# DELIBERATIVE BAYESIANISM: ABDUCTION, REFLECTION, AND THE WEIGHT OF EVIDENCE

by


Lok C. Chan

Department of Philosophy
Duke University


Date: ______________________________

Approved:

______________________________
Kevin D. Hoover, Supervisor

______________________________
Vincent Conitzer

______________________________
Felipe De Brigard

______________________________
Carlotta Pavese

Dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Department of Philosophy
in the Graduate School of
Duke University

2018

# ABSTRACT

# DELIBERATIVE BAYESIANISM: ABDUCTION, REFLECTION, AND THE WEIGHT OF EVIDENCE

by

Lok C. Chan

Department of Philosophy
Duke University

Date: ____________________
Approved:

____________________
Kevin D. Hoover, Supervisor

____________________
Vincent Conitzer

____________________
Felipe De Brigard

____________________
Carlotta Pavese

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Philosophy in the Graduate School of Duke University

2018

# Abstract

In this dissertation, I defend the thesis that an epistemic judgment of probability must be interpreted against the background of the *context of inquiry* in which it is made: in the *abductive* context, judgments of probability are matters of *decision*, made strategically in service to the investigative goal of the inquirer; in *deduction*, probabilities are *derived* based on the premises chosen in abduction, in order to explicate the implied commitments the agent may incur from those decisions; during the *inductive* stage, the inquirer is expected to conduct her empirical investigation in a deliberate manner, in accordance with the assertions and decisions she made during abduction and deduction, collectively referred to as the *deliberative context*.

To develop my position, I set the stage by proposing a pragmatist reading of Bas van Fraassen's Reflection Principle and his *voluntarist* interpretation of assertions of degrees of beliefs as performative locutions to express the intention to undertake a proportional epistemic commitment. I argue for a refinement of this view that I call *deliberativism*, which introduces an abductive dimension in order to understand the normative force that regulates these epistemic judgments. I then elaborate on my claim that decisions made in the context of abduction have inferential repercussions on the validity of inductive inference. In particular, I situate deliberativism in the context of statistical inference by critically examining a problem in the literature called *optional stopping*, which is used to demonstrate how the experimenter's intention to stop can manipulate the statistical significance of the data. The last chapter explores the Pericean idea of deductive reasoning as the strategic interrogation of a provisionally chosen hypothesis by focusing on J. M. Keynes' notion of *the weight of evidence*.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Overview

In this dissertation, I defend the thesis that an epistemic judgment of probability must be interpreted against the background of the *context of inquiry* in which it is made: in the *abductive* context, judgments of probability are matters of *decision*, made strategically in service to the investigative goal of the inquirer; in *deduction*, probabilities are *derived* mathematically based on the premises chosen in abduction, in order to explicate the implied commitments the agent may incur from those decisions; during the *inductive* stage, the inquirer is expected to conduct her empirical investigation in a deliberate manner, in accordance with the assertions and decisions she made during abduction and deduction, collectively referred to as the *deliberative context*.

In order to explore this particular notion of *abduction* in the context of formal epistemology, I will critically examine the *Reflection Principle*, proposed by Bas van Fraassen, as a guiding principle of epistemic rationality. The principle, roughly, says that my degrees of beliefs *now* are rationally constrained by what I think my future degrees of beliefs *will be*. If you think tomorrow you will judge that your degree of belief for rain is, say, 0.5, it would be irrational, the principle says, to assert that your *current* probability for the same event is anything other than 0.5.

To demystify Reflection and illuminate its connection to abduction, I will devote chapter 2 for explaining van Fraassen's probabilist position called *voluntarism*. Basically, voluntarism rejects the Bayesian idea that judgments of probability are *descriptive* reports of the agent's psychological states. Instead, the voluntarist holds that to hold a belief, partial or otherwise, is a matter of making a commitment to stand by the belief in the future. It is, in a philosophically important sense, analogous to making a *promise*, which implies the promiser's commitment to act in accordance with the content of the promise in the future. Thus, judgments of probability, like making promises, are what philosophers of language called *speech acts*. To make the connection between Reflection and abduction, I will provide a *pragmatic* interpretation of voluntarism by appealing to Peirce's later formulation of the Pragmatic Maxim, according to which the practical difference implied by the acceptance of a belief is understood in terms of the contribution the belief makes to the agent's *deliberate conduct*.

In chapter 3, I will discuss this idea of deliberate conduct in the context of philosophy of statistics. One of the many disagreements between the Bayesians and frequentists is the relevance of the so-called "stopping rules", which designate when sampling will be stopped. Traditionally, Bayesian statisticians argue that stopping rules are problematic, because they take into considerations extra-statistical facts, specifically the agent's intention to stop. The thought is that Bayesian methods count facts about the experimenter's deliberations as being irrelevant, due to their adherence to the *Likelihood Principle*. I will argue that such an assumption is not correct: in fact, like frequentists, Bayesians can also manipulate the statistical result by changing their intentions to stop. I will propose that the flight from intentions is not warranted; instead, we need a philosophical account of how intentions are subjective to rational evaluation in terms of epistemic commitments and obligations. I will introduce the problem of stopping rule in a historical context, and also consider a potential objection to my argument based on a result regarding the value of information proven by Frank Ramsey and I. J. Good.

One improvement of Reflection I will propose is that the commitment one incurs from making an epistemic judgment ought to be understood not as an unwavering perseverance in one's belief, but a deliberately structured disposition with degrees of responsiveness to new experience. In other words, evidential weight is a measure of one's *dispositional commitment*: to put oneself under the obligation to revise one's belief under different evidential scenarios in a *self-controlled* manner. Roughly, what I mean is this: there are two senses in which my judgment that there is a 50/50 chance that it will rain tomorrow have repercussion on our conduct in light of future evidence. One is to say I have the obligation to hold steadfastly on the belief, even if I come to encounter evidence that suggests the contrary. On a rigid understanding of the idea that accepting a belief is a commitment, this might sound like the thing to do. The other, one that I think makes more sense in the context of inquiry, is this: I incur a particular *vulnerability* to evidence. For instance, my commitment may be such that though at this point my personal probability of rain is 0.5, I am willing to revise my opinion based on a small amount of evidence. This could be the commitment I ought to have if I had no reason to think one way or another, so that some evidence pointing to one way or another is sufficient for me to be swayed. On the other hand, my commitment in the probability of rain being 0.5 could be *weighty*: perhaps two meteorologists that I trust equally are each giving completely conflicting forecasts, such that I am in a state of ambivalence that cannot be easily resolved unless I find substantial evidence to disqualify one of my sources as reliable. Reflection, for the most part, does not address this. This will the be focus of chapter 4, in which I exploit the notion of the *weight of evidence* in explaining my proposal.

## 1.2 Historical Contexts and Motivations

How should probability constrain our beliefs? Can degrees of belief be rational? These questions are central to philosophical problems that concern me in this dissertation. The dynamic between the freedom of thought and the security of logical thinking occupies an important role in the disagreement between the major historical figures.

### 1.2.1 Two Empiricisms

Historically, empiricism is often associated with a flavor of foundationalism that holds that empirical evidence must in some sense be untainted - if experience is to serve as the objective foundation of knowledge, it must be unsullied by our attitudes, beliefs, and values. It is without a doubt motivated by a conception of rationality familiar to philosophers, because of Descartes' method of doubt:

> Reason now leads me to think that I should hold back my assent from opinions which are not completely certain and indubitable just as carefully as I do from those which are patently false.[1]

The idea is that it is irrational to hold unjustified opinions, and the empiricist response is to find a way in which our knowledge can be justified by some indubitable

1. René Descartes, *The Philosophical Writings of Descartes*, vol. 1 (Cambridge University Press, 1984), 17.

experience. Quine is arguably the last great empiricist in this tradition, even though it is not immediately obvious. In "Epistemology Naturalized", Quine describes what he takes to be 'conceptual' and 'doctrinal' cores of empiricist epistemology. The conceptual core of empiricism, according to Quine, is concerned with the explication of our concepts in terms of more "basic" concepts, which refers directly to our sense experience, and the doctrinal side aims to show that all of our justified beliefs can be inferentially traced to a set of foundational and self-justified propositions.[2] Even though Quine claimed to have jettisoned these traditional empiricist commitments, and replaced them with a naturalized epistemology, he never did abandon this empiricist project: his idea of epistemology naturalized is simply to replace the old-fashioned notion of "experience" with the physicalist equivalent that he deemed more scientifically respectable. In his last major work, "From Stimulus to Science", he still pursues the foundationalist project under the guise of naturalism by suggesting ways in which the basis of science can be reduced to the firing of neural receptors.[3]

Such an attitude also permeates through the works of Carnap, another great empiricist. *The Logical Structure of the World*, also known by its German title *Aufbau*, is often considered by Quine as the most thorough attempt to carry out the project of reductionism, as its aim is "a step-by-step derivation or 'construction' of all concepts from certain fundamental concepts..."[4] Carnap the phenomenal reductionist is Quine's favorite philosophical foil in virtually all of his work.

However, in the *Aufbau* there is also an empiricist instinct that is not antithetical

2. W. V. Quine, "Epistemology Naturalized," in *Ontological Relativity and Other Essays* (New York: Columbia University Press, 1969), 70–78.

3. W. V. Quine, *From Stimulus to Science* (Harvard University Press, 1997), 15–16.

4. Rudolf Carnap, *The Logical Structure of the World* (Open Court Class, 2003) 5.

to foundationalism but also emphasizes the role of decision, intention, and volition in our practice of knowledge. Carnap explains that there is nothing metaphysically necessary about his decision of choosing sense experience as the basis for all concepts—its appropriateness depends ultimately on the decider's "standpoint."[5]

In *The Logical Syntax of Language*, he codifies this normative stance as *the Principle of Tolerance*:

> *It is not our business to set up prohibitions, but to arrive at conventions... In logic, there are no morals.* Everyone is at liberty to build up his own logic, i.e., his own form of language, as he wishes. All that is required of him is that, if he wishes to discuss it, he must state his methods clearly...[6]

This stance was evidently prompted by the disputes between logicians and mathematicians regarding the "correct" language of logic. But the standard of correctness does not exist until a language is chosen, so in an important sense, trying to determine which language is better than another is ultimately futile. However, once we adopt the Principle of Tolerance, "before us lies the boundless ocean of unlimited possibilities."[7] In "Empiricism, Ontology, and Semantics," Carnap expresses this attitudes as the *semantical method*, so-called because it takes the role of linguistic framework as central to the analysis of philosophical problems.[8] In particular, Carnap argues that many philosophical questions needlessly arise due to confusions about questions internal to a framework and ones external to it. As soon as one recognizes that the choice of a framework is a pragmatic decision, Carnap argues, many philosophical problems, including the existence of the external world, simply dissolve.

5. Carnap, *The Logical Structure of the World* 94–95.

6. Rudolf Carnap, *The Logical Syntax of Language* (K. Paul, Trench, Trubner & Co., 1937), 51–51.

7. ibid., xv, 50.

8. Rudolf Carnap, *Empiricism, Semantic, and Ontology* (University of Chicago Press, 1947)

Carnap, I think, embodies two competing philosophical attitudes: one is the desire for an algorithmic ideal of rationality that allows only stepwise movement in the space of reasons, and the other is the recognition that rationality permits leaps over epistemic gaps through extra-logical elements such as decisions and perspectives. The former point of view demands clarity of thought, so that each inferential step is perfectly secure. The second attitude, on the other hands, recognizes that the contributions of volition can never be wholly eliminated from our epistemic practice. Our inquiries often have to begin with assumptions that cannot be justified by the current body of evidence, but we must take a leap of faith with the hope that its provisional acceptance could lead the the gain in knowledge or information.

The goal of this dissertation can be summarized as a philosophical exploration of the dynamic between these two styles of thinking within the context of probabilistic reasoning. My main contention is that both philosophical attitudes are needed in order to make sense of our rationality, and their interaction must be understood in the respective role they play in the different *contexts* of inquiry.

The notion of inquiry is an explicit reference to the philosophy of C. S. Peirce, whose view is an important basis of and inspiration for the position I aim to defend and develop here. Peirce, as I understand him, has a systematic understanding of the context-sensitivity of epistemic judgments. That is, the evaluation of an assertion or a belief cannot be detached from the context in which it occurs: an assertion of probability has different normative implications in the context of abduction, deduction, and induction. My proposal is that an assertion has no normative force unless it is underwritten by an epistemic practice that incentivizes the agent to stand by the obligations imposed upon her.

Neither Carnap nor empiricism play a explicit role in the position I endeavor to defend in this dissertation; however, the ineliminable volitional element in our empirical rationality is central to the core issues to be discussed in virtually every chapter. This is an especially important context to keep in mind in our discussion of van Fraassen, who makes explicit reference to Carnap's view on the role of decision in "Empiricism, Ontology, and Semantics"[9]. Van Fraassen's voluntarism, in particular, should be see as belonging to the same tradition of empiricism that I ascribed to Carnap earlier.

### 1.2.2 Keynes and Ramsey on Rational Degrees of Belief

Probability, in J. M. Keynes's view, is defined as a logical relation between a premise and a conclusion. Probability relations are logical, because this relation belongs to the same conceptual category as the entailment relation between the premises and conclusion in a deductive argument. Keynes says:

> Inasmuch as it is always assumed that we can sometimes judge directly that a conclusion follows from a premiss, it is no great extension of this assumption to suppose that we can sometimes recognise that a conclusion partially follows from, or stands in a relation of probability to, a premiss.[10]

Keynes' logical interpretation of probability has the advantage of providing a direct explanation of why probability is *normative*: we *should* reason in accordance with probability for the same reason that we should respect a deductive rule like *modus ponens*: the degrees of a person's partial belief should correspond to the degree to

9. Bas van Fraassen, "Against Naturalized Epistemology," in *On Quine*, ed. Paolo Leonardi and Macro Santambrogio (Cambridge University Press, 1995), 83—84

10. John Maynard Keynes, *A Treatise on Probability* (Macmillan & Co., Limited., 1921), 57.

which the premises render the conclusion probable, whereas in a deductive proof, one has to accept the conclusion as necessarily true, were the premises true. Hence Keynes talks about probability not as something we can attribute to a proposition, nor one's attitude of it, but a logical and objective property of an argument.[11]

As Frequentists often find little use for the idea of degrees of belief, it is often seen as a notion exclusive to subjective theories of probability; however Keynes' rational degrees of belief are concevied to be objective relations independent of the human mind. More important, in Keynes, we find the foundationalist attitude that he inherited from the logical atomism of Russell. Keynes accepts Russell's idea of knowledge by acquaintance: some propositions, such as "I have a sensation of yellow", are justified by virtue of being perceptually acquainted with it.[12] Furthermore, acquaintance yields not only knowledge of the senses, but also of logical relations:

> When we know something by argument this must be through direct acquaintance with some logical relation between the conclusion and the premiss. In all knowledge, therefore, there is some direct element; and logic can never be made purely mechanical. All it can do is so to arrange the reasoning that the logical relations, which have to be perceived directly, are made explicit and are of a simple kind.[13]

This can be a puzzling position, considering Keynes's goal is to develop a formal system of probability; however, If logical relations are not analyzable in terms of empirical features, but through an irreducible relation knowable to us only by perception or intuition, then why would we need a formal system? Keynes' answer is that every-

11. There is a technical consequence for this. If probability is a relation, it follows that it never makes sense to talk about the probability of an event without in relation to any other proposition, so for Keynes all probabilities are conditional probabilities.

12. Keynes, *A Treatise on Probability*, 11–12.

13. ibid., 14.

one's intuitive faculty is different—"the perceptions of some relations of probability may be outside the powers of some or all of us'—so we need principles to make these perception explicit and justified.[14]

One important principle is the *Principle of Indifference.* Laplace is well-known for articulating a version of it:

> When the probability of a simple event is unknown, one may suppose that it is equally likely to take on any value from zero to one... the probability of each of these hypotheses, given the observed event, is a fraction whose numerator is the probability of the event under this hypothesis, and whose denominator is the sum of similar probabilities under each of the hypotheses.[15]

In other words, when we are in complete ignorance regarding the outcome of the event, the probability of each possible outcome is:

$$\frac{1}{\text{\# of total possible hypotheses}}$$

Many have criticized this principle. Peirce vehemently rejects it, as he argues that in many cases, especially when the number of possible outcomes is ambiguous, using the principle will lead to contradictions.[16]A simple example would be to determine the probability of an unknown object, for example, a marble, having a certain color, say, red. Suppose I have no information about this marble, so I have no reason to think the marble is red or not.[17] Following the principle, it would seem that $P(R) = P(\neg R) = 0.5$. However, we are led to contradictions when we asks if the

14. Keynes, *A Treatise on Probability*, 18.

15. Pierre-Simon Laplace, *Philosophical Essay on Probabilities* (Spring, 1995), 20.

16. Charles S. Peirce, "The Probability of Induction," in *Writings of Charles S. Peirce: A Chronological Edition, Volume 3: 1872–1878* (Indiana University Press, 1986), section II

17. Assume that the surface of a marble can only have one color.

marble is yellow, black, etc.; because, using the same reasoning, we would say $P(Y) = P(\neg Y) = P(B) = P(\neg B) = 0.5$. The contradiction is that these are mutually exclusive propositions, so axioms of probability say that their sum cannot go beyond 1.

When Keynes wrote *A Treatise on Probability*, he was keenly aware of the paradoxes caused by the Principle of Indifference. However, he thinks that the paradoxes only suggest that the principle is to be restricted, not abandoned altogether. To begin, he notes that the crucial terms used in the principle are not clearly defined:

> The principle states that 'there must be no known reason for preferring one of a set of alternatives to any other.' What does this mean? What are 'reasons,' and how are we to know whether they do or do not justify us in preferring one alternative to another? I do not know any discussion of Probability in which this question has been so much as asked.[18]

Instead, he proposes that the ambiguous clause ought to be explicated in terms of *conditional relevance*:

> This distinction enables us to formulate the Principle of Indifference at any rate more precisely. There must be no relevant evidence relating to one alternative, unless there is corresponding evidence relating to the other; our relevant evidence, that is to say, must be symmetrical with regard to the alternatives, and must be applicable to each in the same manner. [19]

That is, to say that we have no reason to prefer a proposition $H$ over its alternatives is to say that there is no evidence such that it is relevant to $H$ but irrelevant to all the alternatives. Roughly speaking, $E$ is relevant to $H$ on background knowledge $K$ if and only if:

$$P(H|K) \neq P(H|K \wedge E)$$

18. Keynes, *A Treatise on Probability*, 58.
19. ibid., 61.

Thus, one necessary condition for a justified application of Indifference is that for a set of $n$ possible outcomes, $H_1, ...H_n$, there is no evidential proposition $E$ such that it is conditionally relevant to some but not others.

I will flag an important issue here. Keynes is not satisfied with this definition of relevance, because it fails to account for what he calls the *weight* of the evidence.[20] The basic idea is that some intuitively relevant evidence does not raise of lower the probability of the hypothesis it confirms. This will be discussed in chapter 4.

In any case, Keynes argues that the Principle of Indifference should not be used when the alternatives under consideration can be further analyzed, or, using his term, "divisible". Consider again the marble example. The paradox begins with the assumption that red and not red are the two alternatives. While it is true that these outcomes are exhaustive, "not red" should be analyzed before we distribute the probabilities, since it also encompasses the possibility that it is black, it is blue, etc. So according to Keynes' conditions, judging $P(R) = P(\neg R) = 0.5$ is unjustified, since it is not a legal application of the Principle of Indifference.

With Keynes' version of the principle, we have to know a great deal about the setup of the sample space, before even entertaining the possibility of indifference between alternatives. According to Keynes' proposal, this means that most of our intuitive judgments of indifference would be illicit. In fact, Keynes admits just as much: he holds that in many if not most scenarios, the probability of a given event cannot be given a precise value.[21] For many propositions, it would be impossible to say anything about their probability at all. Of course, Keynes is not saying that it is

20. Keynes, *A Treatise on Probability*, 79.

21. ibid.,41-42

psychologically impossible for us to have degrees of belief for a proposition that fails to satisfy these conditions, but he is saying that they would not be *rational* degrees of belief.

Keynes' epistemology behind his theory of probability presupposes a very restrictive view of how it means for a probabilistic judgment to be rational. The Principle of Indifference serves as a compelling *normative* principle that is required for any rational judgment of indifference. Because of Keynes' logical view of probability, to attribute equal probability between two outcomes without satisfying the principle would essentially the same as making an deductively invalid argument.[22] So, the principle serves as a standard of correctness by specifying the conditions under which the uniform distribution of probability among hypotheses is *rational*, i.e., justified. This is very different than saying that, whenever you have no information about the problem, you are allowed to uniformly distribute your degrees of belief. On Keynes' view, you should only do so when you have enough information to know the indifference is warrant; otherwise, you should just *withhold* your judgment.

In his "Truth and Probability", Ramsey responds to Keynes's view by rejecting the conception of degree of belief as perceptible logical relations:

> ...there really do not seem to be any such things as the probability relations [Keynes] describes. He supposes that, at any rate in certain cases, they can be perceived; but speaking for myself I feel confident that this is not true. I do not perceive them, and if I am to be persuaded that they exist it must be by argument; moreover I shrewdly suspect that others do not perceive them either, because they are able to come to so very little agreement as to which of them relates any two given propositions.

However, part of the appeal of Keynes' theory is that probability itself is irre-

22. Keynes, *A Treatise on Probability*, 60.

ducibly normative, since they are objective and logical. It would seem that denying the existence of these logical relations also robs probability of its normative force. Ramsey recognizes this, and explicitly rejects both the Principle of Indifference and Keynes' strongly normative conception of rational degrees of belief. Keynes's position, Ramsey argues, *requires* principles such as Indifference is that Keynes assumes that degrees of belief must be *justified* before it can be admitted in the calculus of probability; however, Ramsey thinks that such a demand cannot be satisfied:

> ...to ask what initial degrees of belief are justified... seems to me a meaningless question; and even if it had a meaning I do not see how it could be answered.[23]

Once this requirement of rationality is jettisoned,

> the Principle of Indifference can now be altogether dispensed with; we do not regard it as belonging to formal logic to say what should be a man's expectation of drawing a white or a black ball from an urn; his original expectations may within the limits of consistency be any he likes; all we have to point out is that if he has certain expectations he is bound in consistency to have certain others. This is simply bringing probability into line with ordinary formal logic, which does not criticize premisses but merely declares that certain conclusions are the only ones consistent with them.

Thus, Ramsey recognizes the normative nature of Keynes' use of the principle: it constrains our beliefs such that when the conditions for Indifference are met, one is required by rationality to assign equal probabilities to the outcomes. As the quote above makes clear, Ramsey's view makes no such demand. As far as he is concerned, it is not probability's business to tell people what degrees of belief they *should* have. Instead, the normative force of probability is the same as that of deductive logic: it serves as a tool for checking the *consistency* of beliefs, and no more.

23. Frank P. Ramsey, *Philosophical Papers* (Cambridge: Harvard University Press, 1990), 88.

Ramsey, then, is advocating what came to be the standard notion of Bayesian rationality: coherence. The question has shifted from "what are rational degrees of belief?" to "is this set of partial beliefs consistent?" Ramsey's analogy to deductive logic is helpful. Consider this argument:

**Premise 1** All Chinese are Martians.

**Premise 2** Socrates is Chinese.

**Conclusion** Socrates is Martian.

This is one of the first lessons every student from an introductory logic class would learn: even though the argument contains all false claims, from a deductive perspective, it is a perfectly valid argument. In the same way, if we have a definitely fair coin, yet the agent still decides that $P(Heads) = 0.999$, she is still rational as long as she also believes $P(Tails) = 0.001$. This normative point is perhaps the most influential among the ideas from "Truth and Probability," as Ramsey gestures toward the use of Dutch book arguments in support of coherence as a requirement of rationality: anyone who violates the axioms of probability "could have a book made against him by a cunning better and would then stand to lose in any event."[24] The argument is based on the assumption that the agent's willingness to bet is based her degrees of belief and utility function, so an inconsistent set of partial beliefs would imply contradictory betting behavior that could be exploited.

For instance, if the agent believe than $P(Heads) = 0.999$ yet also simultaneously holding that $P(Tails) = 0.5$. This means, based on the assumptions required by

24. Ramsey, *Philosophical Papers*, 78

Dutch book arguments, you would be willing to pay respectively \$0.999 and \$0.5 for a bet that pays \$1. So, if a bookie offers the agent both bets at those prices, which comes to the total of \$1.499, these bets should appear to be fair, based on her degrees of belief. But she is guaranteed to lose money from this deal, since landing on heads and on tails are mutually exclusive, she will win at most \$1, so she lose $1.499 - 1 = 0.499$ for sure. This is due to the fact that she violates the axiom of probability that says that the sum of the probability of each possible outcome has to be 1.[25]

I do not wish to dispute the effectiveness of Dutch book arguments here. In section 2.2, we will discuss one particular important use of the argument, but my thesis does not depend on it. For now, let us for the sake of argument assume that probabilistic coherence is a basic requirement of rationality. Still, there are aspects of probabilistic rationality unaccounted for in Ramsey's view, even if we grant the use of Dutch book arguments. Is coherence a sufficient substitute for a normative conception of degrees of belief? L. J. Savage, a strong advocate of Bayesianism and subjective probability, expresses his doubts about this:

> According to the personalistic view, the role of the mathematical theory of probability is to enable the person using it to detect inconsistencies in his own real or envisaged behavior. It is also understood that, having detected an inconsistency, he will remove it. An inconsistency is typically removable in many different ways, among which the theory gives no guidance for choosing.[26]

Savage's point can be illustrated by reconsidering the Dutch book case above: suppose the agent is convinced that her degrees of beliefs, $P(Heads) = 0.999$ and

25. Alan Hajek, "Dutch Book Arguments," in *The Oxford Handbook of Rational and Social Choice*, ed. Paul Anand, Prasanta Pattanaik, and Clemens Puppe (Oxford University Press, 2008), 176.

26. Leonard J. Savage, *The Foundations of Statistics* (Dover, 1954), 57.

$P(Tails) = 0.5$ are incoherent: what normative conclusion should she draw from this? The only advice coherence could give is that the sum should add up to 1, but should she lower $P(Heads)$ to 0.5 or lower $P(Tails)$ to 0.001? In fact, there are, quite literally, infinite ways to for her to resolve the inconsistency. This illuminates the lacuna left open by a coherentist conception of probabilistic rationality, motivated by Ramsey's rejection of Keynes' rational degrees of belief. In Keynes' framework, assuming that the conditions for the Principle of Indifference are satisfied, the *only* rational way to evaluate these probabilities is $P(Heads) = P(Tails) = 0.5$.

In his paper, Ramsey struggles with this issue: at the end he arrives at a somewhat vulgar pragmatism that says that an opinion is reasonable insofar as it works more often than not.[27] Keynes' assessment of his debate with Ramsey is telling:

> [According to Ramsey,] the basis of our degrees of belief... is part of our human outfit, perhaps given us merely by natural selection, analogous to our perception and our memories rather than to formal logic. So far I yield to Ramsey—I think he is right. But in attempting to distinguish "rational" degrees of belief from belief in general he was not yet, I think, quite successful. It is not getting to the bottom of the principle of induction merely to say that it is a useful mental habit[28]

Ramsey was evidently also dissatisfied with the view put forth in "Truth and Probability". Nevertheless, he wrote a note just before he died at the age of 26 that contains a relevant idea to the rest of the dissertation. In the note, he writes:

> The defect of my paper on probability was that it took partial belief as a psychological phenomenon to be defined and measured by a psychologist. But this sort of psychology goes a very little way and would be quite unacceptable in a developed science. In fact the notion of a belief of degree $^2/_3$ is useless to an outside observer, except when it is used by the thinker himself who says

27. Ramsey, *Philosophical Papers*, 93.

28. John Maynard Keynes, *Essays in Biography* (Macmillan & Co., Limited., 1933), 300–301.

> 'Well, I believe it to an extent $^2/_3$... Now what is the point of this numerical comparison? how is the number used?[29]

Ramsey seems to be expressing his frustration with a highly *descrptive* notion of degrees of belief—the elicitation procedure he describes in "Truth and Probability" was intended to measure the psychological state of the agent, but epistemology concerns itself not with discovering a thermometer of the mind, but with the *rationality* of belief. However, to be rational is a public matter, and the rationality of degrees of belief cannot be just about ensuring that the numbers in one's head is consistent.

The lesson I want to draw from the dispute between Keynes and Ramsey is this: one of the crucial questions regarding the rationality of partial belief is the rational status of our prior opinions. Keynes, at least in regard to probabilistic judgments, holds that degrees of belief can be rational if and only if they are justified by logically determined and objective priors. The result is a highly skeptical attitude toward the feasibility of having precise degrees of belief for most cases. Ramsey's plea for merely reasonable degrees of belief could be see as the rejection of such a requirement, but, as Keynes points out and Ramsey later recognizes, this proposal is inadequate unless we can find alternative ways to articulate the criteria of rationality for degrees of belief. We now turn to two opposite ways in which this agenda is further pursued.

### 1.2.3 Bayesianism and Voluntarism

In *Belief and the Will*, van Fraassen tries to offer a version of probabilism that is distinct from Bayesianism.[30] In the context of formal epistemology, *probabilism* can

29. Ramsey, *Philosophical Papers*, 95

30. Bas van Fraassen, "Belief and the Will," *Journal of Philosophy* 81, no. 5 (1984): 235–256

be characterized as the following two theses:[31]

1. The strength of a belief can be measured numerically as *degrees of belief.*
2. The rationality of degrees of belief is governed by the axioms of probability.

Even though 'probabilism' is often used as a synonymy for 'Bayesianism', van Fraassen makes a subtle distinction between the two:

> So: I am a probabilist, though not a Bayesian. Like the Bayesian I hold that rational persons with the same evidence can still disagree in their opinion generally; but I do not accept the Bayesian recipes for opinion change as rationally compelling. I do accept the Bayesian extension of the canons of logic to all forms of opinion and opinion change.[32]

Evidently, he is taking "probabilism" to be a broader term than "Bayesianism". In particular, Bayesianism addresses the following issues on top of probabilism:

1. The rationality of our *existing* opinions and,
2. The rational procedure to *revise* these opinions.

What van Fraassen calls *Orthodox* Bayesians rejects that existing degrees of belief must be justified in order to be rational. Any coherent set of partial beliefs are admissible. From an epistemological perspective, this is the first piece of the puzzle for the Bayesian response to skepticism, for the skeptic begins, as noted in the quote from Descartes above, by assuming that it is irrational to hold any belief that is not justified by logic or evidence.

---

31. Lina Eriksson and Alan Hajek, "What Are Degrees of Belief?," *Studia Logica* 86, nos. 183-213 (2007), 183

32. Bas van Fraassen, *Laws and Symmetry* (Oxford University Press, 1989), 175.

Orthodox Bayesians address the second issue above by holding that the only justifiable way to revise one's opinion is through Bayes' theorem.[33] This is often codified as the so-called rule of

> *Conditionalization*: one is rationally compelled to update one's prior degree of belief for $H$ in light of the acquisition of relevant evidence $E$ via the application of Bayes' theorem, which determines the posterior opinion degree of belief $P(H|E)$.[34]

Using Peirce's terminology, which van Fraassen adopts, conditionalization is an *explicative* procedure, which does not go beyond what's implied by facts and logic, as opposed to be being *ampliative*, which extrapolates beyond them.[35] We already saw the same idea from Ramsey, who sees the normative role of probability as the logic of consistency between partial beliefs. Orthodox Bayesianism, however, further legislates the rational revision of belief by requiring conditionalization as the only rational response to evidence. This presumably fills the normative gap pointed out by Savage and Ramsey. That is, instead of Ramsey's "useful mental habits", Orthodox Bayesians regard the ideal Bayesian agent, who revises her belief by following conditionalization, as the standard of rationality. As van Fraassen engagingly puts it, this purely explicative conception of belief revision allows the Orthodox Bayesian "to live a happy and useful life by conscientiously updating the opinions gained at

33. To be more precise, we should also distinguish *epistemic* Bayesians and *statistical* Bayesians. In this paper, I follow van Fraassen in using the term 'Orthodox Bayesian' to describe an epistemological position that holds a strict view on belief revision just described. A Bayesian *statistician* might think that belief do come in degrees and that Bayesian statistics is the best framework for drawing statistical inferences, but she might not think that all beliefs can be meaningfully given a Bayesian analysis.

34. Bas van Fraassen, "Belief and the Problem of Ulysses and the Sirens," *Philosophical Studies* 77, no. 1 (1995): 7–37,17.

35. Peirce, "The Probability of Induction," 297

his mother's knees, in response to his own experience."[36] This is assuming that conditionalization will allow a set of initially diverse priors to eventually converge into the same posterior.

Van Fraassen, however, rejects the idea that the ideal Bayesian agent is the *only* model of rationality. In particular, he insists that rationality cannot be wholly reduced to following an explicative rule. The result is an expanded notion of rationality. This alternative conception of rationality maintains that:

> what is rational to believe includes anything that one is not rationally compelled to disbelieve. And similarly for ways of change: the rational ways to change your opinions include any that remain within the bounds of rationality—which may be very wide. *Rationality is only bridled irrationality.*[37]

Van Fraassen calls his position *voluntarism*, so let us call the voluntaristic conception of rationality. It is voluntaristic in the sense that an agent is free to adopt any opinion that is "within the bounds of rationality".

At a glance, van Fraassen does not seem to be offering anything different than Ramsey's coherentism. It sounds as though we are simply returning the subjectivist idea that priors do not have to be justified. But van Fraassen's view is actually subtler than that, and to see this we must understand the pragmatist root voluntarism has.

Van Fraassen does not claim to be a pragmatist, but he credits James as the originator of his voluntarism, which, in its most naive formulation, says that the acceptance of belief is a matter of the will, that is, to believe a proposition is to make a *decision* to believe the proposition. Unsurprisingly, this is traditionally a theological position, which has its root in Pascal's Wager: the belief in God, Pascal

36. van Fraassen, *Laws and Symmetry*, 178.

37. Fraassen, 171–172.

tells us, cannot be resolved by the appeal to evidence or reason; instead, it is more akin to making a decision based on expected utility. Because we gain infinite utility from believing in God correctly and losing very little otherwise, the argument goes, the rational thing is to decide that God exists.[38]

| | God exists | God does not exist |
|---|---|---|
| Believer | $\infty$ | 0 |
| Unbeliever | $-\infty$ | 0 |

**Table 1.1**: Pascal's Wager

To be clear, my intention is not to deal with the theological issues here, but to carry out the difficult task at hand that is to explain how van Fraassen puts forth voluntarism as a general epistemological position about degrees of belief. To go from Pascal to van Fraassen, we must understand how James develops his voluntarism in response to Clifford.

Clifford, at least as characterized by James, is arguing for a standard of rationality on which the acceptance of a proposition without sufficient evidence is irrational, *even if the proposition were true.* Clifford is clear that this epistemic duty is categorical: "It is wrong always, everywhere, and for every one, to believe anything upon insufficient evidence."[39] To van Fraassen, this is paradigmatic example of a compulsory conception of rationality.[40] According to this picture of rationality, what is rational to believe for a person is restricted to those for she has justification, which could be in the form of evidence or logical necessity. This, without a doubt, is motivated by same intuition about rationality that motivated the traditional empiricists.

38. *Pascal's Pensees*, `https://www.gutenberg.org/files/18269/18269-h/18269-h.htm`, Accessed: 8-21-2018, sec. 233.

39. William James, *The Will to Believe and Human Immortality* (Dover Publications, 1960), 8.

40. van Fraassen, *Laws and Symmetry*, 171.

James argues against Clifford's view by introducing epistemic values into the discussion: James suggests that we are pulled in two directions in the formation of our opinions:

> There are two ways of looking at our duty in the matter of opinion... *We must know the truth; and we must avoid error*,— these are our first and great commandments as would-be knowers; but they are not two ways of stating an identical commandment, they are two separable laws.[41]

Why are they separate? Consider two types of epistemic agents: a greedy truth-seeker and a unmovable skeptic. The former is driven by nothing but the hunger for information, while the latter would rather accepts nothing that is short of being absolutely certain. To satisfy their respective values, their epistemic policies would be quite different: the truth-seeker should maximize her true beliefs by believing in everything, including contradictions and other falsehood. A skeptic can avoid all errors by refusing to believe in anything. So, not only are the desire for truth and the aversion to errors two separate epistemic values, they lead to epistemic policies that are diametrically opposed to each other.

These extreme epistemic policies seem absurd to us, because, we regard both the attainment of truth and the avoidance of error as being valuable. Satisfying one, however, often undermines another, due to our limited resources: if we had infinite cognitive power, memory, and time, we could perhaps learn in a way that guarantees the accuracy of our information. But this is not what our epistemic life is like: we lack the resource to fully fulfill these competing concerns: reaching for the truth often means opening oneself to the risk of error, and to be cautious against believing falsehood often lowers one's chance of the truth. As a result, we are forced to find a

41. James, *The Will to Believe and Human Immortality*, 17.

measure of balance.

From an argumentative point of view, the purpose of James' introduction of epistemic values is to frame Clifford's position in a new light. That is, Clifford's call to suspend one's judgment is motivated by the *decision* to take priority in the avoidance of error over the the desire for truth.

> "he who says 'Better go without belief forever than believe a lie!' merely shows his own preponderant private horror of becoming a dupe. He may be critical of many of his desires and fears, but this fear he slavishly obeys[42]

To agree with Clifford, then, one must decide that the price of the security of skepticism is missing out on a chance in receiving the truth—this is especially prominent in Clifford's insistence that it is better to suspend judgment to accidentally come into the possession of a true belief. But in some context, this seems hardly the rational course of action:

> It is like a general informing his soldiers that it is better to keep out of battle forever than to risk a single wound.[43]

What James is arguing, then, is that Clifford's position presupposes that the avoiding the risk of error is *always* more important than knowing the truth. James argues that this cannot be right, since different questions have different degrees of *urgency*. James argues that the belief in God is an urgent decision that is imposed upon us, for we cannot wait for that an airtight argument for the existence of God. This is where James makes the Pascalian argument:

> You must either believe or not believe that God is—which will you do? Your human reason cannot say. A game is going on between you and the nature of

42. James, *The Will to Believe and Human Immortality*, 18.

43. ibid., 19.

> things which at the day of judgment will bring out either heads or tails. Weigh what your gains and your losses would be if you should stake all you have on heads, or God's existence: if you win in such case, you gain eternal beatitude; if you lose, you lose nothing at all.[44]

His contention is that in such a situation, one has the right to accept these beliefs, as long as they are committed to make the practical difference implied by the acceptance of these beliefs.

It is clear that van Fraassen aims to create a contrast between James and Clifford in service of his alternative to Orthodox Bayesianism. In James' view, we find a possibility for a more permissive notion of rationality that makes room for leaps of faith that are not possible under the algorithmic view of rationality depicted in Orthodox Bayesianism. Van Fraassen's appeal to James is not superficial. In fact, it's not merely analogous: he holds that our epistemic rationality is for the most part governed by our willing to subject ourselves to the commitment to act in accordance with our understanding of the normative implication of the belief itself. This is less outlandish than it sounds, especially if we make the connection to the probabilist assumption that to believe to degree $x\%$ is to commit oneself to the intention to accept bets in accordance with that degree. This interpretation of voluntarism dovetails the standard interpretation of subjective probability as the willingness to bet. Clearly, more to be said about the nature of this commitment, which is the focus of chapter 2. How epistemic commitments bear on our experimental practice will be examined in 3. Just to anticipate, I think voluntarism is correct but must be contextualized: I accept the voluntaristic conception of rationality *in the context that the freedom of thought is prioritized in service of the goal of the inquiry.*

44. James, *The Will to Believe and Human Immortality*, 5

An implication in James' voluntarism that is less exploited by van Fraassen, but will be thoroughly discussed in this dissertation is the opening to the door to the understanding of the role *value* plays in epistemic judgments. To put James' view in a less theological and more epistemological context, the need to consider how different evaluative concerns are to be balance is often required in statistical reasoning: to structure the experimental space, the investigator often has to balance between the probabilities of false positives and false negatives. If we have infinite time and money, such a balance act would not be necessary, but our epistemic practice should not be constrained by impossible idealization. This is the topic of chapter 4.

# Chapter 2

# Reflection Principle and the Pragmatic Maxim

Should my current opinions be constrained by what I expect myself to belief in the future? This is the concern addressed by the principle of Reflection, which answers affirmatively:

> *General Reflection Principle.* My current opinion about event $E$ must lie in the range spanned by the possible opinions I may come to have about $E$ at later time $t$, as far as my opinion is concerned.[1]

In the context of probabilistic judgment, Reflection implies that your current credence, i.e., subjective probability, for the proposition $K$ now at $t1$ must be one of the values you consider as possible in the future at $t2$. Van Frassen formulates this general version of the principle in order to accommodate imprecise probabilities and vague opinions. I put leave issues regarding imprecise probabilities aside: when talking about probability, in our context it is sufficient to a version of Reflection that presupposes precise probability:

> *Special Reflection Principle.* $P_t(E|p_{t+x}(E) = r) = r$

In words, this formulation says: your subject probability for $E$ currently at $t$, given in the future at time $t + x$ it will be $r$, should also be $r$. For example, if tomorrow you

1. van Fraassen, "Belief and the Problem of Ulysses and the Sirens," 16.

will come to believe that the probability of rain is 0.5, then your current probability of rain should also be 0.5.

Why should what I think I will believe in the future constrain my current opinion? The main purpose of this chapter is to demystify Reflection. In particular, I shall argue for the novel view that Reflection is a principle of *abductive* reasoning: it regulates the rationality of the epistemic judgments made in the context of abduction. My contention is that Reflection is a guiding principle of what Peirce calls the rationality of deliberate conduct. One criticism of voluntarism I will focus on is that it does not recognize the importance of the *context sensitivity* of epistemic judgment.

## 2.1 Moore's Paradox and Self-Sabotaging

In section 1.2.3, I have presented a contrast between Orthodox Bayesianism and van Fraassen's voluntarism: conditionalization provides a rational constraint that is explicative in nature, while voluntarism allows ampliative extrapolation beyond the implication of the evidence. Still, even if voluntarism presupposes a more liberal conception of rationality, it must impose least *some* constraints on our beliefs. This is what the Reflection Principle is supposed to do; it aims to be a lenient replacement for conditionalization as the overarching principle of rationality; however, what how we conceive our future selves has to do with rationality is still rather mysterious. To understand this, we have to consider the implication of violating Reflection. The possibility of being Dutch-booked is part of it, but not the full story. The crucial idea is how the violation of Reflection can lead to the so-called "Moore's Paradox."

To begin, we have to distinguish a Moorean *absurdity* and Moore's *paradox*. A

Moorean absurdity arises when a person utters or thinks that '$P$ and I don't believe that $P$'. For example,

$$\text{It's raining but I don't believe that it is.} \tag{2.1}$$

Logically speaking, an utterance or thought like 2.1 is not contradictory: There is no logical connection between my belief in the rain and whether it is actually raining. This can be demonstrated easily by rephrasing the same proposition from a third person point of view:

$$\text{It's raining but Lok does not believe that it is.} \tag{2.2}$$

Unlike 2.1, 2.2 could easily be a statement about a mistaken judgment on my part. Perhaps last I checked the sky is absolutely clear, so I refused to believe my friend's truthful report that it's currently raining. Moore's *paradox* is the thought that, even though logic tells us that such a proposition is perfectly consistent, it seems absurd for anyone to *assert* such the first-person-perspective version of the proposition.[2] The source of absurdity is from the conjunction of someone asserting that it is raining, and the disavowal of the belief in what she just asserts. Another way to phrase the paradox without appealing to assertion is to say that a Moorean absurdity is a proposition that is logically consistent but not believable. That is, I cannot, without succumbing to absurdity, attribute to myself the belief that it's P and I don't believe that P.

What does the possibility of Moorean absurdity suggest? Consider this proposal:

2. Green Mitchell S. and Williams John N., *Moore's Paradox: New Essays on Belief, Rationality, and the First Person* (Oxford University Press, 2007), 190.

it is the result of a violation of established norms. That is, there are implicit standard that govern the norms between asserting $P$ and believing $P$, such that one is expected to believe what she asserts. Of course, this is not to say that people do not make deceptive assertion. In such a case the norms are being exploited for various reasons, but in those cases the speakers do not announce their intention to deceive. The absurdity comes from the fact that the norms are being explicitly broken.

One explanation for Moore's paradox is that it signifies a violation of the underlying norms that govern a *speech act*. The idea is that many of our linguistic practices carry non-linguistic effects beyond what's being said. Consider a classic example of a speech act: making a promise. Speech acts theorists argue that there is a normative link between uttering "I promise that $p$" and the intention to bring about $p$ in the future. A promise could be deemed infelicitous, when a speaker fail to fulfill the normative conditions necessary for the act of promising to be successful. Consider J. L. Austin's example: [3]

$$\text{I promise to do } X \text{ but I do not intend to do it} \tag{2.3}$$

Like 2.1, this proposition appear absurd, because they violate the implicit norm that when one makes a promise, she is expected to have the sincere intention to fulfill the said promise. More important, the expression of the intention implies that the promisor is willingly placing oneself under an obligation to the promisee.[4] So to make the promise, while simultaneously expressing the lack of intention to carry it out is

3. J. L. Austin, *How to Do Things with Words* (Clarendon Press, 1962), 54.

4. John R. Searle, *Speech Acts: An Essay in the Philosophy of Language* (Cambridge University Press, 1969), 60.

an absurd act of *self-sabotaging*.

In section 1.2.3, we noted that, according to the voluntaristic conception of rationality, any belief that is with the "bounds of reason" is rational. Van Fraassen does not clearly define what those bounds are, but he does spell out some specific conditions. One is that self-sabotaging is always irrational, which depends on a distinction between *ex post* and *ex ante* notions of rational evaluation:

> Any act of decision can be evaluated in two ways. if we evaluate it beforehand, we ask how *reasonable* it is, and, afterward, we ask to what extent it was *vindicated*... Therefore a minimal criterion of reasonableness is that *you should not sabotage your possibilities of vindication beforehand*[5]

For instance, a promise could be unreasonable—it would unreasonable for the promiser to make the promise—if the action required to fulfill the promise is impossible, since this means the promiser will never be vindicated. On the other hand, people do make insincere promises, and sometimes that would be the rational thing to do: I could be vindicated in making a insincere promise, if it turned out that doing saved my life; however, my promise would be absurd, if I were to make an insincere promise *and* announce my insincerity.

Van Fraassen's contention is that an assertion of probabilistic judgment, for example:

$$\text{It seems more likely to me that it will snow than that it will rain.} \tag{2.4}$$

is more like the speech act of making a promise than a description of the asserter's psychological state.[6] To begin, we can consider what it would mean for a probabilistic

5. van Fraassen, *Laws and Symmetry*, 157.

6. van Fraassen, "Belief and the Will" 252–255.

judgment to be a statement of fact. Ramsey's operationalist definition, as discussed in the last chapter, seems to fit the bill: an autobiographical report of one's degrees of belief, elicited through a combination of neutral propositions and utility evaluation, is a causal effect of the reporter's disposition to act, so a probabilistic judgment is interpreted as a description of the agent's psychological state, much like reports of an object's responses to being scratched by different substances are descriptions of the object's hardness.

In contrast, consider a passage, which van Fraassen cites approvingly, from Bruno De Finetti, another progenitor of modern Bayesianism:

> Any assertion concerning probabilities of events is merely expressing of somebody's opinion and not itself an event. There is no meaning, therefore, in asking whether such an assertion is true or false or less probable.
> The situation is different, of course, if we are concerned not with the assertion itself but with whether "somebody holds or expresses such an opinion or acts according to it"; for this is a real event or proposition.[7]

De Finetti's distinction can be phrased within the framework of speech acts theory: it is a statement of fact that Lok Chan made a promise to do $X$ today at 5PM. It would be true if I did make such a promise, but it makes no sense to ask if the promise itself is true. I can make a successful promise by clearly expressing my intention to fulfill my obligation. De Finetti appears to be saying something quite similar: an assertion is an *expression* of opinion that itself cannot be true or false.

In agreement with De Finetti, Van Fraassen argues that making a probabilistic judgment is more like making a promise and reporting an autobiographic report about one's psychological state. To make this argument, van Fraassen tries to demonstrate

7. Bruno de Finetti, *Probability, Induction, and Statistics* (Wiley, 1972), 189.

that this statement

$$\text{my current degree of belief for } A \text{ is } x \text{ but it will be } y \neq x. \tag{2.5}$$

to be an instance of a Moorean absurdity. More specifically, anyone who asserts the above is said to be *self-sabotaging*. It is evidently much less clear that the violation of Reflection is an absurdity, compared to making a promise while announcing the lack of intention to keep it.

Perhaps we can first consider the violation in case of a full belief. Suppose I am invited to the flat earth society meeting tomorrow that features an extremely persuasive speaker. Knowing that I am always easily swayed by impressive rhetoric, I assert that

$$\text{The earth is spherical but I expect to believe in a flat earth tomorrow.} \tag{2.6}$$

Am I sabotaging myself? The argument could go: if I am willing to assert fully that the earth is spherical now, I should be able to stand by my assertion in the future. If I consider my belief in a spherical earth a rational one and regard being swayed by rhetorics but not evidence to be irrational, then I should simply avoid going to the meeting, in which case I do not expect to believe in flat earth tomorrow.

On the other hand, if I know I will have some perfectly good reasons to change my mind about the earth being flat, then whatever those good reasons are, *they are also good now*, so I should change current belief in accordance with the expectation of my future epistemic state.

There is a parallel to making promise: if I make a promise today, it is implied that I will keep the promise tomorrow, and a promise shouldn't keep considered as successful if I know perfectly well that I cannot keep my it. For instance, if I made a promise to someone that I shall never drink alcohol again, I am sabotaging myself by intentionally putting myself in a situation that will cause my promise to be broken.

The idea, then, is that making an assertion puts the agent under an obligation to defend and rationally cultivate the proposition being asserted. This makes enough sense for full belief, but there is a gap in carrying this analysis over to degrees of belief: what is the difference between asserting my degree of belief for $P$ to be 0.3 and to be 0.7?

Van Fraassen uses betting behavior and Dutch book arguments to fill this gap.[8] The idea is that asserting a judgment of probability requires me, quite literally, to put money where my mouth is, and the money involved should be directly proportional to my degrees of belief. This means that Reflection implies that what I think my willingness will be in the future should be my willingness to bet now. Van Fraassen shows that violating in Reflection will lead to the susceptibility to a Dutch book—is a set of bets that will lead to a guaranteed loss for anyone who accepts it. A Dutch book *argument* is aimed to demonstrate that the vulnerability to Dutch books a sign of incoherence and, therefore, irrationality. A Dutch book argument can be made in support of Reflection by showing that by violating this principle, one is opening oneself to being "Dutchbooked".

More specifically, the argument for Reflection requires a *Diachronic* Dutch book, which involves a Dutch book *strategy* to offer the agent bets that would be fair to the agent at the time, but the acceptance of them will ultimately lead to a loss to

8. van Fraassen, "Belief and the Will," 244.

the agent. The trick is to ask the agent to bet on her opinion about what her future opinion about $E$ will be, in addition to offering bets on her opinion about $E$ itself. If the agent gives an answer that violates Reflection, then the bookie will be able to make a Dutch book against her, by offering bets that are fair to her according to her credences at the time. This is supposed to show that violating Reflection is an act of self-sabotage. The Dutch book argument will be elaborated with technical details in the next section, even though it may be skipped without loss of continuity.

## 2.2 Dutchbook Argument for Reflection

Initially van Fraasseen uses a Dutch book argument to argue for the special Reflection principle.[9] However, he has come to downplay its importance.[10] This is partly due to the decision-theoretic assumptions needed for a Dutch book argument to succeed, especially on the simplistic model of the agent's willingness to accept bets.[11] Nevertheless, it is still useful illustration on how the violation of Reflection can lead to irrational behavior.

Suppose the Duke basketball team is playing against UNC tonight at 8pm. It is currently 1pm. Your friend asks you now for your subjective probability 4 hours later that you will be willing to bet on Duke winning at odds 2:1. For the sake of clarity, let us define these propositions:

$D$: Duke will win at 8pm.

9. Fraassen, 244.

10. van Fraassen, "Belief and the Problem of Ulysses and the Sirens," 12.

11. James M. Joyce, "A Nonpragmatic Vindication of Probabilism," *Philosophy of Science* 65, no. 4 (1998): 575–603

$B_5$: at 5pm, P(D) = 1/3.

Upon reflection, I respond that

$$P(B_5) = 0.4$$

Note that $P(B_5) = 0.4$ is just a simpler way to write $P(p_5(D) = 1/3) = 0.4$. That is, currently, there is a probability of 0.4 that in four hours, my credence for Duke winning will be 1/3. Now suppose my friend elicits yet another subjective probability from me. This time, she would like to know my personal probability for the eventuality that Duke loses and I come believe at 5pm that their probability to win is 1/3. In other words, what is the probability that $(\neg D \wedge B_5)$? Suppose I respond that

$$P(\neg D \wedge B_5) = 0.3$$

From this, $P(\neg D|B_5)$ is derivable:

$$P(\neg D|B_5) = \frac{P(\neg D \wedge B_5)}{P(B_5)} = \frac{0.3}{0.4} = 0.75 \tag{2.7}$$

And $P(D|B_5) = 1 - 0.75 = 0.25$. Recall that the current $t$ is 1, so $B_5$ is essentially the same as $p_{1+4}(D) = 1/3$—four hours later, I will come to believe that $P(D) = 1/3$. This means that I have violated the Reflection, for

$$P(D|p_5(D) = 1/3) = 0.25 \neq 1/3$$

Now, with this information, my friend then stages a Dutchbook strategy against me with the following bets:

| Bet | Condition | Reward | Cost |
|---|---|---|---|
| 1 | $(\neg D \wedge B_5)$ | 1 | $(1)P(\neg D \wedge B_5) = 0.3$ |
| 2 | $\neg B_5$ | 0.75 | $(0.75)P(B_5) = 0.45$ |
| 3 | $B_5$ | 0.083 | $(0.083)P(B_5) = 0.03$ |

**Table 2.1**: Dutch Book Strategy(Reflection Violated)

The trick is that, in order to devise a Dutchbook against me, the reward for bet 2 has to be $P(\neg D|B_5)$, for bet 3 it has to be $P(\neg D|B_5)$ minus my subjective probability of $P(\neg D)$ at 5pm, which is $0.75 - 2/3 = 0.083$. Since the costs for these bets were calculated using expected utility, I should regard all of them as being fair. All three bets cost me 0.78 in total. Now, at 5pm, there are two possible outcomes:

1. I do not come to believe that $P(D) = 1/3$: I win bet 2, but lost 1 and 3. This leads to a net loss of $-0.78 + 0.75 = -0.03$.
2. I come to believe that $P(D) = 1/3$. I get 0.083 for winning bet 3. Now bet 1 is now contingent on whether or not $\neg D$. My friend now offers me 2/3 to buy back that bet, which is fair in my light. I sell that bet, which renders the result of the game irrelevant. In this case, I have a net loss of $-0.78 + 0.083 + 2/3 = -0.03$.

My initial probability assignment then has rendered me vulnerable to Dutch books, because I have failed to follow the Reflection Principle. To see how, consider the situation if I had obeyed Reflection. As before, suppose that $P(B_5)$ is 0.4. When my friend asks for my value for $P(\neg D \wedge B_5)$, if I had followed Reflection, I would have realized that $P(D|B_5) = 1/3$. So, assuming independence,

$$\begin{aligned} P(\neg D \wedge B_5) &= P(\neg D|B_5) \times P(B_5) \\ &= (1 - 1/3) \times 0.4 \\ &= 0.27 \end{aligned}$$

So this means that my pre-Reflection respond—0.3—was 0.03 higher than it should be, had I followed Reflection, and this discrepancy is exactly how much I was sure to lose due to being Dutchbooked.

| Bet | Condition | Reward | Cost |
|---|---|---|---|
| 1 | $(\neg D \wedge B_5)$ | 1 | $(1)P(\neg D \wedge B_5) = 0.27$ |
| 2 | $\neg B_5$ | 0.675 | $(0.75)P(B_5) = 0.27$ |
| 3 | $B_5$ | 0.008 | $(0.008)P(B_5) = 0.003$ |

**Table 2.2**: Dutch Book Strategy(Reflection Satisfied)

## 2.3 Epistemic Failings

Still, if Reflection is indeed a general normative principle of rationality, there ought to be reasons to accept it independently of an argument from monetary loss. Dutch book arguments rely on specific and unrealistic assumptions about our betting behavior. Some have raised questions about the reliance on Dutch book arguments, and sought to accomplish what it does with less behavioral assumptions.[12] Even van Fraassen himself has distanced himself from the argument.[13]

The key to understanding Reflection intuitively is that it pertains to the rationality of one's epistemic policies and procedures regarding revising opinions. To begin, note that what Reflection asks is that our current opinions should be constrained by what we currently *consider* to be our future opinions. The idea is that, as a rational agent, I should see my future opinions as the consequence of my adhering to my current standard of rationality. If I have good reasons to think that it is rational for my future self to hold such an opinion, it should be good enough for my current self *now*.

Thus, consider van Fraassen's remark that

> [The] violation of this Principle is a symptom, within the current epistemic state, of a deeper defect: that the person holding this opinion cannot regard him or herself as following a rational policy for opinion change.[14]

An interpretation of this somewhat difficult passage is this. The assumption that my future self will epistemically superior is a normative point that we ought to presuppose in our thinking about our future prospects, since rationality requires us to cultivate and maintain our opinions over time. Another way to articulate this is that Reflection Principle requires a certain *all things being equal* clause being satisfied in our thinking about the rationality of our current and future conducts. In particular, "all things being equal" must include the requirement that in the time between I am committed to uphold and fulfill the same standard of rationality, so that my future self will do at least as well as I am now. For a simple example, suppose that, upon

12. James M. Joyce, "A Nonpragmatic Vindication of Probabilism," *Philosophy of Science* 65, no. 4 (1998): 575–603

13. van Fraassen, "Belief and the Problem of Ulysses and the Sirens," 12.

14. ibid., 17.

reflection, I conclude that my future self, with life experience I do not have now, will find my current spending habit quite irresponsible. If I could conceive thinking *that* in the future, my rational course of action *now* to take heed of my future self and revise my spending habits.

Consider van Fraassen's example of a meteorologist Piero.[15] Suppose Piero announces his forecast for the day at 8 a.m. every morning, and that he calculates the probability of rain for the next day based on his total evidence at 6 p.m. every night. If at 6 p.m. he is confident that, perhaps based on historical data, he will likely to see evidence at midnight that will drastically change his current prediction, he should base his opinion what his future self at midnight *would* have based on his data now, and what he confidently thinks his future self *would* have.

These two examples demonstrate the two senses in which my future self could be considered as worthy of my deference: my future self could make better judgments and my future self could also have more information.[16]

For instance, consider some well-known counterexamples to Reflection. They often involve cases in which I expect myself to make worse epistemic judgments: in the future, I might lose information, and I might be worse at making judgments. A prominent example of information loss is that one may reasonably anticipate future memory loss, so it would be reasonable not to defer to one's future opinion, thereby violating Reflection.[17] The basic idea is that it is not unreasonable to refrain from relying on future opinions when you are certain that in the future you will have

15. van Fraassen, "Belief and the Problem of Ulysses and the Sirens," 15.

16. Adam Elga, "Reflection and Disagreement," *Noûs* 41, no. 3 (2006): 478–502, 481.

17. William Talbott, "Two Principles of Bayesian Epistemology," *Philosophical Studies* 62, no. 2 (1991): 135–150, 138–41.

forgotten some crucial information. To use the mundane example from the literature: we often forget what we ate one year ago today, so it is reasonable to expect that one year later that I will forget what I eat now, but it does not mean that I should defer to my future self by concluding that I do not know what I ate today.

Another well-known example for violating Reflection is anticipated impaired judgments.[18] David Christense asks us to consider a person $B$ who has taken a state of mind altering drag that causes one to believe strongly that she could fly. Suppose $B$ is considering her probability of being to fly right after taking the drug but before it has taken effect. She knows in a few minutes she will believe strongly that she could fly, but it would make sense to violate Reflection here.

On my view, these counterexamples work by taking Reflection into contexts that we often consider non-epistemic: in general, we are forgiving in people's epistemic failure, due to our limited capacity. This, I think, is the crucial disanalogy between making a promise and making an epistemic judgment: promises are generally expected to be kept, except for extenuating circumstances; assertions about probabilities or facts are in general not regulated—this is what the memory loss example demonstrated. Similarly, in general we do not take people's partial assertions, e.g., "it seems to me likely that...", "$P$ is *probably* true", etc., too seriously: we definitely do not expect these assertions obey the laws of probability. We *know* that people don't.[19]

Failing to discharge the obligation incurred from making a promise often leads

18. David Christensen, "Clever Bookies and Coherent Beliefs," *Philosophical Review* 100, no. 2 (1991): 229–247, 234–37.

19. Daniel Kahneman and Amos Tversky, "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica: Journal of the Econometric Society*, 1979, 263–291, 263–91

to the loss of credibility, and this is why the promiser has the incentive to keep the promise (assuming that she cares about her credibility). The promiser's aversion to the loss of credibility is also why the promisee can be justified in taking the utterance as a genuine expression of the promiser's sincerity—no one would take a pathological lair's promise seriously, since credibility is no longer an issue. The social practice of making a promise runs on the currency of credibility.

Is there a similar institution for judgments of probability? To answer affirmatively, we need to locate a social convention that penalize the violation of Reflection. In general, however, I do *not* have to put money where my mouth is. Contrary to van Fraassen's claim, I do not think making a probabilistic judgment *in general* is like making a promise. There is a clear social convention in how the norms for promising are governed via an economy of credibility, but socially we usually do not hold assertions of probability as being a genuine expression of intention to consistently maintain the opinion asserted.

The close connection between degrees of belief and the willingness to bet provides an important clue. In specific decision-theoretical analyses of, say, business decisions, we are entitled to expect a stakeholder will act in accordance with her expected utility theory. To put the matter more positively, this is also saying that Reflection *does* work in some specific context, so if I were to play a game in which I am contractually obligated to make a probabilistic judgment and accept fair bets. In such a narrow context, I should obey Reflection, because I *will* incur a sure loss if I didn't. This suggests to me that Reflection is required only in contrived settings that makes Dutch book arguments possible—I do not say this as a criticism, but as a crucial insight to be developed.

Epistemic judgments are more context-dependent than making promises. This, I think, is unaccounted for by voluntarism. Probabilistic judgment that serves the aim of knowledge is exception to how we usually behave and not a norm. Therefore, my proposal is that, instead of trying to demonstrate how Reflection would work in multifarious contexts, we should focus on how Reflection would bear on assertions made *in the different contexts of inquiry*: making conjectures about a phenomenon of interest, verifying a prediction, etc. We need a richer conception of our epistemic practice, which, I suggest, can be found in Peirce's pragmatism, to which we shall now turn.

## 2.4 The Pragmatic Maxim

### 2.4.1 Operationalism and Counterfactual Context

Discussions of the Pragmatic Maxim often begin with Peirce's "How to Make our Ideas Clear", as it contains perhaps the most well-known articulation of the principle:

> Consider what effects, that might conceivably have practical bearings, we conceive the object of our conception to have. Then, our conception of these effects is the whole of our conception of the object.[20]

The Pragmatic Maxim, even on a limited reading, suggests that the meaning of a word is tied to some publicly perceivable phenomenon. So, the Pragmatic Maxim can definitely be read to support a kind of *operationalism*, which holds that the "practical bearings" mentioned above refers the *publicly accessible causal effects* of the object

20. Charles S. Peirce, "How to Make our Ideas Clear," in *Writings of Charles S. Peirce: A Chronological Edition, Volume 3: 1872–1878* (Indiana University Press, 1966). 266.

denoted by the word.[21] This operationalism is motivated by a commitment to define terms in a way that is conducive to scientific inquiry. Many ideas handed down to us by tradition do not naturally admit empirical investigation. Thus, we should clarify them so that questions about them can be resolved by appealing to evidence.[22] The operationalist reading of the Pragmatic Maxim emphasizes on Peirce's view on how we should analyze our ideas in service of scientific endeavors.

The operationalist reading of the Pragmatic Maxim may sounds suspiciously like the logical positivists' verificationism, which says, roughly, the meanings of words are entirely reducible to the difference they make in sense experience. However, this would be a misreading of the maxim, as it does not say a concept is reduced to its actual effects; instead, it states that concepts are delineated by the effects "that might *conceivably* have practical bearing, we *conceive* the object of conception to have". The distinction between the two is not verbal. Peirce takes the full understanding of an object to require the knowledge of not just how the object *has* behaved but also how it *would* behave in counterfactual situations. This emphasis on counterfactuals can be understood as the recognition of the *context-sensitivity* of the conditions of success for the application of concepts. Consider the classic discussion of the context-sensitivity of counterfactual conditionals by Nelson Goodman.[23] A counterfactual conditional is a proposition about the states of affair $C$ that would follow, if the antecedent of the conditional $A$ were true. Goodman points out that a small change of the content of the antecedent would change the truth-condition of the conditional itself, in a

21. F. Thomas Burke, *What Pragmatism Was* (Indiana University Press, 2013), 43.

22. Kevin Hoover, "Pragmatism, Pragmaticism, and Economic Method," in *New Directions in Economic Methodology* (Routledge, 1994), 292.

23. Nelson Goodman, "The Problem of Counterfactual Conditionals," *Journal of Philosophy* 44, no. 5 (1947): 113–128 113–28.

way inconsistent with how material conditionals behave in formal logic. Consider the following is a property of the material conditional but not the counterfactual conditional:

$$\text{‘If } A\text{, then } B\text{’ implies ‘if } A \text{ and } C\text{, then } B\text{’} \tag{2.8}$$

This is because the conditional in 2.8 is understood as $A$ cannot be true without $B$ being true, so the presence of $C$ ought not defeat it. But counterfactual conditionals do not behave this way. Consider:

If the match had been struck, it would have been lighted. (2.9)

If the match had been wet and then struck, it would have been lighted. (2.10)

Even though the former is true, the latter is not. The point is that counterfactual conditionals like the above presuppose a certain context. When I say that *the match would have lighted if it was struck*, I am assuming that the person to whom I am speaking understood the context relevant to this statement. If the other person asks whether the match was underwater, I would be entitled to somewhat incredulous, since I would assume that's the context. Thus, we are entitled to claim to have a understanding of the match's lighting behavior, only when we have our ability to pinpoint the exact context in which the counterfactual conditional is true.

While Peirce's operationalism ties our understanding of a concept to its empirical effects, it does not buy into crude reductionism to the strictly observable and the physical. For instance, Peirce says that the point of the diamond example is not a metaphysical reduction of the property of being hard into nothing but a list facts pertaining the diamond being scratched, but an epistemological point about how properties such as hardness can be probed through learning it *would* behave under all conceivable scenario, including counterfactual ones.[24] Only then we are entitled to say that we have thoroughly learned what the property of hardness is. Most important, these facts and counterfacts are evidence *for* the existence of abstract and theoretical entities, and not a reduction of them.

### 2.4.2 Inferentialism and the Context of Deliberate Conduct

It is clear that Peirce often intends the Pragmatic Maxim to be a principle that guides our analysis of concepts in service of empirical investigation. But in his later works, Peirce also expresses the Pragmatic Maxim as a thesis about how we ought to understand the normative implication of one's *acceptance* of a belief. This is the *inferentialist* reading of the maxim, which interprets the term "practical bearings" as *how concepts and beliefs bear on our practice.* That is, the acceptance of a belief puts a constraint on the agent's other beliefs and her actions.[25] Christopher Hookway's interpretation of the Pragmatic Maxim is an example of an inferentialist reading:

> If I believe or assert a proposition, I commit myself to the expectation that future experience will have a particular character. If this expectation is disappointed, than I will probably have to abandon the belief or withdraw the assertion. Clarification of a concept using the pragmatist principle provides an account of just what commitments I incur when I believe or assert a proposition in which the concept is ascribed to something.[26]

24. Charles S. Peirce, *Collected Papers of Charles Sanders Peirce* (Cambridge: Harvard University Press, 1931), 5.453.

25. Burke, *What Pragmatism Was*, 44.

26. Christopher Hookway, *Truth, Rationality, and Pragmatism: Themes From Peirce* (Oxford University Press, 2000), 60.

Consider the following formulation of the Pragmatic Maxim from Peirce's later work:

> [Pragmatism is] the maxim that the entire meaning and significance of any conception lies in its conceivably practical bearings, - not certainly altogether in consequences that would influence our conduct so far as we can force our future circumstances but which in conceivable circumstances would go to determine how *we should deliberately act*, and how we should act in a practical way and not merely how we should act as affirming or denying the conception to be cleared up.[27]

One striking difference between these formulations and the one from "How to Make Our Ideas Clear" is the emphasis on deliberative actions and rational conduct. "Deliberate conduct," Peirce further explains, is "self-controlled conduct."[28] The crucial idea here is that since the meaning of a word manifests itself through its causal and practical effects, a belief, which consists of the use of these words, should have a practical effect on those who accepts the belief. But the effect here is not one of a causal one, but a rational one.

Thus when Peirce identifies beliefs as habits, he does not mean a blind response to stimuli but a "deliberate, or self-controlled, habit."[29] What Peirce has in mind in particular is that the accepting belief implies a rational constraint on the believer's future conduct, for "future facts are the only facts we can, in a measure, control", so this is what is meant by the idea that beliefs can have rationally binding repercussions on our future conduct.[30]

Putting the two together, what emerges is a more nuanced interpretation of Re-

27. Charles S. Peirce, *The Essential Peirce, Volume 2: Selected Philosophical Writings (1893-1913)*, ed. Peirce Edition Project (Indiana University Press, 1998), 145, my emphasis.

28. ibid., 348.

29. Peirce, *Collected Papers of Charles Sanders Peirce*, 5.480.

30. Peirce, *The Essential Peirce, Volume 2: Selected Philosophical Writings (1893-1913)*, 359.

flection: making a probabilistic judgment may imply a certain commitment, such that I will have to act deliberately in a particular way that I would not have otherwise. But this normative link is highly contextual: just like the lighting of a match could fail if one of the necessary conditions is not met. I will devote the rest of this chapter into developing this by diving further into Peirce's philosophy.

Peirce's view on the normative dimension of assertion is especially relevant. According to Peirce, the assertion of a proposition entails a normative commitment: "to assert a proposition is to make oneself responsible for its truth."[31] The very idea of responsibility has both the prospective and deliberative elements that the inferentialist reading of the pragmatic maxim exploits: to undertake the responsibility of task usually means the responsible party will deliberatively carry out the task involved some time in the future.

Peirce also sees a link between asserting a proposition and the willingness to bet: "Both are acts whereby the agent deliberately subjects himself to evil consequences if a certain proposition is not true."[32] The analogy Peirce is making here is that the acceptance of a belief cannot be understood without appealing to normative concepts such as commitment, intention, and responsibility: to believe that $p$ involves the expression of the agent's intention to assume the responsibility of fulfilling the normative commitments entailed by the belief. Betting is an instance of this: betting on $p$ being true means making a commitment to act a certain way if the proposition were false: making the bet means the agent is now responsible. For the bet to be genuine, the bettor must express her intention to pay if she happens to lose the bet:

31. Peirce, *Collected Papers of Charles Sanders Peirce*, 5.543.

32. Peirce, *The Essential Peirce, Volume 2: Selected Philosophical Writings (1893-1913)*,140.

she either loses money, or, if she refuses to pay, loses credibility.

Peirce's view then anticipates and overlaps with van Fraassen's voluntarism. What makes Peirce's view, I think, superior to voluntarism is his recognition that the normative force of assertion is *context dependent*. In Peirce's words, the "measure of assurance" implied by an assertion is a function of the context in which it is asserted[33] An assertion in the court of law, for instance, presupposes a high degree of assurance, and implies the responsibility on the asserter's part to demonstrate its truth:

> If a man desires to assert anything very solemnly, he takes such steps as will enable him to go before a magistrate or notary and take a binding oath to it.... it would be *followed by very real effects, in case the substance of what is asserted should be proved untrue*... if a lie would not endanger the esteem in which the utterer was held, nor otherwise be apt to entail such real effects he would avoid, the interpreter would have no reason to believe the assertion[34]

The crucial point here is that there is nothing magical about an assertion in and of itself: to make an assertion to count as a genuine expression of the intention is to assume the responsibility for the truth of the asserted proposition, the speaker has to enter into a proper social context in which the speaker's failure to fulfill this responsible will lead to "very real effects", such as the risk of legal jeopardy. Peirce recognizes the subtle differences of context make a world of difference in terms of the normative force implied by the assertion. For instance,

> Nobody takes any positive stock in those conventional utterances, such as "I am perfectly delighted to see you," upon whose falsehood no punishment at all is visited.[35]

33. Peirce, *Collected Papers of Charles Sanders Peirce*, 4.54.

34. ibid., 5.546 my emphasis.

35. ibid., 5.546.

Thus, assertions made during "small talks" are not usually interpreted in assuming the epistemic responsibility for standing up for the truth of the statement, and thus are not subject to the same standard of criticism as a serious assertion. If two different people asked how I was doing, and the answer I give each of them is inconsistent with each other, this is not usually considered the violation of norms. This distinction is helpful in defending Reflection against the counterexamples. In the cases of memory loss, in many if not most social contexts, it is often excusable or expected to be forgetful. Facts pertaining what one had eaten in the past are rarely relevant in situation where the speaker's intention will be called into question. On the other hand, if a speaker is testifying in a court of law under oath, she is expected to stand by any assertion that is made in future. Inconsistency due to faulty memory is strictly regulated and punished through the hearsay rule and cross-examination.[36]

## 2.5 The Abductive Context of Inquiry

Peirce's pragmatism provides a helpful framework to see why rationality of epistemic judgment cannot be totally captured in voluntarist terms. James' and, by extension, van Fraassen's voluntarism is based on an interpretation of the Pragmatic Maxim that, I want to suggest, leads to a problematic interpretation of Reflection.

Peirce makes a very helpful remark in a letter to James, in response to receiving a copy of *The Will to Believe*, a book dedicated to Peirce himself.[37] Peirce points out that James has conflated the crucial distinction between *provisionally* accepting

36. David Greenwald, "The Forgetful Witness," *The University of Chicago Law Review* 160, no. 1 (1993): 167–195, 167.

37. Peirce, *Collected Papers of Charles Sanders Peirce*, 8.250–251.

a belief for the sake of further inquiry, and accepting a belief without any evidence. From Peirce's point of view, James has taken the general lesson of the Pragmatic Maxim, that "everything is to be tested by its practical result", to its logical extreme, that "mere action as brute exercise of strength that is the purpose of all."[38]

The thought is that if the Pragmatic Maxim is interpreted as putting forth the thesis that the rationality of the acceptance of beliefs and the application of concepts is evaluated in terms of whether their practical implications are satisfied, then one may conclude, as James does, that (some) beliefs can be justified simply by virtue of committing into the mode of behavior required by the belief. This is why the central argument in *The Will to Believe* is that one's belief in God can be justified as long as one is committed into acting *as if* God exists, i.e., following the ritualistic standard in the appropriate tradition.

This feature is carried over to van Fraassen's voluntarism. The analogous move here is the claim that if we treat reports of degrees of beliefs as nothing but perfomative speech acts that subject the speaker to making a commitment in standing by the assertion in proportional to its degree, such as willingness to bet, then an epistemic judgment is considered as rational so long as the agent's act in accordance with it in the future. This is what Reflection is supposed to codify.

From a Peircean perspective, however, this is an incomplete picture of our epistemic practice. The matter can be divided into two different points. First, ampliative reasoning, i.e., extrapolation beyond what our evidence says, can be justified in certain contexts. As is well known, Peirce is responsible for distinction between abduction, deduction and induction. Following Issac Levi, we can understand abduction,

38. Peirce, *Collected Papers of Charles Sanders Peirce*, 8.250.

deduction, and induction broadly not only as distinct types of *inferences*, and but also different *tasks* involved in our practice of inquiry.[39] Abduction concerns how the question of the inquiry is *framed*, such as choosing the hypothesis for testing. Once a framework for inquiry has been chosen, it is the task of deduction to tease out the necessary consequences, such as sub-hypotheses, so that the criteria for empirical adequacy are clear. *Induction* then takes place by testing the hypothesis against the deliverance of experience.

What I shall argue below is that Reflection and voluntarism should be understood in the context of *abduction*. My contention is that the seemingly perplexing combination of the permissive rationality of voluntarism and the normative constraint of Reflection makes sense as part of the exploratory stage of inquiry, in which we have to make decisions about our experimental commitment *before* having any evidence at hand. I must, however, begin by addressing the thorny issue of inference to the best explanation(IBE). This is essential, because (1) contemporary philosophers of science often use abduction as a synonymy of IBE, and (2) van Fraassen is well-known for arguing against any attempt to incorporate IBE in the context of probabilistic reasoning. A critical discussion of these issues will not only clarify the sense of abduction needed for my position, but also dislodge in advance van Fraassen's criticism.

### 2.5.1 Van Fraassen's Anti-Abductivism

For the sake of brevity, let us call *explanationism* the position that accepts inference to the best explanation(IBE) as an indispensable rule of *inductive* inference. In its

39. Isaac Levi, "Beware of Syllogism: Statistical Reasoning and Conjecturing According to Peirce," in *The Cambridge Companion to Peirce*, ed. C. J. Misak (Cambridge University Press, 2004), 257, 281.

most naive form, IBE says that we should infer that the hypothesis that best *explains* the evidence we have is the one we should *accept*. Van Fraassen is well-known for arguing against IBE in this naive form. In its most powerful form—a view that van Fraassen does ascribe to some philosophers—IBE can be construed as a solution to Hume's problem of induction, which holds that there is no justification, ampliative or explicative, for extrapolating inductively beyond the evidence we have. IBE gets us out of this problems by giving justificatory force to explanatory virtues, so that the best explanation is the one we *should* accept, in that we are inductively licensed to make the ampliative leap from *this hypothesis is the best explanation* to *this hypothesis is true*. Van Fraassen attacks this position relentlessly. One often cited argument of his is that we never pick the best explanation *simplicter*, but the best *out of the explanations available to us*. Van Fraassen argues this is a horrible justification for a belief, since for some reason we might only have horrible explanations available to us, so 'our selection may well be the best of a bad lot.' [40]

Van Fraassen suggests that, in light of the criticisms against naive PIBE, the strongest recourse available the supporters of IBE is *entrenchment*, which amounts to the repackaging IBE into a rule that works well with Bayesianism. The more plausible way to do this, according to van Fraassen, is to give explanatory virtues a place in the revision of belief in light of new evidence:

> Combining the ideas of personal probability and living by rules, the new rule of IBE would be a recipe for adjusting our personal probabilities while respecting the *explanatory* (as well as predictive) success of hypotheses.[41]

Let's call this new rule 'probabilistic inference to the best explanation' (PIBE),

40. Bas van Fraassen, *The Scientific Image* (Oxford University Press, 1980), 143

41. van Fraassen, *Laws and Symmetry*, 149.

which entitles us to raise the probability of the best explanation. For instance, consider Peter Lipton's description of his position in *Inference to the Best Explanaton*:

> [My] version [of IBE] claims that the explanation that would, if true, provide the deepest understanding is the explanation that is likeliest to be true. Such an account suggests a really lovely explanation of our inferential practice itself, one that links the search for truth and the search for understanding in a fundamental way.[42]

Van Fraassen, however, argues that this cannot do. To begin, if IBE is to be harmonized with Bayesianism, it must not clash the Bayesian procedure of belief revision, i.e., conditionalization, but van Fraassen argues that PIBE is inconsistent with the Bayesian standard of rationality. The problem again is Bayesianism's allergy to ampliative reasoning: PIBE is ampliative, since explanatory virtues goes beyond what is logically implied by our evidence, so it is incompatible with the explicative nature of Bayes' theorem—the posterior probability is nothing but an arithmetic consequence of conditional probability. Since PIBE is the rule that confers 'bonus' probability to a belief based on its explanatory virtue, this bonus conflicts with Bayesianism.

Van Fraassen uses a Dutch book argument against explanationism. The gist of the argument is that an explanatioinst will violate conditionalization when they raise the probability of the best explanation. This leads to a guaranteed loss when the bookie offers bets based on the explanationist's prior belief, and then based on the explanationist's belief after *both* conditionalization and PIBE. A simplified version of his argument is presented in the next section. Readers uninterested in the technical details can skip it without loss of continuity.

42. Peter Lipton, *Inference to the Best Explanation* (Routledge, Taylor / Francis Group, 2004),63.

### 2.5.2 Van Fraassen's Dutch book argument against PIBE

Suppose we are interested in the bias of a certain coin, $\theta$, which indicates the probability of the coin landing on heads. Suppose we know that there are 3 equally probable hypotheses: (A) $\theta = 0.9$, (B) $\theta = 0.5$, and (C) $\theta = 0.1$. Suppose our evidence gathering process is described as follows: $X_i = 1$ denotes 'the coin has landed on heads on the $i$ toss' and $X_i = 0$ otherwise. Suppose we have toss the coin 4 times, and they all landed on heads. So the evidence $E$ is $\sum_{i=0}^{4} X_i = 4$. The marginal probability for $E$ is:

$$P(E) = P(A)P(E|A) + P(B)P(E|B) + P(C)P(E|C) = 0.24$$

Using Bayes' theorem, the posterior probabilities are: $P(A|E) = 0.9129$, $P(B|E) = 0.0869$, and $P(C|E) = 0.0001$. So far so good—4 consecutive heads favors the hypothesis that the coin is bias toward heads, which is what conditionalization is showing us. Where does PIBE comes in, then? Van Fraassen asserts that an argument from PIBE would be as follows: out of the three hypotheses, $A$ best *explains* why we see nothing but heads: it's because it's highly biased. So PIBE should recommend the redistribution of the probabilities such that $P(A)$ would come out even higher. Suppose we raise $P(A)$ to 0.999. This amounts to giving the best explanation a bonus of 0.086 in probability. To accommodate this, we can lower the probabilities of the other hypotheses accordingly. For instance: $P(A) = 0.999$, $P(B) = 0.00086$, and $P(C) = 0.00014$. So we have extrapolated ampliatively beyond what the evidence tells us by using PIBE.

This line of reasoning, however, flies in the face of the Bayesian notion of *coherence*, since it renders one subject to a set of bets that ensures whoever takes these bets a loss. Imagine that we are back at the beginning before we tossed 4 heads. Before tossing the coin for four times, we were offered the following set of bets. Let $E$ again be 'the coin is tossed 4 times and they are all heads' and $H$ be the $X_5 = 1$, that is, 'the fifth toss turns up heads'.

1. \$10,000 if $E$ and $\neg H$.
2. \$1300 if $\neg E$.
3. \$300 if $E$.

Now, we can calculate the values of these bets based on our prior probabilities:

1. Bet 1: $(10000)\frac{0.8^4(0.2)+0.5^5+0.2^4(0.8)}{3} = 323.16$
2. Bet 2: $(1300)(1 - 0.158) = 988.56$
3. Bet 3: $(300)0.158 = 71.87$

So these bets would be worth \$1383.6 in total. Suppose we bought these bets for exactly that much from a bookie, who then proceeded to toss the coin for 4 times. Either $E$ is true or it is false. Suppose it's false—at least one toss landed on tails. In this case, we would have won bet 2 but lost 1 and 3. We would receive \$1300 but this would still lead to a total loss of $-1383.6 + 1300 = -83.6$.

On the other hand, suppose $E$—all tosses turned up heads. We would receive \$300 per bet 3, and now bet 1 would depend entirely on the fifth toss. Now, van Fraassen asks, what should our degree of belief for $\neg H$, that the fifth toss will land on tails? Recall that we have used PIBE to give a bonus to the most explanatory hypothesis, $A$, which effectively has raised the marginal probability of $H$ to 0.9. At this point, bet 1 is now worth $(10000)P(\neg H) = (10000)0.1 = 1000$. Suppose the bookie offers us exactly \$1000 to buy this bet back. We would regard it as fair and accept his offer. In this scenario, we end up with $-1383.6 + 300 + 1000 = -83.6$—we incur exactly the same loss as we would if $E$ were false.

### 2.5.3 Abduction, Predesignation, and Reflection

Van Fraassen's argument is often misunderstood. When explanationists cite van Fraassen's anti-IBE argument, it may sound as though van Fraassen takes Bayesianism as the correct position by fiat, and, a this line of thought goees, any position that contradicts it is *de facto* an inviable one. Indeed, this is how many explanationists read him. Take Lipton for instance:

> In its simplest form, the threatening argument says that Bayesianism is right, so Inference to the Best Explanation must be wrong.[43]

But given our discussion of voluntarism, this assessment of the situation is not quite right, at least in the original context of the argument. The voluntarist argument is not that it is irrational to use conditionalization or IBE, but that (1) they are not rationally compelling in and of themselves, and (2) repacking IBE as some probabilistic rule is inconsistent with Bayesian conditionalization.

Many explanationists, including Lipon who cleaely misread van Fraassen's argument, attempts to challenge van Fraassen's negative argument heads on, by defending that both conditionalization and PIBE are rationally compelling. By taking this approach, explanationists have to argue for the legitimacy of PIBE within the stringent requirement of Bayesianism. To do so, they must give an account of how explanations

43. Lipton, *Inference to the Best Explanation*, 104.

can influence probabilities without subjecting oneself to a Dutchbook Argument. The strategy here is a mixture of neutralizing PIBE so that it will not get in the way of conditionalization, and providing incentives for Bayesians to adopt it by amplifying or emphasizing PIBE in areas where conditionalization comes up short. The result of this approach is often an unsavory stew, since it has to dance around neutralizing and strengthening PIBE without either trivializing it or over-promising what IBE can do. Some examples: Samir Okasha argues for a view that appeals to the phenomenology of explanation.[44] By appealing to behavioral economics, Lipton argues for a view that interprets PIBE as a heuristic for how people, as a matter of psychology, do Bayesian computation in everyday reasoning.[45] Jonathan Weisberg proposes to do away conditionalization altogether, but he proposes to couple PIBE with a even stricter conception of rational degrees of belief like Keynes has (1.2.2), requiring the recommendations of PIBE to somehow coincide perfectly with the probability assignments of *the* one true probability function.[46]

These attempts try to repackage PIBE into something that is less like a normative rule and more like an empirical story. This is necessitated by accepting that conditionalization is the sole possible move in response to evidence. I think that these philosophers have played right into van Fraassen's hands by engaging armchair psychology while trying to appease the Orthodox Bayesian. Alternatively, we could take the path of less resistance: instead of hamstringing explanatory reasoning in order to accommodate a stringent requirement of rationality, we should exploit the liberating

44. Samir Okasha, "Van Fraassen's Critique of Inference to the Best Explanation," *Studies in History and Philosophy of Science Part A* 31, no. 4 (2000): 691–710

45. Lipton, *Inference to the Best Explanation*, chapter 7

46. Jonathan Weisberg, "Locating IBE in the Bayesian Framework," *Synthese* 167, no. 1 (2009): 125–143

nature of voluntarism. Like explanationist, I think that explanatory reasoning is an indispensable element in our probabilistic and statistical reasoning, but I also agree with van Fraassen that PIBE is a non-senscial rule.

I advocate a largely Peircean position that can accommodate both of these points. In Peirce's ideas we can find a view that agrees with explanationism that abduction is indispensable to induction but also with van Fraassen's voluntarist point that the acceptance of a hypothesis have practical repercussions on the inquirer's future behaviors. The crucial point is that making an epistemic commitment is not "a brute exercise of strength"—instead such a commitment can satisfy the aim of inquiry only if it is situated with framework that allows *vindication* if the hypothesis withstands the tribunal of experience, and *correction* if it fails. The development of this idea is the project of the whole dissertation. As a transition to the next chapter, I will discuss a Peircean way of thinking about the role explanations play in inference.

The important idea is that explanatory reasoning does not belong to the domain of explicative reasoning. Accepting this, however, means that we should not look for reliable inferences in explanations. The permissiveness of voluntaristic conception of rationality is characteristic of ampliative reasoning that occur with in the context of Peirce's conception of abduction. This, I think, is the context in which we should understand how explanatory reasoning can inform probabilistic judgments. At this stage of inquiry, different hypotheses can be introduced to the logical space of reasons for non-evidential reasons, hence abduction is "the first starting of a hypothesis and the entertaining of it."[47] For instance, hypothesis that in any way explain the phenomenon under investigation is permissible. In his early years, Peirce uses the

47. Peirce, *Collected Papers of Charles Sanders Peirce*, 6.525.

following syllogistic schema to illustrate one example of abduction:

1. The surprising fact, C, is observed;
2. But if A were true, C would be a matter of course.
3. Hence, there is reason to suspect that A is true.[48]

While Peirce, at least in his later years, does not think that this schema encompasses all possible forms of abductions, it is nevertheless a helpful starting point: $C$ is the observed *fact* that calls for a hypothesis that would, in Peirce's words, *rationalize* it.[49] Peirce sometimes refers to $A$ as an *abductive suggestion* that follows the perception of the surprising fact.[50] They are *possible* accounts for the phenomenon observed, and are not presented as true proposition, but hypotheses that we may accept provisionally for the sake of further testing. Thus, in the context of abduction, ampliative inference is justified, because the goal here is the introduction of hypotheses, without which neither induction nor deduction can proceed. Hence Peirce holds that the role of abduction is paramount to the growth of knowledge, as it is "the only logical operation which introduces new ideas."[51] This is why explanatoriness is a relevant factor in abduction: During abduction, explanatory virtue is clearly a relevant consideration, for the construction of a framework pertains to the choice of a hypothesis to be tested, and how much the hypothesis, *if true*, explains would provide ground for making the hypothesis a genuine contender.

We see a crucial difference in the Peircean's conception of abduction and the

48. Peirce, *Collected Papers of Charles Sanders Peirce*, 5.189.
49. Peirce, *The Essential Peirce, Volume 2: Selected Philosophical Writings (1893-1913)*, 107.
50. ibid., 227.
51. ibid., 216.

explanationist conception. Peirce clearly sees no conceptual connection between a hypothesis's explanatoriness and its probability. Abductive reasoning is a creative but risky process:

> The abductive suggestion comes to us like a flash. It is an act of insight, although of extremely infallible insight.[52]

> ... abduction is, after all, nothing but guessing.[53]

Therefore, in the abductive context, an agent can accept a hypothesis that is not only improbable but also with low explanatory values, if the hypothesis is *informative.* For instance, we may choose to accept an improbable hypothesis provisionally if we know, through deduction, that the confirmation or rejection of the hypothesis will necessarily rule out many others. Peirce also suggests that a hypothesis may be adopted for pure economical reasons. So, unlike the explanationists, Peirce sees no reason to think that an explanation that explains is one that is also probable. Explanatory virtue is but one of many relevant factors in abduction.

Peirce's view on abductive rationality, then, *is* voluntarism: in the context of abduction, we are free to extrapolate ampliatively. Since the choice of a hypothesis is the result of a balancing act between many competing concerns, such as the economy of research or explanatory virtues, they are not justified by any mechanical rules. An abduction, however, does give rise to regulative constraints that require deliberate adherence. This is made clearly by Peirce here:

> An Abduction is a method of forming a general prediction without any positive assurance that it will succeed either in the special case or usually, its justifi-

52. Peirce, *The Essential Peirce, Volume 2: Selected Philosophical Writings (1893-1913)*, 227.

53. Peirce, *Collected Papers of Charles Sanders Peirce*, 1931, 7.219.

> cation being that it is the only possible hope of regulating our future conduct rationally.[54]

As discussed, the notion of deliberate conduct is central to Peirce pragmatism, and to understand the role it plays in relation to abduction, we must also understand Peirce's statistical understanding of inquiry.

As Peirce sees it, the validity of inductive reasoning is strictly regulated by the statistical structure deductively by the commitments she makes in the abductive context. This means that the epistemic judgments made during abduction cannot be changed during the inductive context, otherwise the inference would be invalidated altogether. This can be seen as a requirement to hold an experimental version Reflection: I have to stand by my decision to provisionally accept the abduced hypothesis, before my obligation is satisfactory discharged at the very end of the experiment. I cannot, for instance, rationally change my hypothesis to fit the data better in the middle of the data-gathering process. Peirce codifies this requirement as *the rule of predesignation*, which he calls a necessary guiding principle of induction.[55]

The idea is that the inquirer's commitments, intentions, and judgments all make contribution to the context in which the inductive inference is made. This is why the decisions the agent makes during the abductive context constitute the experimental framework for the testing of the chosen hypothesis, and the rule of predeisgnation imposes a rational constraint on the inquirer's opinions during the inductive context. In particular, the decision the agent makes during abductive phase requires the agent to stand by these commitments during the inductive stage of the inquiry. This is how Peirce explains it:

54. Peirce, *Collected Papers of Charles Sanders Peirce*, 2.270.

55. Peirce, *The Essential Peirce, Volume 2: Selected Philosophical Writings (1893-1913)*, 48.

> If in sampling any class, say the M's, we first decide what the character P is for which we propose to sample that class, and also how many instances we propose to draw, *our inference is really made before these latter are drawn*, that the proportion of P's in the whole class is probably about the same as among the instances that are to be drawn, and the only thing we have to do is to draw them and observe the ratio.[56]

In other words, making statistical inference presupposes a commitment to the assumptions implied by what it means to carry out random sampling. In particular, the responsibility must be taken up prior to the sampling itself, by committing into the stipulations in the experimental setup, such as the statistical hypothesis to be tested and the length of the trial. Hence, Peirce says that once those details have been settled, the inquirer must take the responsibility of carrying it out exactly as she described; otherwise, the inference is illegitimate. The thought is that the presupposition that the experimenter has faithfully fulfilled her duty is a background assumption in the context of induction, because statistical sampling can easily be gamed:s

> But suppose we were to draw our inferences without the predesignation of the character P; then we might in every case find some recondite character in which those instances would all agree. That, by the exercise of sufficient ingenuity, we should be sure to be able to do this, even if not a single other object of the class M possessed that character, is a matter of demonstration.

Using van Fraassen's voluntaristic terminology, such an action is *self-sabotaging*, because if the length of the trial is unfixed, the investigator can keep on sampling until they find a sample that supports her hypothesis. The sampling here would not be random, and the inference would not be valid. After the investigator has made her experimental commitments clear, as Peirce says, "the only thing we have to do is to

56. Charles S. Peirce, "A Theory of Probable Inference," in *Writings of Charles S. Peirce: A Chronological Edition, Volume 4: 1879–1884* (Indiana University Press, 1989), 434.

draw them and observe the ratio." This is why Peirce insists that rationality *requires* what he calls deliberate conduct—I must stand by the decisions I made in abduction, which is to faithfully carry out the sampling process dictated by the probability model, and accept or reject the hypothesis based on the criteria I specified.

## 2.6 Conclusion

This serves as an important transition to the next chapter. Peirce's Pragmatic Maxim and tripartite classification of different contexts of inquiry provide a unified way of understanding how epistemic commitments figure into our probabilistic inferences and experimental procedures. Making inductive inferences based on experimental procedures *presupposes* that we can rely on the experiment to behave deliberately in accord to the setup of the experiment promised at the outset. If the experiment is stopped earlier or later than promised, the character of the inference could be drastically change. How inferences could be dependent on whether or not these practical commitments are satisfied is the focus of the coming chapter. The conclusion I put forth in the this chapter, then, leaves open the question of how exactly deliberative elements can be criticized from a Bayesian perspective. This will be addressed in next chapter.

# Chapter 3

# Deliberation and the Problem of Optional Stopping

In the last chapter, I tried to incorporate the Peircean notion of abduction into a broadly probabilist framework through an inferentialist-pragmatist reading of van Fraasseen's Reflection Principle and voluntarism about belief. I sketched out a framework that aims to accommodate the voluntarist idea that epistemic judgments are speech acts with normative implications, and Peirce's conception of rationality as deliberate conduct. Peirce's rich notion of abduction, which goes beyond the mere appeal to explanatory values, supplies the context-sensitivity needed to understand the voluntarist commitments codified by the Reflection Principle.

The main lesson, as I suggested, is that the normative force that regulates epistemic commitments one undertakes in making a probabilistic judgment cannot be understood without taking into account the *context* of inquiry in which the judgment is made. An assertion, as Peirce suggests, has no normative force unless the assertion is underwritten by an epistemic practice that incentivizes the agent to stand by the obligations imposed upon her. Dutch book scenarios, though in general unrealistic, can be seen as one such context, since in the setup the agent is stipulated to make bets and revise her degrees of belief in a specific way.

More needs to be said about these contexts, especially how they work within experimental practices of hypothesis testing. In this chapter, I endeavor to develop my position in the specific context of inductive and statistical inference. To summarize the position, I defend the following slogan:

**The Deliberativist Thesis** Inductive inference must be interpreted in light of the agent's commitments, intentions, and values that frame the deliberative framework in which the deliberate data gathering procedure is structured.

What I call a *deliberative framework* is the context of induction chosen in the abductive stage: it is the provisional selection of the hypothesis to be provisionally accepted and probed, the probabilistic judgments deduced from the statistical model used to represent the phenomenon of interest, and experimental procedure to practically engage in these assumptions. The term "deliberative" is designed capture Peirce's idea, as discussed in 2.4.2, that accepting a belief or making a judgment means one is rationally required to conduct oneself in a deliberate way, in accordance with the relevant inferential commitments. Even a provisionally accepted hypothesis implies the agent's commitment to test the viability of it. Many elements in this deliberative framework, moreover, cannot be determined mechanically, and must be decided upon according to on the experimenter's intentions and values.

I will defend and develop the deliberativist thesis by critically examining the *problem of optional stopping*, one of the many points of contention in the Bayesian-frequentist debate. The gist of this problem is that the frequentist conception of statistical evidence is dependent on the intentions of the experimenters, such that the experimenter may manipulate the evidence just by changing her intention to stop the sampling. This, Bayesians argue, is detrimental to the validity of frequentist inference. Instead, many Bayesians hold that the *total* evidential import of data can be fully captured by what is called the *likelihood function*, which is impervious to the effects of the agent's intentions. This is called *the Likelihood Principle.*

My goal of this purpose is to resist the Likelihood Principle by addressing the problem of optional stopping. Since denying the involvement of intentions in statistical inference implies the rejection of the deliberative thesis, I do not propose to show that the effects of optional stopping do not exist; instead, I argue that deliberative elements such as the intention to stop are relevant also to Bayesian reasoning, so the deliberativist thesis holds within the Bayesian context.

In section 3.1, I will give a historical presentation of the problem of optional stopping in the context of its used by parapsychologists to, essentially, cook up evidence for the existence of "extra-sensory perception" (ESP). In section 3.2, I will sharpen Bayesians' argument that the statistical involvement of agent-intentions is to be blamed for optional stopping, and why they think this is a critical flaw of frequentism. In section 3.3, however, I will explain how the same problem can apply to Bayesan reasoning. In the last section, 3.4, I examine a potential response to my argument by considering a proof about *the value of evidence* given by Ramsey and Savage.

## 3.1 ESP and Optional Stopping

On April 24th, 1940, the mathematician W. Feller delivered a lecture on his critique of the statistical method used in parapsychological research at a Duke mathematic seminar.[1] At that time, J. B. Rhine was spearheading Duke's parapsychology research: to make parapsychology scientifically respectable, Rhine believed, statistical evidence must be used to support the conclusions he hoped to demonstrate.[2] Feller points out, however, many results of these experiments involve on a trick called "optional stopping", which is used to abuse statistics to get their desired outcome. Feller argued that such experimental practices invalidate the result of parapsychological studies. Feller's specific criticism against parapsychology, however, became the starting pointing of a general critique of frequentist statistical methods, often mobilized by Bayesian statisticians. The argument is that, while the parapsychologists no doubt had questionable experimental practices, it is a flaw of the frequentist methods they employed. The problem, Bayesians argue, is that a statistical conclusion ought not be influenced by extra-statistical concerns such as when the experimenter decides to stop.

One of the phenomena for which parapsychologists claimed to have found statistical evidence is *extrasensory perception* (ESP), i.e., that some people can perceive certain facts without the use of any of the five senses. How can such a claim be examined experimentally and statistically?

One often used experimental setup for testing ESP is an activity called "card guessing," using the so-called Zener cards, after Karl Zener, the Duke psychologist who suggested the design to Rhine.[3] A Zener card can have one of 5 unique symbols. A typical deck of Zener cards contain 5 cards for each symbol. A trial of this

1. William Feller, "Statistical Aspects of ESP," *The Journal of Parapsychology* 2, no. 4 (1940): 271–298, 281

2. Seymour Mauskopf and Michael McVaugh, *The Elusive Science* (K. Paul, Trench, Trubner & Co., 1980), 108

3. J. B. Rhine, *Extrasensory Perception* (Branden Publishing Company, 1964),48

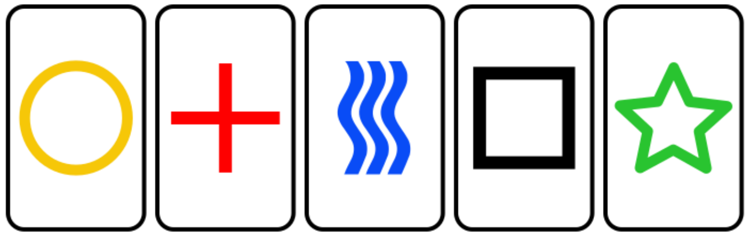

**Figure 3.1**: Zener Cards, by Mikhail Ryazanov. Licensed under CC BY-SA 3.0

experiment typically involves a subject being repeatedly asked to guess the face of a randomly chosen card, while the investigator would note the actual face of the card and the subject's answer. After each trial, the subject's reported sequence would then be compared to the observed result, and a score would be calculated based on the number of correct guesses.

Of course, one could achieve a high score by chance: no one would think I had ESP if I successfully had predicted the outcome of a coin flip, because we know that, no matter what my prediction is, I would still have a 0.5 probability of getting it right. But what if I guessed correctly the result of 10 consecutive tosses? The probability of that is $0.5^{10} = 0.001$. We are much more inclined to say that I have some sort parapsychological ability, because it would have been an extraordinary coincidence if I got it right purely by chance. This is the sort of statistical argument that Rhine and his followers tried to make. Their contention is that if a subject can attain a score that is too extraordinary to be explained by chance, then we have statistical evidence for the person's ESP ability. In statistics, the probability of an outcome, assuming that it is by chance, is called the $p$ value.

Since there are 5 faces, a subject has the probability of 0.2 of getting it right just by guessing alone, so someone with ESP should do better than that. This is how "better" can be explicated: suppose a trial with 100 attempts has been carried out on a subject. If the subject is just guessing, then we would expect that she would get around 20 cards right. In fact, according to the theory of probability, we can ascertain that out of 100 cards, there is a probability of 0.944 that the subject can get 26 or fewer correct guesses. Conversely, the probability that a guesser can get 27 or more cards right is pretty low: *if* that is what we observed, the p value is 0.056. The frequentist idea is that using probability, we can explain what "better than chance" is with much greater clarity, but it still requires an epistemic judgment: I may decide that witnessing something that only has a 0.056 probability of occurrence is enough to imply a discrepancy between my provisionally chosen hypothesis and experience that I am willing to admit it is wrong. If I decide to follow standard statistical practice, I would need a value below 0.05, in which case I would as a rule reject my initial hypothesis if the subject can get 27 hits, since the probability for that would be 0.03. A parapsychologist who desires her field to be accepted by the community, I suspect, would want to choose a standard that is agreeable to the people she wants to convince—skeptical scientists.

Of course, this point between "better than chance" or not also entirely depends on how long the trial is. In other words, we say that 27 is the cutoff, only because we already decided that the trial involves 100 guesses. For 500 guesses, for instance, there is a probability of 0.95 to randomly guess 115 cards correctly. So, using the same standard, scores higher than 115 would be considered as statistical evidence for the hypothesis the subject actually has ESP.

One study claims to have discovered just that: a parapsychologist carried out the

| Subject | Guesses | Hits | Std. Dev. |
|---|---|---|---|
| M.C. | 1250 | 243 | -0.49 |
| H.F. | 1000 | 219 | 1.50 |
| H.H. | 2000 | 416 | 0.90 |
| A.K. | 1000 | 212 | 1.50 |
| H.S. | 1000 | 195 | -0.39 |
| T.S. | 1000 | 222 | 1.73 |
| J.T. | 1000 | 210 | 0.79 |
| E.S. | 1750 | 370 | 1.19 |
| A.M. | 500 | 118 | 2.01 |
| L.S. | 500 | 113 | 1.45 |
| D.A. | 750 | 168 | 1.64 |
| O.W. | 500 | 114 | 1.56 |

**Table 3.1**: Card-guessing performance of manic-depressive patients (Shulman, 1938, 101)

card guessing experiment on 141 patients in a mental hospital.[4] The study claims to have found statistical evidence that manic-depressive individuals have demonstrated the ability to detect the face of a card through extrasensory means. It is said that these subjects consistently scored higher than chance. For instance, consider patient A. M. who got 118 hits out of 500 attempts. As discussed, it would seem that, assuming A. M. was just guessing, it would have been an extraordinary coincidence that he scored 18 higher than normally expected. In fact, the p value—the probability for an outcome like this or better to happen by chance—is 0.027, which is quite improbable. Does this constitute evidence for ESP?

Feller argues that this result is spurious, because the parapsychologists practiced *optional stopping*.[5] The idea is that many of these experiments have no set number of attempts, and often an experiment could stop either exactly when a favorable result

4. Robert Shulman, "A Study of Card-Guessing in Psychotic Subjects," *The Journal of Parapsychology* 2, no. 2 (1938): 95–106

5. Feller, "Statistical Aspects of ESP," 291-292

is obtained, or a "break" would be taken if the subject is not doing as well as they expect. For instance, an experiment could be terminated early in order preserve a significant result. A. M., for instance, has made 500 attempts. His test was then much shorter than his peers, many of who made more than 1000 guesses. The question is, then had A.M. guessed more cards, would his performance regress toward what would expect from guessing?

Of course, it impossible to ascertain this now. But to make this point more tangible, we can take advantage of the power of modern statistical computing: we can simulate experiments with similar stopping rules, and see if we can get $p$ values below 0.05. The difference here is that for the simulation we *know* that the participants are just guessing, so we know that any statistical significance would be a fluke. That is, if we can still force significance using optional stopping, then there is a good reason to doubt the supposed evidence for ESP gathered using optional stopping. We need to stimulate the following scenario: the experimenter will randomly draw a Zener card out of a shuffled deck, with replacement. She will then ask the subject to guess the card, and record the result. For each subject, she will stop under one of these two conditions: *either* the result has reached significance—the probability of the current outcome is less than 0.05—*or* 2000 guesses have been made. The experimenter will then move on to the next subject, until she has examined 1000 subjects. All subjects are guessing, so their probability of success is exactly 0.2.

More precisely, the simulation consists of the following procedure:

1. The simulation will run 1000 times, each of which represents one subject.
2. For each subject $i$, a random sample will be drawn from a Bernoulli distribution,

with the probability of success $p = 0.2$, to simulate a subject who will randomly guess the face of a randomly drawn Zener card, with replacement.

3. For each draw $k$ (of subject $i$), the result $x_{i,j} \in \{0, 1\}$ will be added to the total sum $y_i$ for the subject.

4. Each trial continues to run, *until* one of the following conditions is obtained on the $n$th guess:

   **significance:** When the $p$ value is less than or equal to 0.05. That is,

$$P(Y \geq y_i) = \sum_{j=y_1}^{n} \binom{n}{j} 0.2^j 0.8^{n-j} \leq 0.05$$

   **end:** $n = 2000$: the subject has made 2000 guesses without significance. In this case, we simply move on to the next subject.

Here's a summary of the result, after simulating a test of 1000 subjects:

1. Starting with 3.2, it is helpful to graphically examine the difference in shapes between trials with a fixed $n = 2000$ number of attempts and trials with optional stopping allowed. In accordance with the central limit theorem, the left hand side forms a nice normal distribution.[6] The right hand side with optional stopping forms a bimodal distribution.

2. As seen in 3.4, 371 out of 1000 outcomes has reached significance (p value $\leq$ 0.05). P values for trials with fixed N (left) are almost uniformly distributed.

6. Morris DeGroot and Mark Schervish, *Probability and Statistics* (Addison-Wesley, 2012), section 6.3

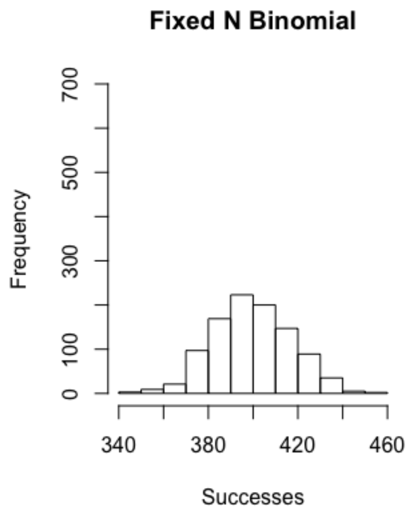


**Figure 3.2**: *Left*: successes out of 1000 random samples drawn from $Binomial(n = 2000, p = 0.2)$. *Right*: successes for all optional stopping simulations.

3. A $p$ value as low as 0.017 was obtained. A distribution of significant p-values with optionally stopping allowed can be see in the right figure in3.4. Simulations that allowed optional stopping clearly have a much higher chance to reach significance, even though we already know the true probability is 0.2.

4. On average, significant results stopped at the 294th attempt. The median is 85. According to figure 3.5, a substantial number of significant results were obtained by stopping before 400 guesses.

5. Successful tests tended to stop early. This makes intuitive sense, since we know that the larger the sample size is, the closer the result will approximate the true proportion. This point will be reexamined in section 3.3, in the context of the argument from intention. There, we will see that the negative binomial distribution, which models trials with no fixed $n$, tend to encourage false positives

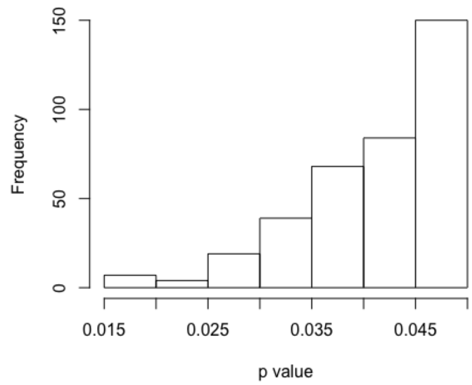


**Figure 3.3**: Histogram of p values($\leq 0.05$)

when the sample size is small.

Thus, using optional stopping, we can easily find 'evidence' for ESP if we look long enough.

Even though Feller was primarily concerned with demonstrating the shoddy experimental practices of parapsychologists, these problems would later be used by Bayesians as part of their critique of frequentism. The basic argument is that parapsychologists could cheat in way they did, because of the way in which the probabilities are computed and interpreted using frequentist methods, and these problems are supposed to avoidable within the framework of Bayesian statistics. In the next section, we will review this argument.

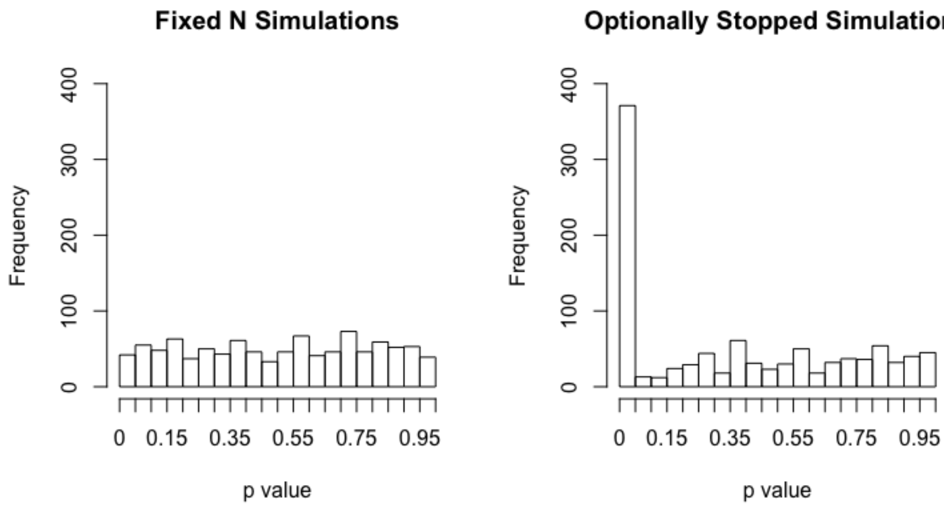


**Figure 3.4**: *Left*: p values for 1000 random samples drawn from $Binomial(n = 2000, p = 0.2)$. *Right*: p values for all optional stopping simulations.

## 3.2 Likelihood and Counterfactual Probabilities

From the problem of optional stopping then, what follows? From the Bayesian perspective, the implication is twofold—a criticism of frequentism and a reason for using Bayesian statistics. The criticism says that the experimenter's intention to stop could directly influence the significance of the result is because of frequentists' reliance on *counterfactual probabilities*. This provides a reason for switching to Bayesian methods, since they, the argument goes, do not require counterfactual probabilities, and, therefore, are immune to the argument from intention. I propose to first examine these two lines of thought, and then transition to the epistemological issues.

Suppose two cards were randomly chosen from the deck, and a subject was able to guess the face of both cards. What is the probability that a subject who is guessing will manage both hits? To begin, there are four ways in which an experiment with

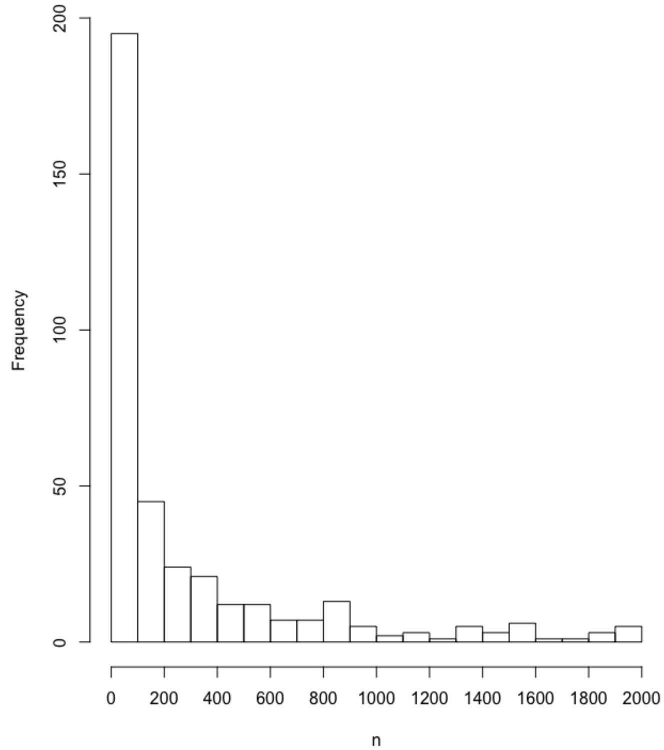


**Figure 3.5**: Histogram of the number of attempts among optionally stopped simulation.

two attempts could have turned out. Let $H$ be "Hit" and $M$ be "Miss". The four possibilities are:

$$MM \quad MH \quad HM \quad HH$$

But to get the probabilities needed, we will have to decide on the probability of a Hit: if we accept *provisionally* that a subject is just making randomly guesses, then the probability of a Hit is 0.2, and the probability of a Miss is 0.8. With these assumptions in place, we can calculate the so-called *sampling distribution*.

| | Factual | Counterfactual | | |
|---|---|---|---|---|
| Hypothesized Reliability | HH | HM | MH | MM |
| $\theta = 0.2$ | 0.04 | 0.16 | 0.16 | 0.64 |
| $\theta = 0.5$ | 0.25 | 0.25 | 0.25 | 0.25 |
| $\theta = 1.0$ | 1.00 | 0.00 | 0.00 | 0.00 |

**Table 3.2**: Comparison between sampling distributions (rows) and likelihood functions (columns)

$$P(HH) = 0.04 \quad P(HM) = 0.16$$

$$P(MH) = 0.2(0.8) = 0.16 \quad P(MM) = 0.8^2 = 0.64$$

This is also represented in the first row of 3.2. The frequentist argument is that if the $p$ value—the probability of the data observed is low enough to be considered statistical significant. Now, Bayesians argue that the $p$ values cannot be seen as an unadulterated summaries of the observed evidence, since any $p$ value is laden with assumptions about what might have happened, based on hypothesized parameters. Consider table 3.2. What frequentists are recommending here is that, when we evaluate a hypothesis, we should consider the probability of the observed data by comparing the relative frequency of its occurrence in comparison to the other values on the same row, which are all calculated based on the assumption that the "random guessing" hypothesis is true. Therefore, when we say that $HH$ is an extraordinary event, we are really saying that it is extraordinary *relative to the three other scenarios that we could see, but did not.*

Bayesians, such as Lindley, criticize the use of probabilities of counterfactual scenarios in the evaluation of statistical hypotheses:

> The usual statistical significance test requires the sample space, or alternatively, the stopping rule to be specified. Many people's intuition says this specification is irrelevant... Of what relevance are things that might have happened, but did not?[7]

Jaynes shares this sentiment:

> The question of how often a given situation would arise is utterly irrelevant to the question how we should reason when it does arise. I don't know how many times this simple fact will have to be pointed out before statisticians of "frequentist" persuasions will take note of it.[8]

How, then, should we look at evidence? The Bayesian perspective focuses *only* on the observed data, but compares its probabilities of occurrence conditional on different hypotheses. The standard Bayesian point of view agrees that the probability of two hits out of two attempts is 0.4, but it disagrees with the notion that the epistemically relevant contrast is the one between different counterfactual probabilities based on one hypothesis of interest. Instead, we should be considering which *hypothesis* would have best predicted what we ended up observing. In the context of table 3.2, the Bayesian way is to look *only* at the column for $HH$, the actually observed result, while ignoring all the counterfactual ones. This is summarized as:

> THE LIKELIHOOD PRINCIPLE[LP]. All the information about $\theta$ [the parameter of interest] obtainable from an experiemnt is contained in the likelihood function of $\theta$ given $X$. Two likelihood functions for $\theta$ (from the same or different experiments) contains the same information about $\theta$ if they are proportional to another.[9]

7. D. V. Lindley and L. D. Phillips, "Inference for a Bernoulli Process (A Bayesian View)," *The American Statistician* 30, no. 3 (1976): 112–119, 114.

8. E. T. Jaynes, "Common sense as an interface," in *Foundations of probability theory, statistical inference, and statistical theories of science*, ed. W. L. Harper and C. A. Hooker (Dordrecht, 1980), 218–257,247

9. James Berger and Robert Wolpert, *The Likelihood Principle* (Institute of Mathematical Statistics, 1988), 19.

Savage's expression of the principle is perhaps more ambitious:

> The likelihood principle says this: the likelihood function... is much more than merely a sufficient statistic, for given the likelihood function in which an experiment has resulted, *everything* else about the experiment—what its plan was, what different data might have resulted from it, the conditional distributions of statistics under given parameter values, and so on—is irrelevant. [10]

We must be clear about the use of the term 'likelihood', since it is technical: it refers to the likelihood function $p(x_{1:n}|\theta)$, which reads: the probability of observations $X_1 ... X_n$ conditional on the parameter $\theta$. $\theta$ in our case is the probability of guessing a card correctly, and, in table 3.2, we are arbitrarily considering a limited set of possible $\theta$s, i.e., $\theta = 0.2, \theta = 0.5, \theta = 1.0$. LP holds that the only epistemically relevant likelihood function is the one represented by the factual column. Here, the result of two hits (HH) is held fixed, while the hypothesized values are varied. In the frequentist method, the hypothesis is held fixed, and the result is varied.

More important, many Bayesians believe that the sort of problem caused by optional stopping can be explained by its violation of LP, because what counts as the set of all possible scenarios—the number of cells in a row—is entirely dependent on the experiment's intention to stop. But it seems peculiar, the argument goes, that the statistical import of data must be dependent on something that goes on the experiment's mind. For instance, take I. J. Good's assessment of the problem:

> Given the likelihood, the inferences that can be drawn from the observations would, for example, be unaffected if the statistician arbitrarily and falsely claimed that he had a train to catch, although he really had decided to stop sampling because his favorite hypothesis was ahead of the game. (This might cause you to distrust the statistician, but if you believe his observations, this distrust would be immaterial.) On the other hand, the "Fisherian" tail-area

10. Leonard J. Savage, "The foundations of statistics reconsidered," *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* 1 (1961): 575–586, 583.

> method for significance testing violates the likelihood principle because the statistician who is prepared to pretend he has a train to catch (optional stopping of sampling) can reach arbitrarily high significance levels, given enough time, even when the null hypothesis is true.[11]

At this point, the debate becomes quite messy, since Bayesians tend to run the problem of optional stopping and the likelihood principle together, as Good has clearly done in the passage above. The assumption is that optional stopping cannot occur once Bayesian methods are adopted. However, the problem of optional stopping is perfectly intelligible on frequentist ground: it draws out undesirable consequences based on assumptions accepted by frequentists. The introduction of LP, however, begs the question against the frequentists. If all the fundamental frequentist methods violate LP, why would any frequentist accept this principle? My suspicion is that they probably won't, so the introduction of LP muddles the water.

However, what could be motivating Bayesians to see LP as being indispensable is that it justifies statistical inferences to be made without taking into consideration the intentions of the experimenters. This version of the problem is summarized as *the argument from intention*, which holds that what optional stopping shows is that Frequentist's reliance on counterfactual probabilities renders their result vulnerable to manipulation, because the experimenter's intention *alone* can radically alter the import of the evidence.

11. Irving J. Good, *Good Thinking: The Foundations of Probability and its Applications* (Univ Minnesota Pr, 1983),135

## 3.3 Intentions and Self-Sabotage Redux

### 3.3.1 Lindley's Argument from Intention

Consider a simple illustration concerning the bias of a coin discussed by Lindley and Phillips.[12] Suppose I was told that the coin was tossed 12 times but out of those times 3 turned up heads. The argument from intention says that, unless you know what goes on inside the tosser mind when she decided to stop the tossing, there is no way to know what the evidence says. And, depending on the answer she gives, the evidential import of the result can alter drastically. For instance, consider these stopping rules:

1. Stop after 12 tosses
2. After 3 heads.

Depending on which of the above rules the parapsychologist used, the significance of the data will change, even if the number of guesses and hits are the same. To begin, note that each rule implies different impossibilities. For instance, if the experimenter stops after 12 tosses, it is obviously impossible for the test to last for more than 12 tosses, but it is possible to get anything from 0 to 12 heads. On the other hand, if the test terminates after 3 heads, the only possible number of heads is 3, but the experiment can take as many tosses as needed to reach that goal. So each different stopping rule implies different counterfactuals, and leads to different sets of probabilities.

So Bayesians have an important point here: intentions are influencing the statistical result through counterfactual probabilities, but, unless there are reasons to think

12. Lindley and Phillips, "Inference for a Bernoulli Process (A Bayesian View)," 113–14.

accounting for intentions is somehow inherently bad, this *supports* the deliberativist thesis: it argues that deliberating on when to stop based on what would happen after a possible course of action is one of the judgments that shapes the epistemic context, in which the evidential import of the data should be interpreted. Consequently, I am not so much concerned with *solving* the problem by arguing that intentions do matter, but rather with *dissolving* it by making sense of its role in induction.

To begin, let us spell out what the argument from intention is. Consider the rule that says to stop after $n = 12$ tosses. Using the frequentist method means that we have to consider the probability of all possible outcomes—0-12 heads, in various combinations. We know that, from probability theory, for a random variable with binary outcomes—success or failure, for instance—the probability of getting $k$ success out of $n$ trials, given the probability of a single success, is

$$\binom{n}{x} p^x (1-p)^{n-x}$$

Let's say a "success" is a coin toss that lands on heads. To carry out an investigation, we have to make two decisions: the first is to choose a hypothesis about $p$ to be tested. In a binomial process with $n = 12$, assuming that the coin's probability of landing on heads is 0.5, what is the probability that she gets 3 or less heads? That is, let $Y = \sum_i^{12} X_i$, where $X_i = 1$ if the coin lands on heads, and 0 otherwise, then

$$P(Y \leq 3) = \sum_{i=0}^{3} \binom{12}{i} (0.5)^i (0.5)^{12-i} = 0.07$$

This is generally considered an insignificant result, but what if the intention was to stop whenever the subject has gotten 3 heads? To model this, we would have to use the so-called negative binomial distribution, which models the probability of making $r$ failures before getting $k$ successes. In this case, the experimental question is in fact quite different, since now we would consider the coin to be biased against heads if it takes an abnormally large number of tosses. So the statistical question is: what is the probability of the coin needing 12 or more tosses in order to get 3 heads? Using a variant of the binomial distribution called the *negative binomial distribution*, we can easily find this. Let $X$ be the number of misses, so

$$P(X \geq 9) = 1 - P(X < 9) = 1 - \sum_{i=0}^{8} \binom{3+i-1}{i} (0.5)^3 (0.5)^i = 0.03$$

This seems to be a much more significant result. It seems like we were able to raise the significance of the data by changing nothing but the stopping rule, but this is not the complete picture.

### 3.3.2 The Deliberativist Analysis

From the deliberativist perspective, the effect of changing stopping rules is a demonstration of the two key ideas of my position: the Peircean idea that decisions made in abduction contextualize inferences made in induction and the voluntarist ideal of rationality as self-controlled deliberate conduct. According to view I defend, the inferential force of induction is at least partly underwritten by the decisions made during the abductive context. The intention to stop is one such decision. Since the commitment to uphold those decisions is required for the inductive inferences to be considered as valid, changing it afterward is an act of self-sabotage: it defeats the very purpose of trying to probe the hypothesis accepted provisionally.

As discussion in 2.5, the abductive context is characterized by the freedom of thought to propose and accept a hypothesis provisionally. There is no context-independent justification for the choice of $n$, which determines the probability of having $k$ successes, but whatever decision one makes has a direct repercussion on how the inference will be made in the inductive context. For instance, suppose I choose $n = 5$, and that I decide on the hypothesis that $p = 0.5$. There decisions will put a obligatory constraint on my behavior during the inductive stage, such as when to stop the experiment, and how to revise my belief in light of the evidence. We can, however, explicate these obligations that these potential decisions imply by engaging in a deductive interrogation, since many of these obligations follow as a necessary consequence of the potential model. For instance, if I choose $p = 0.5$ and $n = 5$, I can deduce that

$$P(X = 0) = \binom{5}{0} 0.5^0 (1 - 0.5)^{5-0} = 0.031$$

$$P(X \leq 1) = \sum_{i=0}^{1} \binom{5}{i} 0.5^i (1 - 0.5)^{5-i} = 0.19$$

These are the probabilities that follow deductively from the decisions I have made in the abductive context. I made what Peirce would call a *probable deduction*, which involves the deductive derivation of probabilistic judgments based a model with known parameters. Even though the conclusion of probable deduction is probabilistic, the *connection* between the conclusion and its premises is necessary.[13] They signify the epistemic obligation I incur: if I accept provisionally the hypothesis that $p = 0.5$ for

13. Peirce, "A Theory of Probable Inference," 417.

$n = 5$, then I am committed to the probabilistic judgment that the probability of tossing the coin five times without heads is 0.031.[14]

As Peirce points out, inference made in the inductive context presupposes the experimenter to hold certain values, including

> first, a sense that we do not know something, second, a desire to know it, and third, an effort,—implying a willingness to labor,—for the sake of seeing how the truth may really be.[15]

We can understand the problem of optional stopping as a special case of *self-sabotage*. The very point of experimentally testing a hypothesis is to see how a it can withstand the deliverance of experience, so knowingly running a test that would likely to give a wrong answer is preemptively sabotaging one's possibility of "seeing how the truth may really be". As an analogy, consider making a promise, a speech act we discussed in section. 2.1 The normative point of making a promise is to express to the promisee your intention to undertake a certain obligation. [16] Because of this, I would be sabotaging myself if I make a promise that is clearly disingenuous. Making a promise I clearly cannot keep is clearly one example, but a more apt analogy would be a promise made in bad faith. For example, consider van Fraassen's example of making a promise with a vacuous conditional statement:[17]

I promise that if I see Santa Claus, I will ask him to bring you a bicycle. (3.1)

14. Of course, I could be dissatisfied by the result of the deduction, in which case I could revise my experimental commitment abductively. I put this issue aside now, as the dynamic between abduction and deduction is the focus on chapter 4.

15. Peirce, *The Essential Peirce, Volume 2: Selected Philosophical Writings (1893-1913)*, 48.

16. Searle, *Speech Acts: An Essay in the Philosophy of Language*, 58-59

17. Bas van Fraassen, *The Empirical Stance* (Yale University Press, 2002), 243

Such a promise is made in bad faith, because the promiser can never break them. It defeats the purpose of the promise to express an intention, since it's phrased in such a way that the promise is unlikely to be perceived as being sincere. Of course, there is nothing self-sabotaging about this example in and of themselves, but one would be if they are genuinely trying to make a promise. It would be like a person saying to reassure her partner by saying that "I promise I will *try* to be faithful". Her partner will hardly be reassured. Analogously, the goal of experimentally probing a hypothesis is to produce evidence that will be accepted by one's epistemic community. The possibility of this acceptance, however, depends on whether or not the agent has genuinely given the hypothesis a chance to face the tribunal of experience, and this is sabotaged by optionally stopping.

This idea can be accounted for by appealing to the frequentist idea of *error probabilities*. For instance, suppose I follow the standard practice and declare that I will reject the hypothesis that $p = 0.5$ when the sample I collect has a less than the probability of 0.05 of occurrence. This would mean that I am committed in rejecting my hypothesis if I get no heads after tossing the coin 5 times. But what if, while the coin was actually biased, it was not *so* biased that it will not land on heads at all? The probability of making such an error can be calculated *ex ante* deductively. For instance, suppose the reality is that the coin is biased such that $p = 0.2$. But if this were true, I have a high probability of keeping my provisionally accepted hypothesis by mistake, because

$$P(X > 0) = 1 - P(X = 0) = 1 - \binom{5}{0} 0.8^5 = 0.67$$

That is, it is improbable even for a biased coin to have 5 out of 5 heads. This means that by accepting $p = 0.5$, I am incurring a pretty high risk of error: if $p = 0.2$, there is a 0.67 probability that my decision would be a wrong one. In fact, the closer $p$ is to 0.5, the more likely it is that I will come to accept $p = 0.5$ erroneously. If $p = 0.3$, for instance, this error probability is 0.83. [18]

Thus, Lindley is entirely correct in pointing out that under different stopping rules would have an impact on what will counted as statistical significance, even if the numerical result will be the same, but from the deliberativist standpoint, these stopping rules imply different sets of epistemic commitments, that is, the experimenter incurs different degrees of the risk of being wrong as indicated by the relevant error probabilities, depending on the decision rules she accepts in the context of abduction.

If the agent intends to stop after 12 trials, then to aim for a level of statistical significance $\alpha$ at 0.05, she would have to commit to reject the fair coin hypothesis if she gets 2 or less heads. The repercussion is that, had the coin been only slightly biased against landing on heads, it is unlikely that she would be able to find the truth. For instance, the probability of getting 2 or less heads, if $p = 0.4$, is only 0.08. For $p = 0.3$, it is 0.25. In other words, if the coin were only slightly biased, it would be unlikely to produce a result that is detectable within this particular deliberative framework. Of course, there is nothing sacred about $\alpha = 0.05$, notwithstanding the preaching of introductory statistics textbooks. If the agent's intention is to determine if the coin is only slightly biased, she is free to adjust $\alpha$ so that her risk of error, had $p$ been 0.4, is smaller. Even though, if $p = 0.5$, the probability of getting 3 or fewer heads out of 12 at 0.07 does not quite reach the textbook standard of statistical

18. Of course, the risk of error is a concern only if I care about avoiding errors and finding the truth. If I care more about just proving my favorite hypothesis, I might just welcome the fact that I get to keep my hypothesis if it was the wrong one. At that point, however, the agent is hardly operating in the context of induction.

significance, it nevertheless would be much better at detecting $p = 0.4$, since the probability of the same outcome occurring would be 0.23. Not perfect, but this is the kind of decision one makes during the abductive context.

None of the above considerations hold if we had changed the stopping rule to "stop after 3 heads." The deliberative framework would be entirely different. To begin, we are now adjudicating not the probabilities of error between different numbers of heads landed, but the number of tails we would tolerate before 3 heads is seen. We saw that having to see 9 tails before 3 heads is a statistically significant enough reason to reject the hypothesis that the coin is fair. What this overlooks, however, is that a biased coin can often get 3 heads before 9 tails; because, probabilities from a negative binomial distribution tend to be "front-loaded". For instance, with a coin that is half as unlikely to land on heads than tails, that is, $p = 0.25$, the probability to see 8 or less tails before 3 heads is

$$P(X \leq 8) = \sum_{i=0}^{8} \binom{3+i-1}{i} (0.25)^3 (0.75)^i = 0.55$$

So looking strictly at the different p-values is a somewhat misleading way to look at the matter. Mayo and Spanos summarizes the frequentist response as follows:

> [the argument from intention] would seem to beg the question against the error statistical [i.e., frequentist] methodology which has perfectly objective ways to pick up on the effect of stopping rules: far from intentions "locked up in the scientist's head" (as critics allege), the manner of generating the data alter error probabilities...[19]

Of course, a defense of intentions in frequentism is not an argument *for* the rel-

19. Deborah G. Mayo and Aris Spanos, "Error Statistics," in *Handbook of the Philosophy of Science, Vol. 7: Philosophy of Statistics*, ed. Prasanta S. Bandyopadhyay and Malcolm R. Forster (Elsevier B.V., 2011), 186.

evance of intention in Bayesianism. Furthermore, Bayesians like Lindley without a doubt were aware of these basic statistical facts from power analysis. What prompted their stance is the assumption that Bayesian methods are impervious to the problem of optional stopping, since the likelihood function is not affected by intentions, nor other facts in the deliberative framework. In the next section, I will demonstrate that optional stopping can also affect results obtained using Bayesian methods, so I am attacking the very idea that Bayesians cannot ignore the effects of deliberation.

## 3.4 Bayesian Optional Stopping

Deborah Mayo suggests that using early stopping to manipulate experimental results is also possible using Bayesian methods, despite the Likelihood Principle.[20] Mayo's argument appeals to the fact that Bayesian methods can obtain the same or similar result when using a flat/uninformed prior, so this opens the door to early stopping using the same method. I will try to get more mileage out of this argument by using simulations.

To begin, we must gain a basic understanding of what Bayesian inference is, which, naturally, begins with Bayes' theorem. Consider some hypothesis or belief $H$ and some evidence $E$.

$$P(H|E) = \frac{P(H)P(E|H)}{P(E)}$$

In its most basic form, Bayes' theorem has three components: The unconditional

20. Deborah Mayo, *Error and the Growth of Experimental Knowledge* (University of Chicago Press, 1996),352-353

probability of $H$, $P(H)$, represents the probability we would initially assign to the belief. Second, the evidence, as we have discussed, is represented by the likelihood $P(E|H)$—the probability of the evidence, given the hypothesis is true. The third ingredient is $P(E)$, the unconditional probability of $E$. To see how this works, consider an example with Zener cards. To begin, suppose that we have a subject in front of us, and we have to determine if she has ESP. Let's say she has 2 out of 2 correct answers. How and what should we learn from this data? What follows is the standard Bayesian story.

For the sake of simplicity, let us suppose for now that there are only two options: either the subject is randomly guessing, or she has ESP, which entails a perfect reliability. In other words, we have two hypotheses. Let $\theta$ be the subject of probability of getting a hit, and

1. $H_0 : \theta = 0.2$
2. $H_1 : \theta = 1$

These are sometimes called "chance hypotheses". Now let $E_i$ refer to the result of the $i$th guess, such that $E_i = 1$ for a hit, and 0 otherwise. So let $E = \sum_i^2 E_i = E_1 + E_2 = 2$. This means that we have the following likelihoods:

3. $P(E|H_0) = 0.2^2 = 0.04$
4. $P(E|H_1) = 1$

Recall that the likelihood principle says that this contains all the information we need to know about the experiment. Now, suppose you are not a believer of ESP, so you are almost certain—say, 99% sure—that the subject will not do better than chance. We then have the priors needed:

5. $P(H_0) = 0.99$
6. $P(H_1) = 0.01$

From the above, we can derive:

$$\begin{aligned} P(E_i = 1) &= P(H_0)P(E_i = 1|H_0) + P(H_1)P(E_i = 1|H_1) \\ &= 0.99(0.04) + 0.01(1) \\ &= 0.0496 \end{aligned}$$

Using Bayes' theorem, we can then revise our belief about the subject's ability to guess cards, producing the following *posterior probabilities*:

$$P(H_0|E) = \frac{0.99(0.04)}{0.0496} = 0.8$$
$$P(H_1|E) = \frac{0.01(1)}{0.0496} = 0.2$$

Having seen two successful attempts in a row, we have warmed up to the idea that the subject might have ESP. An intuitive way to look at this Bayesian procedure is that the posterior probability is a promise between my existing belief—my priors—and evidence, which is summarized by the likelihoods, according to the LP.

To make my point, these basic Bayesian statistical procedures are sufficient, though things will get somewhat messy when we consider more realistic cases. For instance, it is arbitrary to consider only two chance hypotheses. A more applicable model would be to consider all possible values of $\theta$ in $[0, 1]$. For this we have to use some of the well-established distributions. I will again use simulation to demonstrate the effect of optional stopping, but to do so I need to first explain how the situation will be modeled.

Recall that optional stopping from a frequentist context entails falsely rejecting the null hypothesis by sampling over and over again until obtaining an outcome with a probability low enough on the null hypothesis to secure statistical significance. The Bayesian parallel is to continue sampling so we can have $E$ such that $P(H_0|E) < x$ where $x$ is a value that the optional stopper is committed to believing. Note that now we concern ourselves with the probability of the hypothesis itself, whereas in the frequentist setting we were concerned with the probability of the observation.

Fortunately, since there are only two outcomes, a Zener card-type experiment can be modeled as a Beta-Bernoulli process, where the Beta distribution would model our degrees of belief about a subject's *propensity*, and the Bernoulli distribution would represent the Zener card experiment itself. What these models represent is usually clear enough in a practical and statistical setting, but since we are doing philosophy, we need to be clearer about what we mean by degrees of belief and propensity, so that we are clear about the phenomena being modeled.

I suggest we follow the views of D. V. Lindley and David Lewis. Lindley argues that probability is a relation between the agent and the world, so when we say $P(\theta = 0.5)$, it represents our epistemic judgment about some physical event $\theta$.[21] In our case, this has to be an objective reliability of the subject's ability to discern the face of the card, which is a property in the world: even though $\theta$ looks like a probability, in the Bayesian statistical framework we can just treat it as another parameter being modeled, not unlike $\mu$ or $\sigma$ for normal distributions, so a subject's extra-perceptual reliability is an objective feature of the world in a way no different than the fact that

21. Lindley and Phillips, "Inference for a Bernoulli Process (A Bayesian View)," 115.

the average age of Duke students is an objective fact. The degree of beliefs about them, according to the voluntarist conception I accept, is an epistemic judgment made about this objective fact.

This recommendation is compatible with, if not the same as, the influential view presented by David Lewis, who adopts Carnap's pluralistic stance on probability. Carnap thinks there are at least two concepts of probability: *probability*$_1$, which is an epistemic concept about degrees of confirmation and *probability*$_2$, which refers to the empirical concept of long-run limiting frequencies.[22] Lewis suggests that we should instead interpret the epistemic concept as credence or degree of belief and the empirical concept as chance or propensity.[23] So, following Lewis, we can interpret $P(\theta = 0.5) = x$ to be "the degree for the belief that the chance of heads is 0.5 is $x$." For the sake of consistency, I will refer to subjective probability just as *credence* or *degrees of belief*, and objective probability as *chance* or *propensity*.

We can now spell out the type of experiments to be simulated.

Early on, we considered a case in which only two possible hypotheses are considered: either the subject is guessing randomly($H_0 : \theta = 0.2$) or the subject has perfect reliability($H_1 : \theta = 1$). This makes our epistemic attitude relatively easy to summarize, since all we have to do is to assign a value to our credence on each of the two hypotheses. As we noted, this is an oversimplification, since there is no reason to arbitrarily restrict ourselves to just two hypotheses. This, however, means that we need a way to deal with the fact that there are infinitely many possible hypotheses between 0 and 1, which is why we need the beta distribution.

22. Rudolf Carnap, *Logical Foundations of Probability* (University of Chicago Press, 1950), 517.

23. David Lewis, "A Subjectivist's Guide to Objective Chance," in *Philosophy of Probability: Contemporary Readings*, ed. Richard C. Jeffrey (University of California Press, 1980), 83–132.

The beta distribution is really nothing but a function that, based on two parameters we provide, describes our epistemic attitudes toward $\theta$.[24] The two parameters, $\alpha > 0$ and $\beta > 0$, can be thought of as, in our context, our past experience about $\theta$, with $\alpha$ representing past successes and $\beta$ past failures. For instance, if we set $\alpha = \beta = 1$, it should say that we are extremely ambivalent about $\theta$. In fact, it is equivalent to having a uniform distribution over $[0, 1]$—this means that I am utterly indifferent regarding any value for $\theta$.

Our data-collection will be modeled using the binomial distribution. Let $x$ be the number of success, $n$ the number of trials, and $\theta$ the propensity of success:

$$f(x|\theta, n) = \binom{n}{k}\theta^k(1-\theta)^{(n-k)}$$

This is the same distribution we used as the sampling distribution in the frequentist case, but recall that for Bayesian analysis we will no longer concern ourselves with counterfactual probabilities, instead, we are treating $\theta$ as the function of the $x$, the number of success which is constant.

Fortunately, as soon as this is laid out, the rest is very simple, thanks to the fact that the beta distribution is a *conjugate prior* for the binomial distribution.[25] Essentially, what this means is just that if we plug the beta and binomial distributions into Bayes' theorem to get a posterior distribution

$$p(\theta|x) = \frac{p(\theta)p(x|\theta)}{\int p(\theta)p(x|\theta)}$$

24. The distribution has the form: $\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$ where the parameters $\alpha, \beta > 0$ and $0 \leq x \leq 1$($x$ is the random variable being modeled)

25. DeGroot and Schervish, *Probability and Statistics*, section 7.3

the result is simply another beta distribution with parameters $\alpha = \alpha + x$ and $\beta = \beta + n - x$. In words, to learn from experience, all we have to do is to add the number of successes to $\alpha$ and the number of failures to $\beta$. Another useful thing to keep in mind is that the beta distribution's expected value has the form:

$$E(\theta) = \frac{\alpha}{\alpha + \beta}$$

Now, thanks to conjugacy, the *posterior* expected value is simply:

$$E(\theta) = \frac{\alpha + k}{\alpha + \beta + n}$$

So, our experimental procedure is fairly simple: we pick an appropriate set of parameters, and for each trial in which the subject is able to guess the card correctly, we add 1 to $\alpha$; otherwise, we add 1 to $\beta$. For example, consider again the case of a skeptic who observed that a subject has made two correct guesses in a row. Since prior to the observation, the skeptic does not believe that the subject would do better than chance, she can decide her existing opinion is such that

$$E(\theta) = \frac{\alpha}{\alpha + \beta} = \frac{1}{5}$$

There are various ways in which we can make this work mathematically, but for now let's say $\alpha = 2$ and $\beta = 8$.[26] Since the subject has gotten 2 out of 2 correctly, the skeptic's posterior should be the beta distribution with $\alpha = 2 + 2 = 4$, while $\beta$ remains at 8.

Because we are using Bayesian methods, we can ask directly the probability of $\theta$

26. The epistemological relevance of the choice of these parameters will the discussed in chapter 4.

having certain values. A similar question we can ask, then, *given* the evidence we have, what is the probability of the subject's para-perceptual reliability is no better than randomly guessing? In other words, what is the probability that $\theta$ is less than or equal to 0.2? Using the cumulative distribution function for the beta distribution provided by computer programs, we can find out the prior and posterior values:

$$P(\theta \leq 0.2) = 0.56$$

$$P(\theta \leq 0.2|\mathbf{X}) = 0.16$$

We can see that after witnessing the evidence, the skeptic's personal probability for the belief that the subject is doing no better than chance is lowered by quite a bit. The Bayesian version of optional stopping is this: a Bayesian optional stopper can decide to stop gathering more evidence as soon as the posterior is low enough. The idea is that a committed enough optional stopper will eventually find "evidence" for ESP, i.e., subjects with low posterior probability of random guessing.

To simulate this procedure, we will carry out the following:

1. For each simulation in $n$ times:

   a. For each subject $s$, we begin with a flat uniform distribution by setting $\alpha = \beta = 1$.

   b. A random sample $x_i$ will be drawn from a Bernoulli distribution, with $\theta = 0.2$.

   c. Add 1 to $\alpha$ if $x_i = 1$; add 1 to $\beta$ otherwise.

   d. Terminate if either (i) $P(\theta_s \leq 0.2|\mathbf{X})$ is less than threshold $k \leq 0.05$ or

(ii) the number of trial $i$ has exceeded the maximum number $m = 2000$. Otherwise, return to step a with the same subject.

2. If $n$ subjects have been tested, terminate; else, return to step a with a new subject.

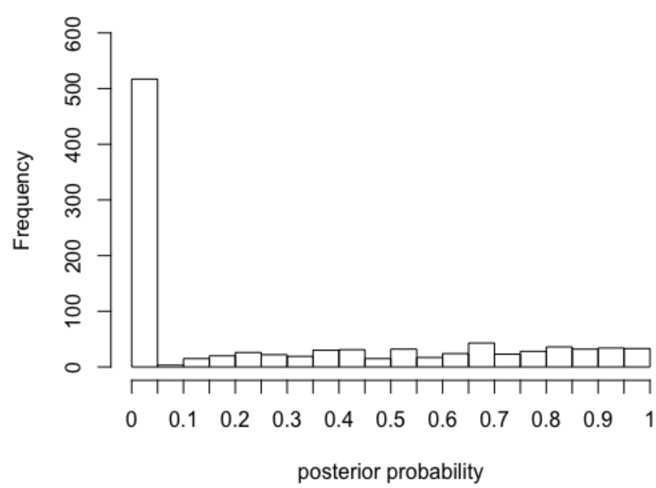


**Figure 3.6**: Histogram of $p_i(\theta \leq 0.2|x, \alpha = 1, \beta = 1)$

$n = 1000$ stimulations were carried out. We clearly see the effect of early stopping here. In fact, using a flat prior, $\alpha = \beta = 1$, 517 out of 1000 stimulations were stopped early due to reaching a low enough posterior probability for $p(\theta \leq 0.2)$, which is more than the number we had for frequentist early stopping. The distribution of the posterior probabilities can be seen in figure 3.6. This is partly due to the fact that, when starting from a flat prior, the experiment can stop after 1 guess if the first guess is the correct one, because with $\alpha = 2, \beta = 1$, the posterior probability for $P(\theta \leq 0.2) = 0.04$, which is already lower than the threshold designated. 208 out of

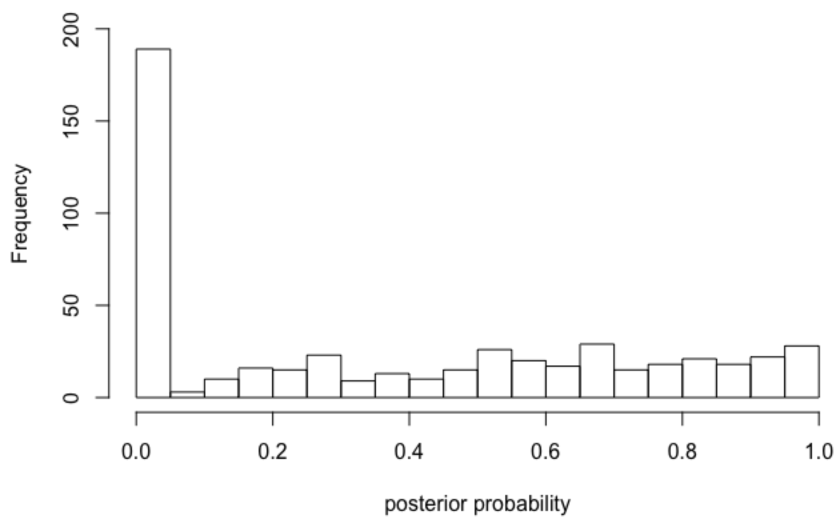


**Figure 3.7**: Histogram of random samples from $p_i(\theta \leq 0.2|x, \alpha = 1, \beta = 4)$

1000 simulations ended this way, which is exactly what was expected, since we are drawing from a Bernoulli distribution with the probability 0.2 of success.

To eliminate that particular result, we can use a weakly informed prior by setting $\alpha = 1, \beta = 4$. The result is summarized in figure 3.7. Since the initial expected value is 0.2, and this means that for the cases in which early stopping is successful, we have essentially manipulated ourselves from the right opinion into the wrong one. The result is nearly identical to the frequentist case, with 367 out of 1000 results being stopped early to get the posterior probability desired.

Based on the above, one may argue that optional stopping could be further prevented by using a even more strongly informed prior. This is true: for instance, if, instead of having $\alpha = \beta = 1$ as parameters, we use something strongly biased in favor of $\theta = 0.2$, such as $\alpha = 10, \beta = 40$, it would be fairly difficult for the optional stopper to game the result. But this makes sense only because we *know* what the true distribution is. This also seems to me a point *for* deliberativism, not against it, because this amounts to saying commitments and intentions matter by sneaking them in through the back door of priors. Furthermore, what is stopping an optional stopper from cheating even more by adopting a set of parameters that is biased *against* $\theta = 0.2$? Of course, we would criticize anyone that adopts such an experimental stance, but it would be made on the deliberativist ground: the standard of criticisms does consider intentions and decisions as relevant information.

## 3.5 Is Optional Stopping Irrational From the Perspective of the Utility Maximizer?

There is, however, a more substantive objection that requires a more thorough exposition. An Orthodox Bayesian could argue that optional stopping is an irrational practice, because from a Bayesian perspective it is never rational to refuse evidence, because additional evidence *always leads to an increase in expected utility.* This is in fact a result that has been proven in various occasions and forms, first by Frank Ramsey, and then later, independently, I. J. Good, and L. J. Savage.[27][28][29] Skyrms has also discussed Ramsey's unpublished note in details.[30]

27. Frank P. Ramsey, "Weight or the Value of Knowledge," *British Journal for the Philosophy of Science* 41, no. 1 (1990): 1–4

28. Irving J. Good, "On the Principle of Total Evidence," *British Journal for the Philosophy of Science* 17, no. 4 (1966): 319–321

29. Savage, *The Foundations of Statistics*, sec 6.2.

30. Brian Skyrms, *The Dynamics of Rational Deliberation* (Harvard University Press, 1990), chapter 4.

Some context is helpful. In his *A Treatise in Probability*, J. M. Keynes points out that subjectivists and expected utility theorists often implicitly assume that we should always get more evidence. Bernoulli, for instance, suggests that rationality demands the utilization of all evidence available to us. This implies, Keynes thinks, that it's always rational to get more evidence, but then it raises another critical question about whether or not one could ever be rational in refusing new evidence.[31] If the answer for the former question is positive, and the latter question negative, then we have to conclude that rationality dictates that we should never stop looking for more evidence. Keynes remains ambivalent about the notion itself. Ramsey, who clearly read *A Treatise*, responded to the problem in an unpublished note.

Many years later Ayer raises the same question in response to Carnap's *Logical Foundation of Probability*, in which Carnap essentially restates Bernoulli's maxim as "the requirement of total evidence".

> *Requirement of total evidence*: in the application of inductive logic to a given knowledge situation, the total evidence available must be taken as basis for determining the degree of confirmation.[32]

Ayer asks the Keynesian question: should "total evidence" include relevant evidence that I do not yet have in possession?[33] If finding the truth value of some proposition $P$ could potentially sway the balance of my evidence, then it seems that I definitely acquire it. Thus the principle of total evidence seems to suggest that I am also rationally compelled to consider some evidence I do not yet have. I. J. Good interprets Ayer's as questioning "why... we should bother to make new observa-

31. Keynes, *A Treatise on Probability*, 84–85.

32. Carnap, *Logical Foundations of Probability*, 211.

33. A. J. Ayer, *Probability and Evidence* (Macmillan, 1972), 56.

tions."[34] In the context of optional stopping, this is particularly salient: if I already have the result I want, why should I bother get more evidence?

Ramsey interpreted Keynes' question in the same way and addressed it an unpublished note from an expected utility perspective. Ramsey's argument is roughly that, *if* we assume an agent to be a perfect Bayesian and that new information does not cost anything, then she will never be worse off getting new evidence. In fact, she is guaranteed to be *better* off as long as the new evidence will tell her something new. A perfect Bayesian agent is someone who studiously updates her opinions based on Bayes' rule and then acts by choosing the action that maximizes her expected utility. Note that this assumes a few things: first, for any decision problem she faces, there is always going to be at least one option that maximizes her expected utility. Second, this assumes that she can always assign precise probabilities to the options she faces, because, as Good later realizes, if the agent's judgments of probability are vague, that is, she may judge the probability of an event to fall within a range of values, instead of one precise numerical value, then it is possible that confounding observations could muddle the water by widening this range, so that she could be worse off by getting more evidence.[35] Third, as Skyrms points out, this also implies that the agent knows that she will always *remain* perfectly Bayesian in the future. What we have here, then, is the ideal Bayesian agent, someone who revises her opinions strictly with an explicative rule.

I will make use of an intuitive example rather than reproducing the proof here.[36]

34. Good, *Good Thinking: The Foundations of Probability and its Applications*, 178.

35. ibid., 181-182

36. This example is adapted from Isaac Levi, "The Weight of Argument," in *Fundamental Uncertainty: Rationality and Plausible Reasoning*, ed. Silva Marzetti Dall'Aste Brandolini and Roberto Sczzieri (Palgrave MacMilan, 2011), 39–58, 49.

Suppose we have three hypotheses about the content of an urn in front of us:

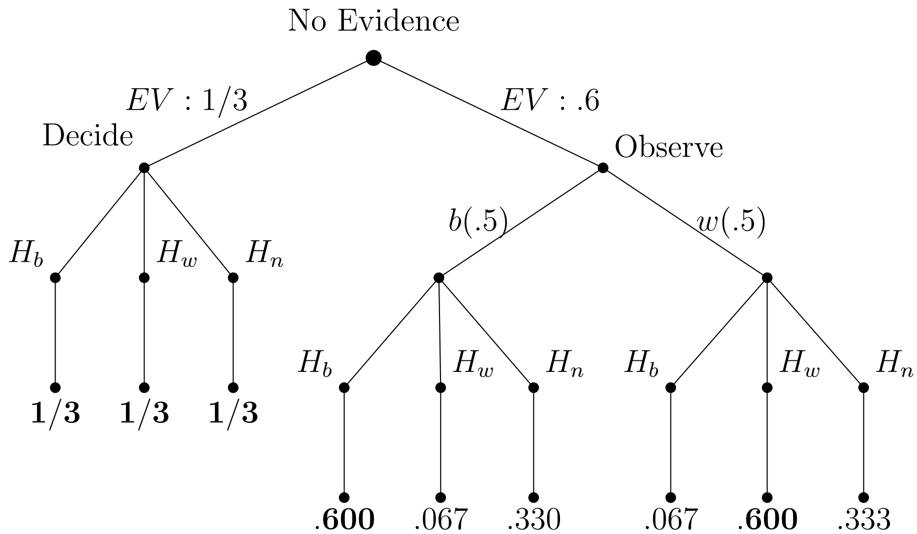


**Figure 3.8**: Decision tree for no existing evidence(uniform prior).Utility maximizing options are bolded.

1. $H_b$: 90 black balls and 10 white balls
2. $H_w$: 10 white balls and 90 black balls
3. $H_n$: 50 white balls and 50 black balls.

Suppose we start by assuming $P(H_b) = P(H_w) = P(H_n) = 1/3$. There is a reward of \$1 for picking the correct hypothesis. Our expected payoff for choosing each hypothesis would be the same at 1/3. Nevertheless, we are allowed to sample with replacement as many times as we wish. Should we get more evidence? Yes, according to Ramsey, we should, and this can be demonstrated in an expected utility analysis.

This situation is represented as a decision tree in figure 3.8: at the beginning, the probability of getting a black ball is the same as getting a white ball. Let $E_b$ be "a black ball is drawn" and $E_w$ for white balls. So:

$$\begin{aligned} P(E_b) &= P(H_b)P(E_b|H_b) + P(H_w)P(E_b|H_w) + P(H_n)P(E_b|H_n) \\ &= \frac{1}{3}(.9) + \frac{1}{3}(.1) + \frac{1}{3}(.5) \\ &= .5 \\ P(E_w) &= 1 - P(E_b) \\ &= .500 \end{aligned}$$

So, in the event of drawing a black ball from the urn, we would update our belief like so:

$$\begin{aligned} P(H_b|E_b) &= \frac{P(H_b)P(E_b|H_b)}{P(E_b)} \\ &= \frac{1/3(.9)}{.5} \\ &= .600 \end{aligned}$$

Similarly, by applying the calculation to the other hypotheses, we get:

$$P(H_w|E_b) = .067$$

$$P(H_n|E_b) = .333$$

Similar arguments can be made by assuming $E_w$, that is, a white ball is chosen. In that case $P(H_w|E_w) = .600$. If we were an ideal Bayesian agent, we should pick $H_b$ if $E_b$, and pick $H_w$ if $E_w$. Since an ideal Bayesian would choose the option that maximizes our expected utility, in either case the expected value after drawing from the urn once is .600, which is an improvement, since before drawing our expected utility is 1/3 for all options. The net gain in expected utility from getting new information would be $.60 - .33 = .27$. This is referred to as *the value of information* in the literature.[37][38] Interestingly, Ramsey seems to think that the value of information or evidence captures what Keynes calls the *weight* of evidence—see chapter 4.

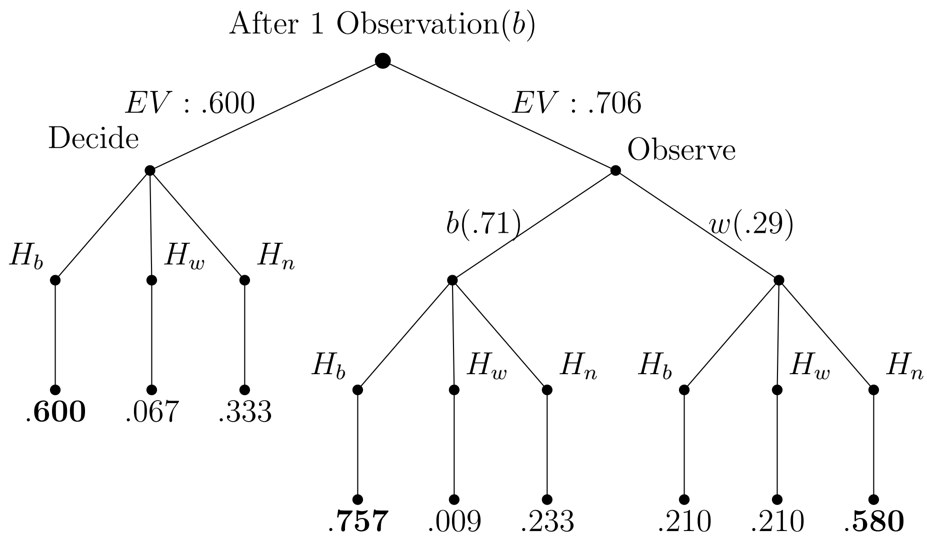


**Figure 3.9**: Decision tree after observing one black ball). Utility maximizing options are bolded.

It turns out that we would be even better off if we were to draw from the urn again. This situation is represented as a decision tree in figure 3.9: suppose the first draw yields a black ball. So now we have one piece of evidence in hand. Let us refer to our state of belief after the first draw as $H'_b, E'_b, ..$ and so on. For instance, $P(H'_b) = P(H_b|E_b)$ and $P(E'_b) = P(E'_b|E_b)$. One notable change is that $P(E'_b) = .71$

37. Howard Raiffa and Robert Schlaifer, *Applied Statistical Decision Theory* (Harvard, 1964), 89–90

38. Robert Winkler, *An Introduction to Bayesian Inference and Decision* (Probabilistic Publishing, 2010), section 6.3.

and $P(E'_w) = .29$. If we draw again and get a black ball, this means:

$$P(H'_b|E'_b) = .757$$

$$P(H'_w|E'_b) = .009$$

$$P(H'_n|E'_b) = .233$$

If a white ball were to be drawn:

$$P(H'_b|E'_w) = .210$$

$$P(H'_w|E'_w) = .210$$

$$P(H'_n|E'_w) = .580$$

Thus, if the second sample is a black ball, we would choose $b$ since it has the maximum expected utility at .757, and if we get a white ball, we choose $n$ with the expected value at .580. So, the expected utility, if we were to draw from the urn again, is: $.7132(.757) + .2867(.580) = .706$, which is an improvement over just drawing once. The net gain is $.706 - .600 = .106$. Ramsey's proof shows that we can keep on getting more evidence and we will never be worse off.

To see a bird's-eye view of the situation, consider figure 3.10, in which the sequential revisions of expected value after drawing from different hypothesized distributions are plotted. Each line represent the *net change* of expected value of getting more evidence conditional on having $x$ observation. We can see that the expected net change will never dip below 0, meaning that one will often be rewarded and never punished for getting more evidence.

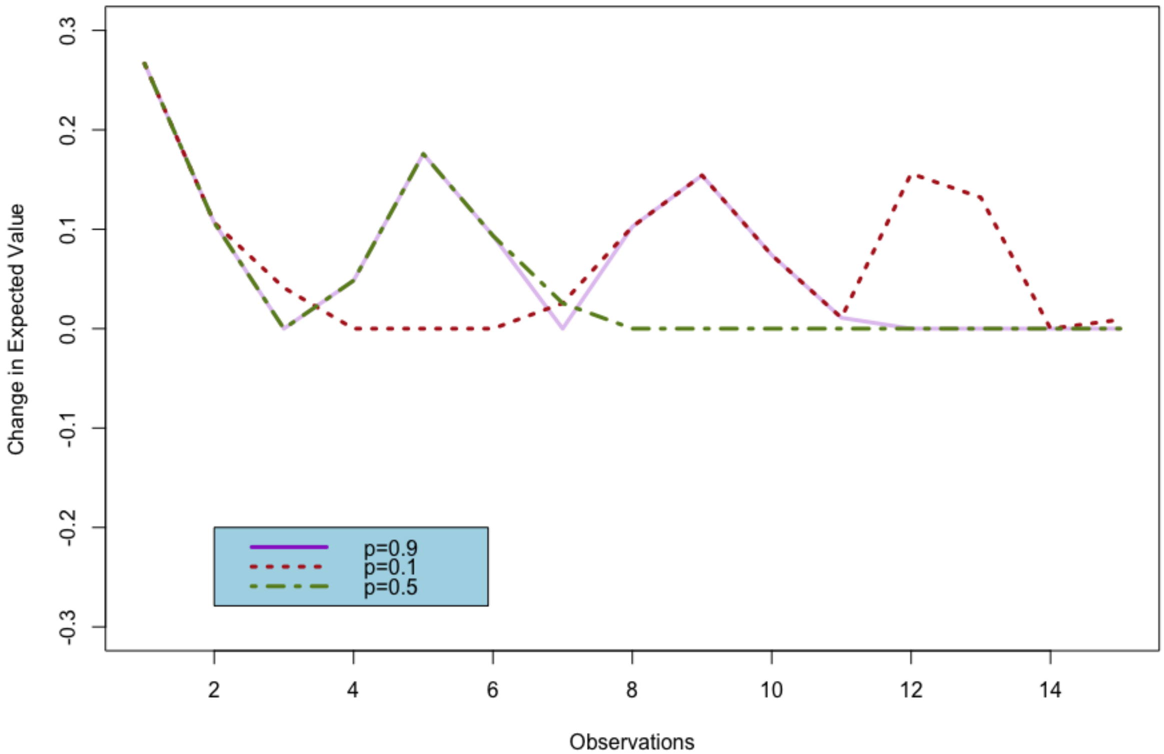


**Figure 3.10**: Net gain in expectation after $x$ observations from $\theta = .9, \theta = .1, \theta = .5$.

This result could be used to answer the Bayesian version of optional stopping in this way: since getting more evidence always yields better expected values, the ideal Bayesian agent will always opt for more evidence, instead of stopping ahead just because the posterior has reached her favorite degree.

However, I do not think this answer will do. To begin, the crucial assumption here is that evidence costs *nothing*. The scenario we imagined quickly breaks down once we starts to introduce some sort of cost. It was assumed in the example that it costs us neither money nor time to draw from the urn, but suppose it costs us 25 cents for each sample. This means that we would be gaining only $.27 - .25 = .02$ in expected payoff for the first draw, and the second draw would definitely not be worth the additional 25 cents. Or suppose that one dollar is not worthy any endeavor that lasts longer than 15 seconds, and it takes 30 seconds to draw from the urn. As soon as minimally realistic assumptions are introduced, Ramsey's result no longer holds.

Cost might also enter into consideration in different forms. Savage discusses a case in which a very ill person, who is given a cost-free option to find out if the disease she has is terminal. Savage points out that an argument can be made that in this case refusing information could be rational. The thought is that the patient may decide that, based on an assessment of her own personality, she would live the rest of her remaining life in agony if she were to find out that her disease is very serious, whereas she could live relatively happily without knowing. Savage's point is that in this case the information is not really free; it has a *psychological* cost.[39]

Ramsey's and Good's proofs, while extremely valuable from a logical and mathematical perspective, are somewhat tone-deaf to the practical problem posed by Ayer and Keynes. The actual complaint was that the Principle of Total Evidence *presupposes* that we know ahead what "total evidence" amounts to, since the decision to get more evidence or simply stick to our current body of evidence is not one that can be resolved just by appealing to probability, because the rationality of such a decision is highly context-sensitive. One important context is the *urgency* of decision. In many cases time is of the essence: "a general who refused to launch an attack until he had

39. Savage, *The Foundations of Statistics*, 107.

ascertained the position of every enemy solider would not be very successful."[40]

The economist G. L. S. Shackle makes a similar point engagingly by retelling the thought process of a certain Chinese guard who had to decide on the spot whether to join the rest of the guards to partake in a rebellion or to be the lone loyalist to stand in defense of the empire. He argues that it would be rather foolish to suggest that the guard should maximize his utility by looking for more evidence:

> [Had the guard taken heed of the advice given by the expected utility theorist,] he might have argued thus: 'I find in the record of history a thousand cases similar to my own, wherein the person concerned decided upon treachery, and in only four hundred of these cases the rebellion failed to and he was beheaded. On balance, therefore, the advantage seems to lie with treachery, provided one does it often enough'... Had the sentry decided to support the rebellion, he might have had time, just before the axe fell, to reflect that he would never, in fact, be able to repeat his experiment a thousand times, and thus the guidance given him by actuarial considerations had proven illusory.[41]

My point, of course, is not that making decisions based on probability and utility is irrational. Far from it, but that rational inductive thinking presupposes a deliberative framework. The context of the story makes it clear that for the guard, "total evidence" really just means whatever he has in mind at the moment, and it would be irrational to suggest that he should get more evidence just because his expected utility will improve. As Ayer suggests, Carnap's inductive logic presupposes that the agent already know what total evidence is, and the extent to which new evidence is warranted.

Good, who proved the same result independently of Ramsey, tries to address this

40. Ayer, *Probability and Evidence*, 57.

41. G.L.S. Shackles, *Uncertainty in Economics and Other Reflections* (Cambridge University Press, 1955), 2.

issue by distinguishing what he calls Type I and Type II rationality.[42] Type I rationality is that of the ideal Bayesian agent, one who lives her life by abiding by the principle of maximizing expected utility. Good recognizes, however, that Type I rationality provides no guidance in regard to when an investigation should be concluded. This is where the more emcompassing Type II rationality comes in: it consists of the principle of maximizing expected utility plus the consideration of "the cost of theorizing." More important, the goal of type II is "a sufficient maturity of judgments."[43] Good's two types of rationality could be interpreted as a concession to the idea that there is a level of rational criticism that cannot be captured within the strict framework of expected utility. Phrases such as "cost of theorizing" and "maturity of judgment", it seems to me, are evaluations of the rationality of the agent's the intention to stop, so intentions are, after all, relevant in Bayesian reasoning.

## 3.6 Conclusion

My focus of this section was to elucidate the very idea that the commitment made in the abductive context has repercussions on the investigator's deliberate conduct during the inductive stage of inquiry. The problem of optional stopping in parapsychology has served as a helpful case to demonstrate the issues at stake. Even though this problem has been traditionally associated with frequentist methodologies, I have tried to show that it should also concern Bayesians by reproducing a similar result using statistical simulations.

42. Good, *Good Thinking: The Foundations of Probability and its Applications*, 29-30. As far as I could tell, this has nothing to do with the distinction between Type I and Type II error in frequentist statistics.

43. ibid., 29.

My contention is that aspects of inductive reasoning have to be criticized in light of the deliberation the inquirer undergoes prior the experiment, such as the commitments and intentions considered and decided by the experimenter, all of which directly impact the statistical model used for inductive inference. Error probabilities are especially important in the evaluation of the experimenter's intention to allow her hypothesis to be confronted by experience in a fair manner.

I explored a Bayesian response to optional stopping, which relies on a mathematical result, proven by Ramsey and others, that one's expected value never decreases from gathering new evidence, and in many cases it actually increases. I dismissed this line of thought, because the result only holds when the evidence is cost-free, which is almost never the case.

# Chapter 4

# Abduction, Strategic Interrogation, and The Weight of Evidence

The main purpose of this chapter is to propose a way in which the interplay between abduction and deduction can be understood, in service of understanding how the deliberative elements of an experiment could be criticized. The process of abduction, I suggest, is intertwined with deduction: deduction is used to strategically interrogate the hypothesis chosen provisionally, in order to see a better hypothesis should be chosen before committing into inductively testing the hypothesis. During deliberation, we switch back and forth between making educated guesses abductively and derive the necessary connections from it deductively.

To see how this is done in the context of probabilistic judgment, I shall examine Keynes' notion *the weight of evidence*. My claim is that the weight of evidence is a *deductive* tool a reasoner could use to evaluate the worthiness of a provisionally chosen hypothesis. In particular, I suggest that the intentions and goals of an experimenter can be criticized by appealing to notions of *counterfactual priors and hypothetical data*. My suggestions here are heavily influenced by the works of statisticians who have argued to move away from the orthodox approaches to Bayesian analysis, such as James Berger's "Robust Bayesianism",[1] Gelman and Shalizi's "hypothetico-deductive

1. James Berger, "The robust Bayesian viewpoint," in *Robustness of Bayesian Analyses*, ed. J. D. Kadane (Elsevier Science, 1984)

Bayesianism",[2] and Peter Walley's theory of imprecise probability.[3]

I will set the stage with a discussion on Peirce's idea that abduction is an interrogative process in section 4.1. An exposition on Keynes' ideas on the weight of evidence will be carried out in section 4.2. To motivate my proposal, I will introduce Popper's *paradox of ideal evidence* in 4.3. I will overview a statistical solution to the paradox in section 4.4, and a formal epistemological perspective in 4.5. Finally, I will articulate my proposal in 4.6.

## 4.1 Deduction as Strategic Interrogation

In section **??**, we discussed Peirce's idea that the abduced hypothesis is a highly fallible conjecture, insofar so he would sometimes refer to it as a guess. However, Peirce also makes it clear that, even though abduction is ultimately guessing, he does not mean that abduction is nothing but a game of blind luck. One characterization that Peirce often use in the context of abduction is that it involves an "interrogation":

> A hypothesis ought, at first, to be entertained interrogatively. [4]
>
> The first starting of a hypothesis and the entertaining of it, whether as a simple interrogation or with any degree of confidence, is an inferential step which I propose to call abduction. [5]
>
> It is to be remarked that, in pure abduction, it can never be justifiable to accept the hypothesis otherwise than as an interrogation. But as long as that condition is observed, no positive falsity is to be feared; and therefore the whole question

2. Andrew Gelman and Cosma Rohilla Shalizi, "Philosophy and the Practice of Bayesian Statistics in the Social Sciences," in *The Oxford Handbook of Philosophy of Social Science*, ed. Harold Kincaid (Oxford University Press, 2012)

3. Peter Walley, *Statistical Reasoning with Imprecise Probabilities* (Chapman / Hall, 1991)

4. Peirce, *Collected Papers of Charles Sanders Peirce*, 6.524

5. ibid., 6.525

> of what one out of a number of possible hypotheses ought to be entertained becomes purely a question of economy. [6]

A helpful way to explain what Peirce means by "interrogation" is to use own analogy of the game of "twenty questions", in which one party has to think of an object, while another party has to find out what the object by asking 20 or less questions.[7] Peirce's idea is that posing a question in the game is akin to proposing a hypothesis in an inquiry—for both cases, we make a guess about the nature of the subject of our inquiry, and the difference is that in the game the answer we receive is certain and direct, while a hypothesis has to be tested or broken down into sub-hypothesis. Peirce has an important point in mind, however, with this example. That is, the choice of a question can drastically alter the course of an inquiry. Since the question of the game has to figure out what the answer is by asking at most 20 questions, she must ask them *strategically.* Each question the inquire asks should be as informative as possible—they should, for instance, ask questions that narrow the space of possibilities as quickly as possible. Abduction, which is the logic of hypothesis selection, is guided by a similar goal.

Hintikka, who calls abduction "the fundamental problem in contemporary epistemology", proposes that the interrogative element of abduction could be understood as the strategic process in which the informational gain, in light of the agent's epistemic goal, from inquiry is maximized.[8] Hintikka raises an interesting example of how the strategic freedom in abductive thinking is crucial for deductive logic, notwithstanding

6. Peirce, *Collected Papers of Charles Sanders Peirce*, 6.529

7. Peirce, *The Essential Peirce, Volume 2: Selected Philosophical Writings (1893-1913)*, 109

8. Jaakko Hintikka, "What Is Abduction? The Fundamental Problem of Contemporary Epistemology," *Transactions of the Charles S. Peirce Society* 34, no. 3 (1998): 503–533

its emphasis on mechanical rules of inference. Suppose a student is asked, in an exam for a logic class, to demonstrate that

$$\vdash [[A \rightarrow (B \wedge \neg C)] \wedge (\neg B \vee D)] \rightarrow (A \rightarrow D) \tag{4.1}$$

In other words, the student has to prove that formula 4.1 is a logical truth. Before proving that, she must make a decision on which method to use. Since she is pressed for time, it is in her interest to prove in a strategic manner. For instance, suppose she choose the *semantic* strategy of showing that the proposition is always true. The most obvious way would be to use the truth table to show that the formula is true under all possible interpretations. However, that would be a inefficient choice, since there will be $2^4 = 16$ rows and, because of her strategic decision, she cannot skip any row—she has to demonstrate that the proposition is true in all rows.

Thus, the strategic dimension of reasoning involves finding a way to gain as much information as possible in light of economical constraints. In the game of twenty questions, the constraint is that the questioner only gets 20 questions. For the student taking a logic exam, she is bound by the time she has. In both cases, the reasoner has to make a decision on confronting the problem she faces in an efficient manner. In particular, before even tackling the problem they are interested, both reasoners must select one of many methods in which the problem could be addressed. Reasoning about how our resources should be allocated in service to our epistemological goal is often categorized by Peirce as the logic of the economy of research:

> Proposals for hypotheses inundate us in an overwhelming flood, while the process of verification to which each one must be subjected before it can count as at all an item, even of likely knowledge, is so very costly in time, energy, and money—and consequently in ideas which might have been had for that

> time, energy, and money, that Economy would override every other consideration even if there were any other serious considerations. In fact there are no others.[9]

Hinttika's point the connection between abduction and deduction is relevant here. The economical implication of a hypothesis can often be probed in a deduction, which is both relatively cost-free and secure.

For example, reconsider formula 4.1. Suppose the logic student realizes that using truth table would be too uneconomical for this particular problem, and decides to use natural deduction instead. This, however, calls more further strategic thinking, because since there is no premise given, the student must *choose* her own assumed premise, and often there are many different ways in which a proof can begins . Introductory textbooks often recommend the strategy of assuming the negated version of the logical truth, and prove the result by using *Reductio ad Absurdum*, if the student is unsure how to begin.

In this case, however, this might not be the best decision, for the result seems rather unwieldy:

$$\neg\{[[A \rightarrow (B \wedge \neg C)] \wedge (\neg B \vee D)] \rightarrow (A \rightarrow D)\} \vdash ???? \tag{4.2}$$

She would end up with a long negated conditional, without a clear way forward. Of course, it is not impossible to find a contradiction from formula 4.2. If the student keeps out unpacking it deductively, she will eventually run into a contradiction by brute force, just as in the game of twenty questions, a player might eventually run into the correct answer if she had unlimited chances to ask question.

9. Peirce, *Collected Papers of Charles Sanders Peirce*, 5.602

The point, however, is that there are more economical choices to begin the inquiry. One strategy would be to take advantage of the conditional nature of the logical truth. That is, we can prove the result conditionally assuming the antecedent of the conditional, and derive the consequent:

$$[[A \rightarrow (B \wedge \neg C)] \wedge (\neg B \vee D)] \vdash (A \rightarrow D) \tag{4.3}$$

The right hand side of the $\vdash$ presents itself a clear way forward: assume $A$, and then derive $D$. A student with a keen eye for deduction can see that the assumption given in 4.3 is a conjunction that can be decomposed into two conjuncts, giving us the propositions needed to derive $B \wedge \neg C$ using $A$, and then get $D$, using $B \wedge \neg C$. Depending on the rules available to her, such a method would take 10 or less lines.

The example is supposed to demonstrate how the question of economy of a proposed premise can be investigated deductively. Similar considerations apply in the game of twenty questions. Suppose we are considering for the first question the choice between asking "is it steak?" and "is it food?" It is a matter of deduction that the latter is much more informative: both an affirmative or negative answer would greatly reduce the size of the possible question. It would, of course, be nice if the answer happens to be steak, in which case the former would be the question to ask, but the probability, given we know nothing about the object in question, is astronomically low, and knowing that it is *not* steak only rule out *one* possible answer.

## 4.2 The Weight of Evidence

In *A Treatise in Probability*, Keynes discusses a great deal about how probability ought to reflect our epistemic judgments. One type of such judgments is the *judgment of relevance*. Keynes' observation is that we often can judge whether one proposition $E$ counts as being relevant to another proposition $H$ by considering whether the probability of $H$ would change on the supposition that $E$ is true. Keynes's example is that, in a typical urn example with some black and white balls, if we want to know the probability of a white ball being randomly chosen, the color of the ball would not change its probability of being chosen, so the idea is that a ball's probability of being chosen conditional on being (say) white is the same as the probability of the ball being chosen in general.[10] So, Keynes proposes that evidence $E$ is irrelevant to the proposition $H$ if and only if:

$$P(H|E) \neq P(H)$$

Now that the notion of relevance has been introduced, we come to Keynes' idea of the weight of evidence. Keynes is troubled by the fact that the degree of a probability does not scale straightforwardly with the amount of the evidence we have at hand. In a well-known passage, Keynes says:

> As the relevant evidence at our disposal increases, the magnitude of the probability of the argument may either decrease or increase, according as the new knowledge strengthens the unfavourable or the favourable evidence; but something seems to have increased in either case,—we have a more substantial basis upon which to rest our conclusion. I express this by saying that an accession of new evidence increases the weight of an argument. New evidence will sometimes decrease the probability of an argument, but it will always increase its

10. Keynes, *A Treatise on Probability*, 59.

'weight.'

The crucial idea here is the weight of evidence is closely tied to the absolute amount of evidence and is conceptual distinct from the "magnitude" of a probability. Keynes explains this as the distinction between the *balance* and the *weight* of the evidence: he first brings our attention to the fact that when we consider the conditional probability of the hypothesis in question under all relevant evidence, the resultant number constitutes the balance between favorable and unfavorable evidence.[11] For instance, we may say that when $P(H) < 0.5 < P(H|E)$, then evidence $E$ is somewhat in favor of the hypothesis. Of course, the balance changes as we gather more relevant evidence, and it might go from favorable from unfavorable depending on the nature of the new evidence.

However, as Keynes points out, this is not the only epistemologically significant relation between probability and evidence, for we not only care about how much the current evidence favors the hypothesis, but we also concern ourselves with the *amount* of evidence involved in calculating the balance of the evidence, and Keynes calls this measure the *weight* of evidence. But, unlike the balance of the evidence, which can go either direction, the weight of evidence can only go up as we gather more relevant evidence. In Keynes' words, "New evidence will sometimes decrease the probability of an argument, but it will always increase its 'weight.'"[12]

To see what Keynes means, imagine two urns $A$ and $B$ with unknown proportions of black and white balls. Suppose you sample (with replacement) 100 balls from the urn $A$ and find 50 black balls and 50 white balls. Justifiably, you infer that the

11. Keynes, 78.

12. Keynes, 78.

proportion of black balls in $A$ - call it $\theta_A$ is about 0.5. You then decide to sample from $B$, but this time you only manage to draw 4 samples, 3 of which are black balls. Your best estimate for $\theta_B$ is 0.75. At this point, I offer you another chance to draw from one of the urns, and if you manage to draw a black ball from that urn, you get \$100. Which urn would you pick?

Clearly, $\theta_B > \theta_A$, but it is not clear that $B$ is obviously the better choice, because the amount of evidence you have for $\theta_A = 0.5$ is higher than for $\theta_B = 0.75$. This is a problem for probabilism, because, in terms of just comparing the probabilities alone, picking urn $B$ clearly has a higher probability of winning; however, all the facts in the situation are different than what the probability lets on, so the probability has failed to reflect some crucial information about the evidence.

Keynes was not the first to notice the problem of the weight of evidence, it is one of many criticisms Peirce has for what he calls *conceptualism*, a subjective position on probability that was heavily influenced by Laplace. In particular, conceptualists accept of the so-called Principle of Indifference, which says roughly that complete ignorance should be modeled as a uniform distribution over all hypotheses. In a typical case of estimating a unfamiliar coin's probability of heads, this would mean that the expected value is 0.5.

We already reviewed some of Peirce's problems with the Principle of Indifference in section **??**. It is in this context that Peirce appeals to the notion of the weight of evidence. Peirce's argument is pragmatic: conceptualists say that you should adhere to the Principle of Indifference when you either have no information. This means that the degree of belief you *should* have for a unfamiliar coin landing on heads on the next toss is 0.5. However, Peirce argues there is a behavioral difference between betting on a fair coin and a unfamiliar coin. For the former you should know exactly how much to bet, in the latter case you should simply refrain from betting. The Conceptualist model, however, cannot make sense of this, since both entail the degree of belief of 0.5. Nowadays this is usually known as the *Ellsberg's Paradox*.[13]

In other words, conceptualism fails to distinguish between *rational indecision* and *indifference*. What distinguishes the two is the difference in the weight of evidence. To recycle the example earlier, further consider another urn $C$, from which you draw 2 balls, and one of them is white, so your best estimate would be $\theta_c = 0.5$. If probabilities can perfectly reflect the evidence, then it must mean that your epistemic attitude toward $A$ and $B$ ought to be the same, but Peirce insists that this cannot be the case.

> In short, to express the proper state of our belief, not one number but two are requisite, the first depending on the inferred probability, the second on the amount of knowledge on which that probability is based.

The weight of evidence, then, is a crucial piece of the puzzle for the position I am trying to defend. The goal of Deliberative Bayesianism aims to situate Bayesianism within the Peirce's framework of abduction, deduction, and induction. My claim is that the weight of the evidence something is relevant in *abduction*. This is why Keynes was perplexed about it.

## 4.3 The Paradox of Ideal Evidence

Like Peirce before him, Karl Popper was highly critical of the subjective interpretation of probability and the epistemologies that sprung out of it. Popper has further develop

13. Ellsberg, "Risk, Ambiguity, and the Savage Axioms."

Peirce's criticism as the *paradox of ideal evidence.* The alleged paradox arises out of the contradiction that, by accepting the notion of conditional relevance proposed by Keynes, some evidence is both relevant and irrelevant.

Popper asks us to consider a certain coin with an unknown bias: let $N$ be the proposition "the next toss of the penny will yield heads."[14] Now, what should $P(N)$ be? He suggests, either by appealing to intuition or the Principle of Indifference, Bayesians would suggest that $P(N) = 0.5$.[15]

Now let $I$ be what he calls *the ideal statistical evidence* in favor of the idea that the coin in question is a fair one. Popper's example is to let $I$ be a statistical report that says 'in a million tosses, the coin landed on heads roughly half a million times.' The exact number is not important, as long as the number of heads and tails would make it practically certain that the coin is fair—the same point could be made using 10 millions instead of a million. Now, given we have ideal evidence $I$, what is the probability of $N$? Popper claims that it would have to be 1/2. So

$$P(N|I) = P(N) = \frac{1}{2}$$

However, as discussed earlier, evidence $I$ is relevant to the hypothesis $N$ if and only if

$$P(N|I) \neq P(N)$$

14. Karl Popper, *The Logic of Scientific Discovery* (Routledge, 2002), 425.

15. It should be noted that Popper is not attacking the principle of indifference in this context. That is, for this argument he is willing to grant that Bayesians have some way of arriving at $P(N)$—it could be by indifference, through elicitation, etc.

If $P(N|I) = P(N) = 1/2$, this means that the ideal evidence is also irrelevant evidence. Popper then concludes:

> Now this is a little startling; for it means, more explicitly, that our so-called 'degree of rational belief' in the hypothesis, [$N$], ought to be completely unaffected by the accumulated evidential knowledge, [$I$]; that the absence of any statistical evidence concerning [the hypothesis that the coin is fair] justifies precisely the same 'degree of rational belief' as the weighty evidence of millions of observations which, prima facie, support or confirm or strengthen our belief.[16]

What is 'startling' about this? Popper's point appears to be that we *expect* the awareness of evidence $I$ should change our attitude toward $N$ *in some way*, but if our prior for $N$ is already $1/2$, $I$ will not change it in anyway, so on Keynes' account, $I$ is irrelevant. This seems to contradict with our intuition that the ideal evidence should be relevant.

We can interpret Popper to making this following argument:

1. $I$ is ideally favorable to $N$.
2. $P(N|I) = P(N) = 1/2$.
3. According to the notion of conditional relevance, $I$ is irrelevant to $N$.
4. But according to premise 1, $I$ is relevant to $N$.
5. 3 and 4 are contradictory.

The inferential step from premise 1 to premise 4 is not via the technical notion of conditional relevance, since it appeals to the intuitive idea of what "ideally favorable" is. As the argument stands, there is nothing stopping the Bayesian from biting the bullet and insisting that $I$ is irrelevant to $N$, or that $I$ was never ideally favorable to begin with. Of course, this response is not satisfactory unless Bayesians have a way

16. Popper, 426.

to say something about what exactly $I$ is doing to our state of belief. This returns to Keynes' initial observation: clearly *something* is changed by the ideal evidence, but it is not $P(N)$.

One answer we want is clearly that it's the weight of evidence that changed, and it is manifested as a property of $P(N) = 1/2$. But what does that mean?

## 4.4 The Statistical Perspective

The natural reaction to Popper's paradox is to deal with it statistically. In fact, there is a clear statistical answer from the Bayesian perspective that could address the paradox. I will try to explain this briefly but suggest why the answer, while making perfect statistical sense, is not sufficient as an epistemological answer.

In section 3.4, we discussed how Bayesian methods could be used to represent our degrees of belief about a hypothesis regarding physical chance. The same can be done here to address Popper's paradox easily.

The kind of trials involved in the paradox of ideal evidence can be modeled as Beta-Bernoulli process, where the Beta distribution would model our state of belief and the Bernoulli distribution the coin tossing process. This can be seen as a special case of the binomial distribution we used earlier. Here, the Bernoulli distribution has the parameter $\theta$, which is often interpreted as the probability of success of a binary event, e.g., landing on heads, and thus in this sense we are talking probabilities of a probability. Again, we canthink of the parameter $\theta$ as representing the *propensity* of the coin. We then use the Beta distribution to model the propensity, representing the degree of our belief in various hypotheses of $\theta$ having a certain value $x$ where

$0 \leq x \leq 1$.

More precisely, let $\theta$ be the propensity of the coin to land on heads and let

$$X_i = \begin{cases} 1 & \text{the coin lands on heads on toss } i, \\ 0 & \text{otherwise.} \end{cases}$$

Now these random variables can be modeled as follows:

$$\theta \sim Beta(\alpha, \beta)$$

$$X_1, ...X_i \sim Bern(\theta)$$

As mentioned, the Beta distribution has two parameters, $\alpha > 0$ and $\beta > 0$, which can be thought of as, in our context, our initial opinion about the coin's propensity. What we did not discuss previously, however, is that the beta distribution can represent different states of belief that have the same expected value. Consider three beta distributions:

1. $Beta(1, 1)$:
2. $Beta(11, 11)$
3. $Beta(500001, 500001)$

Note that all three distributions have the same expected values:

$$\frac{1+0}{2+0} = \frac{1+10}{2+20} = \frac{1+500000}{2+1000000} = \frac{1}{2}$$

However, even though these distributions produce identical expected values, if we

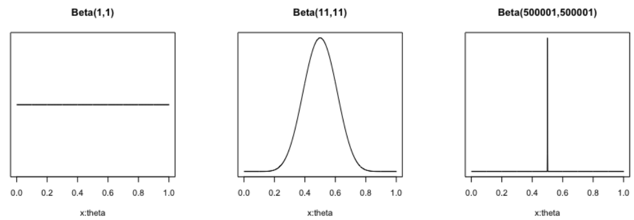


**Figure 4.1**: Beta Distributions

plot them, we can see that how they represent states of belief that are drastically different:

Intuitively, we can think of the first distribution as representing your state of belief about the probability of getting a head on the next flip. This distribution is plotted in Figure 1: note that it is wholly flat, capturing the sort of state of indifference that the Principle of Indifference is supposed to capture: One finds no ground in thinking one probability is more credible than another.

The second distribution can be seen as our state of belief after witnessing 10 flips of the coin, and 5 turn up heads and 5 tails. Naturally, the peak - the mode of the distribution - is at $\theta = 0.5$, which seems sensible, because it reflects the evidence that exactly half of the samples are heads. But we can see that at this stage we are quite uncertain about $\theta$, evidenced by the width of the distribution. While $\theta = 0.5$ is the peak, there is a substantial area covering $\theta > 0.55$ and $\theta < 0.45$.

The third distribution, modeling the state of belief after one million trials with half of them being heads, is intended to be an approximation of Popper's ideal evidence scenario. The peak is again at 0.5, but this plot has a noticeably narrower spread: we are much more confident in our assessment that the coin has an equal propensity to land on heads as tails. Also notice that at this stage, any value of $\theta$ other than 0.5

are practically impossible after receiving the ideal evidence.

To state these observations more precisely, we can calculate the exact probability using the corresponding cumulative distributions. Since beta distributions are continuous distributions, we can only deal with intervals of values. Still, we can provide a reasonably close approximations. For instance, conditional on the ideal evidence, we would be absolutely sure that the probability is between 0.49 and 0.51. The relevant probabilities are summarized in the following table:

| Distribution | $P(0.46 < \theta < 0.54)$ | $P(0.49 < \theta < 0.51)$ |
|---|---:|---:|
| $Beta(1, 1)$ | 0.08 | 0.02 |
| $Beta(11, 11)$ | 0.29 | 0.07 |
| $Beta(500001, 500001)$ | 1 | 1 |

Now, let $E$ be the ideal evidence, $X_1, \ldots X_{1000000}$, where $\sum_{i=1}^{1000000} X_i = 500000$, and let $\theta$ be the coin's propensity to land on heads. We now see that the following inequality holds, since the left-hand side is 0.02, and for the right it's 1.

$$P(0.49 < \theta < 0.51) < P(0.49 < \theta < 0.51|E)$$

To make things more official-sounding, perhaps we can describe this as *Higher Order Relevance(HOR)*. Recall Keynes's idea is that for some evidence $E$ and hypothesis $H$

$$\text{Evidence } E \text{ is relevant to } H \text{ if and only if } P(H) \neq P(H|E)$$

HOR takes this on step further and suggests that, in addition to $H$ and $E$, consider specific values $x$ and $y$, where $0 \leq x \leq y \leq 1$

Evidence $E$ has a higher order relevance to $\theta$ iff

$$P(x \leq \theta \leq y) \neq P(x \leq \theta \leq y)|E)$$

Under this analysis, we can see that Popper's argument contains a sleight of hand that shifts between two ways of thinking about N—the next coin toss landing on heads—'s probability. The argument begins by asking, rather innocuously, for your prior for $N$, but the ideal evidence $I$ Popper immediately introduced is not for $N$ but for the hypothesis that $\theta = 0.5$. Popper is reasonably explicit about *that*, but what he is not explicit about is *this*: he has convinced us that $I$ is both evidentially ideal for and conditionally relevant to $\theta = 0.5$. That, however, is a misdirection, because he immediately starts talking the conditional probability on $I$, *not* of $H_{0.5}$, but of $N$.

While I think that HOR provides a technical response to Popper's paradox, it is not quite the same as accounting for the phenomenon in question. In fact, by focusing on overcoming the difficulty raised by the paradox caused by an absurd amount of evidence, we might have overlooked what is truly at stake: rarely, if ever, do we have ideal evidence for any substantive hypothesis, so situations where we have an overabundance of evidence is an incomplete benchmark for the adequacy of the account. In fact, our analysis shows that when we have perfect information, evidential weight essentially becomes a non-issue, because it eliminates the uncertainty that calls for probabilistic reasoning to begin with.

The important question, instead, is whether HOR can help with decision making in situations where evidence is severely lacking. To this end, it remains to be seen how higher order relevance can trickle down to first order probability, on which decision making is based within the classical Bayesian framework.

We need to, then, ask ourselves if HOR as a concept can make any practical difference in decision making. It is in fact difficult to do so within the basic framework of Bayesianism. To see this, imagine a perverse game in which you will be shot to death if a coin flip lands on head. Clearly, you don't want heads. You are given a choice between two coins: the first coin $P$ is similar to Popper's coin from the ideal evidence scenario, except now the ideal evidence actually shows that there is a slight bias in favor of heads, say the expected value is 0.51. The other coin $U$ is one you have never seen before, so on an ignorance prior your expected value $E(\theta_U)$ is 0.5 . Now, which would you choose? An argument can be made that you probably still want the Popperian coin, because you know you are almost certainly getting $\theta_P = 0.51$. From the perspective of expected utility, however, it is hard to rationalize such a decision, because the unknown coin still has a lower expected value. That is,

$$\frac{510000}{1000000} > \frac{1}{2}$$

So it seems that we are back to where we started - the relevance demonstrated on a higher order simply vanishes when we consider the matter on the level of decision making, which is entirely based on a precise point-estimate of the first order probability.

If the point estimate is to be blamed, the natural response is that we do rely on an interval estimate instead. This solution is reminiscent of the call to abolish the use of $p$ values in Frequentist statistics, and instead we should report the confidence interval of our findings. The idea is that point estimates are inherently misleading, since they, by design, summarize the data by discarding information such as higher order relevance. This problem is somewhat analogous to the one we are running into with respect to expected values. So one possible solution is that we should only insist on making our decisions based on *credible intervals*, which is the Bayesian version of the confidence interval. For instance, suppose $\theta_P \sim Beta(480000, 520000)$ and $\theta_U \sim Beta(1, 1)$. We can deduce that

$$P(0.479 \leq \theta_P \leq 0.481) = 0.99$$

$$P(0.005 \leq \theta_U \leq 0.995) = 0.99$$

In other words, we can say there is a 0.99 probability that $P$'s propensity to land on heads is between 0.479 and 0.481 (practically 0.48) and for $U$ it's between 0.005 and 0.995.

However, it seems to me that we are simply restating higher order relevance in terms of credible intervals, without dealing with the crux of the problem - unless we are rationally allowed to refuse to follow the precise expected value, even if the weight of evidence is low, we will always have to match it to our best point estimate, which is the expected value. Of course, the point is not that HOR doesn't matter, because intuitively it does. The point is rather that we need a richer philosophical framework to rationalize this intuition.

## 4.5 The Concept of Resiliency

Skyrms credits Richard Jeffrey as the first who notices that Popper's paradox brings light to the very idea of resiliency.[17] Jeffrey points out that once we stop fixating on the probability of $N$, the next toss coming up head, we can see that our state of belief prior receiving the ideal evidence has a degree of malleability.[18] Consider, for instance, instead of asking only for the probability of $N$, we ask the probability of the next 5 tosses coming up heads. Once we think about how our belief responds to how these 5 tosses would act as potential evidence, given our posterior state of belief, we have very little choice but to believe that probability is $(0.5)^5$, but we would be a lot less compelled to do so with the prior state of belief.

Skyrms has proposed the notion of *resiliency* to capture Jeffrey's observation in a generalized manner: even though evidential weight is not reflected by the probability, it is captured by its stability.[19] The idea that there is a probabilistic representation for a stable state of belief can be illustrated as follows: if I have in front of me an urn $U$, with an unknown proportion of black and white balls. If I randomly draw 2 balls from it with replacement and find one ball for each color, my intuitive estimate of the proportion of black balls would sensibly be somewhere around 1/2. But my state of belief should be relatively unstable: it would be irrational for me to fixate on this estimate, especially new light of conflicting evidence. If I sample two more balls from the urn and they are both black, it would make sense for me to raise my estimate for the proportion of black balls to more than 1/2—perhaps to 3/4. But suppose I

17. Brain Skyrms, "Resiliency, Propensities, and Causal Necessity," *Journal of Philosophy* 74, no. 11 (1966): 704–713,704

18. Richard Jeffrey, *Logic of Decision* (McGraw-Hill, 1965), 184.

19. Skyrms, "Resiliency, Propensities, and Causal Necessity," 705

continue to sample from for 996 more times. Out of the total 1000 draws, 500 are black. At this point a sensible would be back to around 1/2, but unlike my state of belief after only 2 samples, after 1000 samples my state of belief is stabilized: suppose I sample again and I draw five black balls in a row. Now, even though drawing 5 black balls in a row seems rather extraordinary, against the body of my evidence it would not warrant me to revise my belief in any significant measure.

The intuition here is that the increase in the amount of evidence, expressed here in terms of the number of samples, corresponds to the increase of stability of the estimate. Skyrms has introduced a notion called the *resiliency* to capture this intuition sense of stability.

Conceptually, this is an attractive way to capture to notion of evidential weight. Keynes' puzzlement about how relevance and weight could come apart is addressed; When a belief $B$ is resilient, the conditional probability on some new evidence $E$ should be approximately the same, that is, $P(B|E) \approx P(B)$, even if $E$ would be highly relevant were $B$ not resilient. If, for instance, a resilient belief is one where the weight of the evidence for it is high, then it is a logical consequence that evidence could make a belief more weighty without changing its degree, for the weight is in fact stabilizing this particular value.

Nevertheless, how the resilience of a belief can make a practical difference is yet to be explained. In fact, for the most part, resiliency does not do much better than higher order relevance: within the structure of expected utility calculus, a highly resilient belief will still recommend the same action as a non-resilient belief with the same expected value. Skyrms' original motivation for the concept, however, provides an important clue: he intended the concept of resiliency to explicate the concept of the laws of nature, so the idea is that a probability statement $A$ based on a law of nature is one that remains resilient against various extreme scenarios. What this means is that we are supposed to calculate likelihoods based on data that *might have happened*, and see whether $A$ is resilient against it.[20]

## 4.6 Counterfactual Priors and Hypothetical Data

The notion of resiliency opens the door to how counterfactual reasoning can be relevant in Bayesian reasoning. The crucial point is that to see how resilient a probability is, it is not something we can determine by looking at data that we already have. We have examine what *might have happened.* So notwithstanding Lindley's rhetorical question:

> Of what relevance are things that might have happened, but did not?[21]

Counterfactual considerations are relevant in deliberating on how a chosen hypothesis *would* respond in light of confounding scenarios, so that we can decide if it is a worthy hypothesis and, if chosen, how we ought to further probe it. More specifically, the weight of evidence is puzzling, because it needs to be understood in the context of abduction, since its goal is to give information about how we ought to structure our inductive space.

20. James Joyce has recently proposed that what evidential weight does is to stabilize the distance between one's subjective probability of an event and her underlying hypotheses of the objective chance about it. Joyce's account has some important technical advantage, but conceptually it remains the same as Skyrms' account. Essentially, Joyce's suggestion is that the basic resilient quantity is the *expected loss* of a particular chance hypothesis. Providing a proper treatment would require an exposition on statistical decision theory, so I have decided to omit it here. See James M. Joyce, "How Probabilities Reflect Evidence," *Philosophical Perspectives* 19, no. 1 (2005): 153–178, 166-168.

21. Lindley and Phillips, "Inference for a Bernoulli Process (A Bayesian View)," 114.

Skyrms suggests that the resilience of a belief manifests itself as "a reluctance to change."[22] He suggests that we can measure weight directly in terms of the difference between prior and posterior probabilities. Perhaps we can call this the measure of instability, which signifies a lack of weight:

$$\text{Instability: } |P(X|E_j) - y|$$

Where $j$ is in the set of $n$ possible states of affairs, $E_1, ..., E_n$. Skyrms' idea is that we should pick a $j$ that creates the biggest difference.

To see what this means, consider how resiliency can be demonstrated using the beta distribution, though keep in mind that it is not meant to be a generalizable result, since different distributions work differently. In any case, recall that different sets of parameters can produce the same expected value. For instance, consider $Beta(1, 1)$ and $Beta(5000, 5000)$. Their expected values are, of course, the same, since $\frac{1}{2} = \frac{5000}{10000}$. But the second distribution is way much resilient than the first. Suppose these are distribution functions that represent the opinions of two different agents, but they both receive the same piece of evidence $Y$ from a Bernoulli process such that $Y = \sum_{i=1}^{5} X_i = 5$, that is, getting 5 heads in a row. We do not have to do the calculation to see that they will respond to the evidence pretty differently: the first agent will raise her posterior expected value from 1/2 to 6/7. The difference is .36, which is a considerable increase.

The second agent's opinion, however, would barely be changed:

22. Skyrms, "Resiliency, Propensities, and Causal Necessity," 707.

$$\frac{5005}{10005} - \frac{5000}{10000} = 0.00025$$

It is barely a rounding error. In fact, we can see that the second agent would have to see 200 successes *in a row* in order to raise the credence by 0.01:

$$\frac{5000 + 200}{5000 + 5000 + 200} = 0.51$$

Notice that this analysis requires counterfactual thinking in two ways: we have to consider what our priors *would* be like in different circumstances, and we have to consider its response to some hypothetical data. This is why the weight of evidence is a dispositional commitment: one is committed to respond to the the deliverance of experience in a deliberate manner, as dictated by the model decided upon after considering various counterfactual scenarios.

This way of understanding evidential weight provides an important insight on the voluntaristic idea that degrees of belief ought to be understood as taking up an epistemic commitment during the context of abduction. To deliberate on the appropriate hypothesis to be accepted provisionally, I have to know what *practical difference* its acceptance would make on my future conduct. To do so, I have to draw deductive inferences based on various possible models that I could possibly accept, based on various hypothetical scenarios that come up during experiment.

To see what I mean, it might help to see the interaction between the evidence and posterior probability directly—this is shown in the figure. The x-axis represents the number of heads in a row, so the higher x is, the more extreme the hypothetical evidence is. The y-axis is the posterior probability after receiving the x heads out of

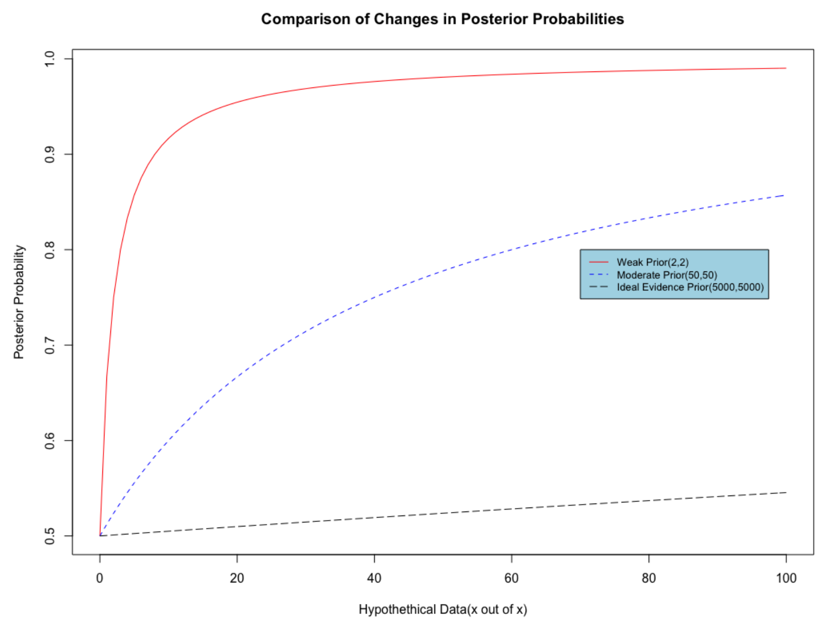


**Figure 4.2**: Comparison of stabilities of different prior distributions: values in parentheses are parameters for the beta distribution.

x throws as indicated by the x-axis.

We see that these counterfactual priors behave quite differently in slight of extreme evidence. Here, the weight of evidence clearly corresponds to the sum of the parameters $\alpha + \beta$, and the higher it is, the less responsive it is to evidence. This is especially clear when $\alpha = \beta = 500$—we see that with such as a weighty prior distribution, it makes absolutely no difference how the extreme the data is, until we get more than 60 heads in a row. Even if we tossed 100 out of 100, the expected value stays very close to 0.5. A "flat prior", i.e., $\alpha = \beta = 1$, is, as expected, not resilient against almost any form of extreme evidence. The posterior is expected to almost 0.8 after seeing 10 heads in a row, and rapidly approaches unity.

From the deliberativist point of view, there is no *a priori* justification for one over another. There are circumstances in which the extremely recalcitrant prior would be appropriate. Perhaps we can consider the probability of a person's guilt based on the evidence. In such a case, it would be rational to adopt a prior that is extremely resilient, so that the posterior would be unresponsive unless the evidence is beyond any reasonable doubt. Even in a less dramatic situation such as testing the effectiveness of a drug, a resilient prior could still be advisable when the result could mean have life-altering consequences.

This also has an implication on the voluntarist interpretation of degrees of belief. Recall that van Fraassen argues that assertions of probability should not be a description of the agent's psychological state. An argument for the voluntarist reading can be made in this context. Suppose personally I think the coin is extremely likely to be fair. It seems *less* rational for me to adopt a prior to reflect such a state, e.g., $Beta(500, 500)$; because, it would look as though I am rigging the experiment in favor of *my* hypothesis. The rational thing, as a matter of fact, should be the opposite: *because* I am confident that my opinion is true, I am intentionally adopting the opposite prior, with the assumption that the data will overwhelm it. This is akin to a gambler with inside information who is willing to make a bet with extremely unfavorable odds. As the psychologist John Kruschke points out, it might be advisable to choose a prior to satisfy a skeptic.[23]

23. John Kruschke, "Bayesian Estimation Supersedes the T Test," *Journal of Experimental Psychology: General* 142, no. 2 (2013): 573–603, 575.

## 4.7 Conclusion

The central concern of this chapter is the dynamic between the abductive and deductive contexts of inquiry. We saw that Peice makes the tantalizing suggestion that abduction is *interrogative*. To explain what this means, I borrowed Hintikka's helpful connection between the economics of choosing a hypothesis for inductive testing and choosing an assumption for premiseless proofs. Both involve a strategic element in thinking about the deductive implication of what *would* happen, had the choice been made a certain way. I tried to apply this insight on Keynes' idea of the weight of evidence, and suggested that the evidential weight represents a *dispositional commitment*, revealed through a deductive probability calculus from the chosen prior. We saw that in the case of Popper's paradox of ideal evidence, what has been changed is not the posterior probability, but our willing ness to revision in light of new evidence.

# Bibliography


Austin, J. L. *How to Do Things with Words*. Clarendon Press, 1962.

Ayer, A. J. *Probability and Evidence*. Macmillan, 1972.

Berger, James. "The robust Bayesian viewpoint." In *Robustness of Bayesian Analyses*, edited by J. D. Kadane. Elsevier Science, 1984.

Berger, James, and Robert Wolpert. *The Likelihood Principle*. Institute of Mathematical Statistics, 1988.

Burke, F. Thomas. *What Pragmatism Was*. Indiana University Press, 2013.

Carnap, Rudolf. *Empiricism, Semantic, and Ontology*. University of Chicago Press, 1947.

———. *Logical Foundations of Probability*. University of Chicago Press, 1950.

———. *The Logical Structure of the World*. Open Court Class, 2003.

———. *The Logical Syntax of Language*. K. Paul, Trench, Trubner & Co., 1937.

Christensen, David. "Clever Bookies and Coherent Beliefs." *Philosophical Review* 100, no. 2 (1991): 229–247.

de Finetti, Bruno. *Probability, Induction, and Statistics*. Wiley, 1972.

DeGroot, Morris, and Mark Schervish. *Probability and Statistics*. Addison-Wesley, 2012.

Descartes, René. *The Philosophical Writings of Descartes*. Vol. 1. Cambridge University Press, 1984.

Elga, Adam. "Reflection and Disagreement." *Noûs* 41, no. 3 (2006): 478–502.

Eriksson, Lina, and Alan Hajek. "What Are Degrees of Belief?" *Studia Logica* 86, nos. 183-213 (2007).

Feller, William. "Statistical Aspects of ESP." *The Journal of Parapsychology* 2, no. 4 (1940): 271–298.

Gelman, Andrew, and Cosma Rohilla Shalizi. "Philosophy and the Practice of Bayesian Statistics in the Social Sciences." In *The Oxford Handbook of Philosophy of Social Science*, edited by Harold Kincaid. Oxford University Press, 2012.

Good, Irving J. *Good Thinking: The Foundations of Probability and its Applications.* Univ Minnesota Pr, 1983.

———. "On the Principle of Total Evidence." *British Journal for the Philosophy of Science* 17, no. 4 (1966): 319–321.

Goodman, Nelson. "The Problem of Counterfactual Conditionals." *Journal of Philosophy* 44, no. 5 (1947): 113–128.

Greenwald, David. "The Forgetful Witness." *The University of Chicago Law Review* 160, no. 1 (1993): 167–195.

Hajek, Alan. "Dutch Book Arguments." In *The Oxford Handbook of Rational and Social Choice*, edited by Paul Anand, Prasanta Pattanaik, and Clemens Puppe. Oxford University Press, 2008.

Hintikka, Jaakko. “What Is Abduction? The Fundamental Problem of Contemporary Epistemology.” *Transactions of the Charles S. Peirce Society* 34, no. 3 (1998): 503–533.

Hookway, Christopher. *Truth, Rationality, and Pragmatism: Themes From Peirce.* Oxford University Press, 2000.

Hoover, Kevin. “Pragmatism, Pragmaticism, and Economic Method.” In *New Directions in Economic Methodology.* Routledge, 1994.

James, William. *The Will to Believe and Human Immortality.* Dover Publications, 1960.

Jaynes, E. T. “Common sense as an interface.” In *Foundations of probability theory, statistical inference, and statistical theories of science*, edited by W. L. Harper and C. A. Hooker, 218–257. Dordrecht, 1980.

Jeffrey, Richard. *Logic of Decision.* McGraw-Hill, 1965.

Joyce, James M. “A Nonpragmatic Vindication of Probabilism.” *Philosophy of Science* 65, no. 4 (1998): 575–603.

———. “How Probabilities Reflect Evidence.” *Philosophical Perspectives* 19, no. 1 (2005): 153–178.

Kahneman, Daniel, and Amos Tversky. “Prospect Theory: An Analysis of Decision Under Risk.” *Econometrica: Journal of the Econometric Society*, 1979, 263–291.

Keynes, John Maynard. *A Treatise on Probability.* Macmillan & Co., Limited., 1921.

———. *Essays in Biography.* Macmillan & Co., Limited., 1933.

Kruschke, John. "Bayesian Estimation Supersedes the T Test." *Journal of Experimental Psychology: General* 142, no. 2 (2013): 573–603.

Laplace, Pierre-Simon. *Philosophical Essay on Probabilities.* Spring, 1995.

Levi, Isaac. "Beware of Syllogism: Statistical Reasoning and Conjecturing According to Peirce." In *The Cambridge Companion to Peirce*, edited by C. J. Misak, 257. Cambridge University Press, 2004.

———. "The Weight of Argument." In *Fundamental Uncertainty: Rationality and Plausible Reasoning*, edited by Silva Marzetti Dall'Aste Brandolini and Roberto Sczzieri, 39–58. Palgrave MacMilan, 2011.

Lewis, David. "A Subjectivist's Guide to Objective Chance." In *Philosophy of Probability: Contemporary Readings*, edited by Richard C. Jeffrey, 83–132. University of California Press, 1980.

Lindley, D. V., and L. D. Phillips. "Inference for a Bernoulli Process (A Bayesian View)." *The American Statistician* 30, no. 3 (1976): 112–119.

Lipton, Peter. *Inference to the Best Explanation.* Routledge, Taylor / Francis Group, 2004.

Mauskopf, Seymour, and Michael McVaugh. *The Elusive Science.* K. Paul, Trench, Trubner & Co., 1980.

Mayo, Deborah. *Error and the Growth of Experimental Knowledge.* University of Chicago Press, 1996.

Mayo, Deborah G., and Aris Spanos. "Error Statistics." In *Handbook of the Philosophy of Science, Vol. 7: Philosophy of Statistics*, edited by Prasanta S. Bandyopadhyay and Malcolm R. Forster. Elsevier B.V., 2011.

Okasha, Samir. "Van Fraassen's Critique of Inference to the Best Explanation." *Studies in History and Philosophy of Science Part A* 31, no. 4 (2000): 691–710.

*Pascal's Pensees.* `https://www.gutenberg.org/files/18269/18269-h/18269-h.htm`. Accessed: 8-21-2018.

Peirce, Charles S. "A Theory of Probable Inference." In *Writings of Charles S. Peirce: A Chronological Edition, Volume 4: 1879–1884.* Indiana University Press, 1989.

———. *Collected Papers of Charles Sanders Peirce.* Cambridge: Harvard University Press, 1931.

———. "How to Make our Ideas Clear." In *Writings of Charles S. Peirce: A Chronological Edition, Volume 3: 1872–1878.* Indiana University Press, 1966.

———. *The Essential Peirce, Volume 2: Selected Philosophical Writings (1893-1913).* Edited by Peirce Edition Project. Indiana University Press, 1998.

———. "The Probability of Induction." In *Writings of Charles S. Peirce: A Chronological Edition, Volume 3: 1872–1878.* Indiana University Press, 1986.

Popper, Karl. *The Logic of Scientific Discovery.* Routledge, 2002.

Quine, W. V. "Epistemology Naturalized." In *Ontological Relativity and Other Essays.* New York: Columbia University Press, 1969.

———. *From Stimulus to Science.* Harvard University Press, 1997.

Raiffa, Howard, and Robert Schlaifer. *Applied Statistical Decision Theory*. Harvard, 1964.

Ramsey, Frank P. *Philosophical Papers*. Cambridge: Harvard University Press, 1990.

———. "Weight or the Value of Knowledge." *British Journal for the Philosophy of Science* 41, no. 1 (1990): 1–4.

Rhine, J. B. *Extrasensory Perception*. Branden Publishing Company, 1964.

S., Green Mitchell, and Williams John N. *Moore's Paradox: New Essays on Belief, Rationality, and the First Person*. Oxford University Press, 2007.

Savage, Leonard J. *The Foundations of Statistics*. Dover, 1954.

———. "The foundations of statistics reconsidered." *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* 1 (1961): 575–586.

Searle, John R. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, 1969.

Shulman, Robert. "A Study of Card-Guessing in Psychotic Subjects." *The Journal of Parapsychology* 2, no. 2 (1938): 95–106.

Skyrms, Brain. "Resiliency, Propensities, and Causal Necessity." *Journal of Philosophy* 74, no. 11 (1966): 704–713.

Skyrms, Brian. *The Dynamics of Rational Deliberation*. Harvard University Press, 1990.

Talbott, William. "Two Principles of Bayesian Epistemology." *Philosophical Studies* 62, no. 2 (1991): 135–150.

van Fraassen, Bas. "Against Naturalized Epistemology." In *On Quine*, edited by Paolo Leonardi and Macro Santambrogio. Cambridge University Press, 1995.

———. "Belief and the Problem of Ulysses and the Sirens." *Philosophical Studies* 77, no. 1 (1995): 7–37.

———. "Belief and the Will." *Journal of Philosophy* 81, no. 5 (1984): 235–256.

———. *Laws and Symmetry*. Oxford University Press, 1989.

———. *The Empirical Stance*. Yale University Press, 2002.

———. *The Scientific Image*. Oxford University Press, 1980.

Walley, Peter. *Statistical Reasoning with Imprecise Probabilities*. Chapman / Hall, 1991.

Weisberg, Jonathan. "Locating IBE in the Bayesian Framework." *Synthese* 167, no. 1 (2009): 125–143.

Winkler, Robert. *An Introduction to Bayesian Inference and Decision*. Probabilistic Publishing, 2010.