of evidential support that can supplement the balance of evidence. Its significance for other underdetermination problems is largely unexplored and a rich area for profitable inquiry.

The sophisticated formal study of scientific reasoning is a young area, but it has already had a tremendous impact on the philosophy of science. Above all else, my thesis indicates that there are still many unexplored areas and that Evidential Probability has much to offer this endeavour.

BIBLIOGRAPHY

Achinstein, P. (1978). Concepts of Evidence. *Mind 45* (345) p. 22-50.

Abrams, J. J. (2002). Solution to the Problem of Induction: Peirce, Apel, and Goodman on the Grue Paradox. *Transactions of the Charles S. Peirce Society 38* (4) p. 543-558.

Al-Najjar, N. I. and Weinstein, J. (2009). The Ambiguity Aversion Literature: A Critical Assessment. *Economics and Philosophy* 25 (3) p. 249-284.

Antony, L. M. (2004). Who's Afraid of Disjunctive Properties? *Philosophical Issues 13* (1). p. 1-21.

Armstrong, D. M. (1983). *What is a Law of Nature?* Cambridge: Cambridge University Press.

Bacchus, F.; Kyburg, H. E.; and Thalos, M. (1990). Against Conditionalization. *Synthese 85* (3) p. 475-506.

Bar-Hillel, M. (1982). Ideal Evidence, Relevance and Second-Order Probabilities. *Erkenntnis 17* (3) p. 273-290.

Boole, G. (1958). *An Investigation of the Laws of Thought on Which are Founded the Mathematical Theories of Logic and Probabilities*. New York: Dover Publications.

Bradley, Seamus, "Imprecise Probabilities", *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/sum2015/entries/imprecise-probabilities/=

Bradley, S. and Steele, K. (2016). Can Free Evidence Be Bad? Value of Information for the Imprecise Probabilist. *Philosophy of Science 83* (1) p. 1-28.

Burks, A. W. (1953). The Presupposition Theory of Induction. *Philosophy of Science 20* (3) p. 177-197.

Butler, J. (1736). *The Analogy of Religion, Natural and Revealed, to the Constitution and Course of Nature*. London: John and Paul Knapton.

Carnap, R. (1945). The Two Concepts of Probability: The Problem of Probability. *Philosophy and Phenomenological Research 5* (4) p. 513-532.

Carnap, R. (1947). On the Application of Inductive Logic. *Philosophy and Phenomenological Research 8* (1) p. 133-148.

Carnap, R. (1962). *The Logical Foundations of Probability*. Chicago: University of Chicago Press.

Carnap, R. (1968). On Rules of Acceptance. In Imre Lakatos (Ed.), *The Problem of Inductive Logic* (pp. 146-150) Amsterdam: North-Holland Publishing Company.

Couvalis, G. (1997). *The Philosophy of Science: Science and Objectivity*. London: SAGE.

Curd, M. and Cover, J. A. (1998) *Philosophy of Science: The Central Issues*. New York and London: W. W. Norton & Company.

Darwin, C. (1872). *The Origin of Species by Means of Natural Selection: or, The Preservation of Favoured Races in the Struggle for Life*. London: John Murray.

Davidson, B. and Pargetter, R. (1987). Guilt Beyond Reasonable Doubt. *Australasian Journal of Philosophy 65* (2) p. 182-187.

Dow. J. and Werlang, S. R. C. (1992). Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio. *Econometrica 60* (1) p. 197-204.

Eells, E. and Fitelson, B. (2000). Measuring Confirmation and Evidence. *The Journal of Philosophy 97* (5) p. 663-672.

Elga, A. (2010). Subjective Probabilities Should be Sharp. *Philosophers' Imprint 10* (5) p. 1-11.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics 75* (4) p. 643-696.

Fennell, D. and Cartwright, N. (2010). Does Roush show that evidence should be probable? *Synthese 175* (3). 289-310.

Fitelson, B. (2001). A Bayesian Account of Independent Evidence with Applications. *Philosophy of Science 68* (3) p. S123-S140.

Fitelson, B. (2008). Goodman's "New Riddle". *Journal of Philosophical Logic* 3*7* (6) p. 613-643.

Fox, C. R and Tversky, A. (1995). Ambiguity Aversion and Comparative Ignorance. *The Quarterly Journal of Economics 110* (3) p. 585-603.

Franklin, J. (1998). Two Caricatures I: Pascal's Wager. *International Journal for the Philosophy of Religion 44* (2) p. 109-114.

Franklin, J. (2001). Resurrecting Logical Probability. *Erkenntnis 55* (2) p. 277-305.

Franklin, J. (2006). Case comment – United States vs. Copeland, 369 F. Supp. 2d 275 (E.D.N.Y. 2005): quantification of the 'proof beyond reasonable doubt' standard. *Law, Probability and Risk 5* (2) p. 159-165.

Franklin, J. (2012). Discussion paper: how much of commonsense and legal reasoning is formalizable? A review of conceptual obstacles. *Law, Probability and Risk 11* (0) p. 225-245.

Friedman, K. (1973). Son of Grue: Simplicity vs. Entrenchment. *Noûs 7* (4) p. 366-378.

Gärdenfors, P. (1990). Belief Revision and Relevance. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1990 2* p. 349-365.

Gärdenfors, P. and Sahlin, N. (1982). Unreliable Probabilities, Risk Taking, and Decision Making. *Synthese 53* (2) p. 361-386.

Gemes, K. (2007). Irrelevance: Strengthening the Bayesian Requirements. *Synthese 157* (2) p. 161-166.

Giere, R. N. (1970). An Orthodox Statistical Resolution of the Paradox of Confirmation. *Philosophy of Science 37* (3) p. 354-362.

Good, I. J. (1985). Weight of Evidence: A Brief Survey. *Bayesian Statistics 2* p. 249-270.

Goodman, N. (1946). A Query on Confirmation. *The Journal of Philosophy 43* (14) p. 383-385.

Goodman, N. (1983) *Fact, Fiction and Forecast*. Cambridge, Massachusetts: Harvard University Press.

Haack, S. (2003). *Defending Science – Within Reason*. Amherst: Prometheus Books.

Haack, S. (1996). Reflections of a Critical Common-Sensist. *Transactions of the Charles S. Peirce Society 32* (3) p. 359-373.

Hammond, P. J. (1976). Changing Tastes and Coherent Dynamic Choice. *The Review of Economic Studies 43* (1) p. 159-173.

Hájek, Alan, "Interpretations of Probability", *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2012/entries/probability-interpret/>.

Hájek, A. (2007). The Reference Class Problem Is Your Problem Too. *Synthese 156* (3) p. 563-585.

Hájek, A. (2003). What Conditional Probability Could Not Be. *Synthese 137* (3) p. 273-323.

Hempel, C. (1943). A Purely Syntactical Definition of Confirmation. *The Journal of Symbolic Logic 8* (4). p. 122-143.

Hempel, C. (1945a). Studies in the Logic of Confirmation I. *Mind 54* (213) p. 1-26.

Hempel, C. (1945b). Studies in the Logic of Confirmation II. *Mind 54* (214) p. 97-121.

Hempel, C. (1960). Inductive Inconsistencies. *Synthese 12* (2) p. 439-469.

Hempel, C. (1965) *Aspects of Scientific Explanation and Other Essays in Philosophy of Science*. New York: Free Press; London: Collier-Macmillan.

Hempel, C. and Oppenheim, P. (1945). A Definition of "Degree of Confirmation". *Philosophy of Science 12* (2). p. 98-115.

Hindmoor, A. (2006). *Rational Choice*. Basingstoke: Palgrave Macmillan.

Hooker, C. A. (1968). Goodman, 'Grue' and Hempel. *Philosophy of Science 35* (3) p. 232-247.

Hooker, C. A. and Stove, D. C. (1968). Relevance and the Ravens. *The British Journal for the Philosophy of Science 18* (4) p. 305-315.

Horwich, P. (1982). *Probability and Evidence*. Cambridge: Cambridge University Press.

Hosiasson, J. (1931). Why Do we Prefer Probabilities Relative to Many Data? *Mind 40* (157) p. 23-26.

Howard, R. E. (2006). The Tower of the Elephant. In Stephen Jones (Ed.), *The Complete Chronicles of Conan* (pp. 79-100) London: Gollancz.

Howson, C. and Urbach, P. (1993). *Scientific Reasoning: The Bayesian Approach*. Chicago, Ill.: Open Court.

Huber, W. A. (2010). Ignorance Is Not Probability. *Risk Analysis 30* (3) p. 371-376.

Huemer, M. (2001). The Problem of Defeasible Justification. *Erkenntnis 54* (3) p. 375-397.

Jackson, F. (1994). Grue. In Douglas Stalker (Ed.), *Grue! : The New Riddle of Induction* (pp. 132-153) Chicago: Open Court.

Jaynes, E. T. (2003). *Probability Theory: The Logic of Science*. Cambridge: Cambridge University Press.

Jeffrey, R. C. (1960). Review of The Logic of Scientific Discovery. *Econometrica 28* (4) p. 925.

Jeffrey, R. C. (1965). *The Logic of Decision*. New York: McGraw-Hill.

Jeffrey, R. C. (1992). *Probability and the Art of Judgement*. Cambridge, New York: Cambridge University Press.

Joyce, J. M. (2005). How Probabilities Reflect Evidence. *Philosophical Perspectives 19* (1) p. 153-178.

Joyce, J. M. (2010). A Defence of Imprecise Credences in Inference and Decision Making. *Philosophical Perspectives 24* (1) p. 281-323.

Johnson, W. E. (1924). *Logic, Part III: The Logical Foundations of Science*. Cambridge: Cambridge University Press.

Keynes, J. M. (1921). *A Treatise on Probability*. London: Macmillan.

Knight, F. H. (2006). *Risk, Uncertainty and Profit*. Mineola, NY: Dover Publications.

Kreps, D. M. (1990). *A Course in Microeconomic Theory*. New York; London: Harvester Wheatsheaf.

Kyburg, H. E. (1961). *Probability and the Logic of Rational Belief*. Middletown, Conn.: Wesleyan University Press.

Kyburg, H. E. (1968). Bets and Beliefs. *American Philosophical Quarterly 5* (1) p. 54-63.

Kyburg, H. E. (1970). *Probability and Inductive Logic*. Toronto, Ontario: The Macmillan Company.

Kyburg, H. E. (1974). *The Logical Foundations of Statistical Inference*. Dordrecht: Reidel.

Kyburg, H. E. (1976). Local and Global Induction. In Radu J. Bogdan (Ed.), *Local Induction* (pp. 255-266) Dordrecht, Holland; Boston-USA: D. Reidel Publishing Company.

Kyburg, H. E. (1978). Subjective Probability: Criticisms, Reflections, and Problems. *Journal of Philosophical Logic 7* (1) p. 157-180.

Kyburg, H. E. (1980). Reviewed Work: The Probable and the Provable. by L. Jonathan Cohen. *Noûs 14* (4) p. 623-629.

Kyburg, H. E. (1990). *Science and Reason*. New York: Oxford University Press.

Kyburg, H. E. (1992). Getting Fancy with Probability. *Synthese 90* (2) p. 189-203.

Kyburg, H. E. (2007). Bayesian Inference with Evidential Probability. In William Harper and Gregory Wheeler (Eds.), *Probability and Inference: Essays in Honour of Henry E. Kyburg, Jr.* (pp. 281-296) London: College Publications.

Kyburg, H. E. and Teng, C. M. (2001). *Uncertain Inference*. Cambridge: Cambridge University Press.

Ladyman, J. (2002). *Understanding Philosophy of Science*. London: Routledge.

Lakatos, I. (1980). *The Methodology of Scientific Research Programmes*. (Eds. John Worrall and Gregory Currie.) Cambridge; New York: Cambridge University Press.

Laudan, L. (1981). *Science and Hypothesis*. Dordrecht; London: Reidel.

Levi, I. (2007). Probability Logic and Logical Probability. In William Harper and Gregory Wheeler (Eds.), *Probability and Inference: Essays in Honour of Henry E. Kyburg, Jr.* (pp. 255-266) London: College Publications.

Longino, Helen E. (1990). *Science as Social Knowledge: Values and Objectivity in Science*. Princeton: Princeton University Press.

Maher, P. (2010). Explication of Logical Probability. *Journal of Philosophical Logic 39* (6) p. 593-616.

McGrew, T., Alspector-Kelly, M. and Allhoff, F. (2009). *Philosophy of Science: An Historical Anthology*. West Sussex: Wiley-Blackwell.

Mill, J. S. (1882). *A System of Logic, Ratiocinative and Inductive*. New York: Harper & Brothers, Publishers.

Miller, R. W. (1995). The Norms of Reason. *The Philosophical Review 104* (2) p. 205-245.

Musgrave, A. (1974) Logical versus Historical Theories of Confirmation. *The British Journal for the Philosophy of Science 25* (1) p. 1-23.

Nance, D. A. (2016). *The Burdens of Proof: Discriminatory Power, Weight of Evidence, and Tenacity of Belief*. Cambridge: Cambridge University Press.

Neurath, O. (1973). Anti-Spengler. In Marie Neurath and Robert Cohen (Eds.), *Empiricism and Sociology.* (pp. 158-213)). Dordrecht: Reidel.

Nix, C. J. and Paris, J. B. (2007). A Note on Binary Inductive Logic. *Journal of Philosophical Logic 36* (6). p. 735-771.

Norton, J. D. (2003). A Material Theory of Induction. *Philosophy of Science 70* (4) p. 647-670.

Norton, J. D. (2007). Probability Disassembled. *The British Journal for the Philosophy of Science 58* (2) p. 141-171.

Norton, J. D. (2011). Challenges to Bayesian Confirmation Theory. In Prasanta S. Bandyopadhyay and Malcolm R. Forster (Eds.), *Philosophy of Statistics* (pp. 391-440) Oxford: Elsevier B. V.

Norton, J. D. (2013). A Material Dissolution to the Problem of Induction. *Synthese 191* (4) p. 671-690.

Nozick, R. (1993). *The Nature of Rationality*. Princeton, N.J.: Princeton University Press.

O’Donnell, R. M. (1989). *Keynes: Philosophy, Economics and Politics*. Basingstoke: Macmillan.

O’Donnell, R. (1992). Keynes's Weight of Argument and Popper's Paradox of Ideal Evidence. *Philosophy of Science 59* (1) p. 44-52.

Paris J. and Vencovská, A. (2015) *Pure Inductive Logic*. Cambridge: Cambridge University Press.

Pedersen, A. P. and Wheeler, G. (2014). Demystifying Dilation. *Erkenntnis* 79 (6) p. 1305-1342.

Peirce, C. S. (1932). The Probability of Induction. In Charles Hartshorne and Paul Weiss (Eds.), *Collected Papers of Charles Sanders Peirce Vol. II: Elements of Logic* (pp. 82-105) Cambridge, Mass.: Harvard University Press.

Pennock, R. T. (1998). Evidential Relevance and the Grue Paradox. *Kagaku Tetsugaku 31* (1). p. 101-119.

Popper, K. (1972). *Objective Knowledge: An Evolutionary Approach*. Oxford: Oxford University Press.

Popper, K. (1980). *The Logic of Scientific Discovery*. London: Hutchinson.

Reichenbach H. (1949). *The Theory of Probability.* Berkeley: University of California Press.

Reiss, J. (2013). *Philosophy of Economics: A Contemporary Introduction*. New York, NY: Routledge.

Reiss, J. (2014). What's Wrong With Our Theories of Evidence? *Theoria: An International Journal for Theory, History and Foundations of Science 29* (2) p. 283-306.

Resnik, M. D. (1987). *Choices: An Introduction to Decision Theory*. Minneapolis: University of Minnesota Press.

Rinard, S. (2013). Against Radical Credal Imprecision. *Thought 2* p. 1-9.

Rothbard, M. (2009). *Man, Economy, and State with Power and Market*. Auburn, Alabama: Ludwig von Mises Institute.

Roush, S. (2005). *Tracking Truth: Knowledge, Evidence, and Science*. Clarendon Press: Oxford University Press.

Rowbottom, D. P. (2007). The Insufficiency of the Dutch Book Argument. *Studia Logica 87* (1) p. 65-71.

Rowbottom, D. P. (2008). On the Proximity of the Logical and 'Objective Bayesian' Interpretations of Probability. *Erkenntnis 69* (3) p. 335-349.

Runde, J. (1990). Keynesian Uncertainty and the Weight of Arguments. *Economics and Philosophy 6* (2) p. 275-292.

Schmeidler, D. (1989). Subjective Probability and Expected Utility without Additivity. *Econometrica 57* (2) p. 571-587.

Seidenfeld, T. (2007). Forbidden Fruit: When Epistemological Probability May *Not* Take a Bite of the Bayesian Apple. In William Harper and Gregory Wheeler (Eds.), *Probability and Inference: Essays in Honour of Henry E. Kyburg, Jr.* (pp. 267-279) London: College Publications.

Settle, T. (1974). Induction and Probability Unfused. In Paul Arthur Schlipp (Ed.), *The Philosophy of Karl Popper* (pp. 697-749) La Salle (Ill.): Open Court.

Slote, M. A. (1967). Some Thoughts on Goodman's Riddle. *Analysis 27* (4) p. 128-132.

Steele, Katie and Stefánsson, H. Orri, "Decision Theory", *The Stanford Encyclopedia of Philosophy* (Winter 2015 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2015/entries/decision-theory/>.

Stove, D. C. (1965). On Logical Definitions of Confirmation. *The British Journal for the Philosophy of Science 16* (64). p. 265-272.

Stove, D. C. (1982). *Popper and After: Four Modern Irrationalists*. Oxford: Pergamon.

Stove, D. C. (1986). *The Rationality of Induction.* Oxford: Oxford University Press.

Talbott, William, "Bayesian Epistemology", *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2016/entries/epistemology-bayesian/>.

Van Fraassen. (1990). Figures in a Probability Landscape. In Michael Dunn and Anil Gupta (Eds.), *Truth or Consequences* (pp. 345-356) Dordrecht: Springer.

Vickers, John, "The Problem of Induction", *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2016/entries/induction-problem/>.

Von Neumann, J. V. and Morgenstern, O. (1953). *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.

Walley, P: *Statistical Reasoning With Imprecise Probabilities* (1991) London: Chapman & Hall.

Walley, P. (2000). Towards a Unified Theory of Imprecise Probability. *International Journal of Approximate Reasoning 24* (2-3) p. 125-148.

Weingartner, P. (1994). Can There be Reasons for Putting Limitations on Classical Logic? In Paul Humphreys (Ed.), *Patrick Suppes: Scientific Philosopher*, *Vol. 3* (pp. 89-124) Dordrecht: Kluwer Academic Publishers.

Weintraub, R. (2008). Scepticism about Induction. In John Greco (Ed.), *The Oxford Handbook of Scepticism* (pp. 129-148) Oxford: Oxford University Press.

Weir, A. (1995). Gruesome Perceptual Spaces. *Analysis 55* (1) p. 27-36.

Wheeler, G. and Williamson, J. (2011). Evidential Probability and Objective Bayesian Epistemology. In Prasanta S. Bandyopadhyay and Malcolm R. Forster (Eds.), *Philosophy of Statistics* (pp. 307-331) Oxford: Elsevier B. V.

Whewell, W. (1857). *History of the Inductive Sciences, Vol. II*. London: John W. Parker and Son.

White, R. (2010). Evidential Symmetry and Mushy Credence. In Tamar Szabó Gendler and John Hawthorne (Eds.), *Oxford Studies in Epistemology*) (pp. 161-186) Oxford: Clarendon Press.

Wilberforce, S. (1860). On the Origin of Species, by means of Natural Selection; or the Preservation of Favoured Races in the Struggle for Life. *Quarterly Review 108* (215) p. 225–264

Williams, D. C. (1947). *The Ground of Induction*. New York: Russell & Russell.

Williams, M. (2001). *Problems of Knowledge*. Oxford: Oxford University Press.

Williamson, J. (2007). Motivating Objective Bayesianism: From Empirical Constraints to Objective Probabilities. In William Harper and Gregory Wheeler (Eds.), *Probability and Inference* (pp. 151-179) London: College Publications.

Williamson, J. (2010). *In Defence of Objective Bayesianism.* Oxford: Oxford University Press.

Williamson, J. (2011). Objective Bayesianism, Bayesian conditionalisation and voluntarism. *Synthese 178* (1). p. 67-85.

Williamson, T. (1998). Conditionalizing on Knowledge. *The British Journal for the Philosophy of Science 49* (1). p. 89-121.

Worrall, J. (2010). For Universal Rules, Against Induction. *Philosophy of Science 77* (5) p. 740-753.

Wright, J. (1991). *Science and the Theory of Rationality.* Aldershot, Hants; Brookfield, Vt.: Avebury.

Zahar, E. (1973a). Why Did Einstein's Programme supersede Lorentz's? *The British Journal for the Philosophy of Science 24* (2). p. 95-123.

Zahar, E. (1973b). Why Did Einstein's Programme supersede Lorentz's? *The British Journal for the Philosophy of Science 24* (3). p. 223-262.