Inductive Logic

1. Brief Historical Background and Motivation

The idea of inductive logic as providing a general, quantitative way of evaluating arguments is a relatively modern one. Aristotle's conception of 'induction' (ἐπαγωγή) — which he contrasted with 'reasoning' (συλλογισμός) — only involved moving from particulars to universals (Kneale and Kneale (1960: 36)). This rather narrow way of thinking about inductive reasoning seems to have held sway through the Middle Ages, and into the 17th century when Francis Bacon (1620) developed an elaborate account of such reasoning. During the 18th and 19th centuries, the scope of inductive reasoning began to broaden considerably with the advent of more sophisticated inductive techniques (e.g., those of Mill (1843)), and with precise mathematical accounts of the notion of probability. Intuitive and quasi-mathematical notions of probability had long been used to codify various aspects of uncertain reasoning in the contexts of games of chance and statistical inference (see Stigler (1986) and Dale (1999)), but a more abstract and formal approach to probability theory would be necessary to formulate the general modern inductive-logical theories of non-demonstrative inference. In particular, the pioneering work in probability theory by Bayes (1764), Laplace (1812), Boole (1854) and many others in the 18th and 19th centuries laid the groundwork for a much more general framework for inductive reasoning. (See PROBLEM OF INDUCTION for parallel historical developments in philosophical thinking about the possibility of inductive knowledge, most famously articulated by Hume (1739, 1740, 1758).)

The contemporary idea of inductive logic (as a general, logical theory of argument evaluation) did not begin to appear in a mature form until the late 19th and early 20th centuries. Some of the most eloquent articulations of the basic ideas behind inductive logic in this modern sense appear in John Maynard Keynes' *Treatise on Probability*. Keynes (1921: 8) describes a "logical relation between two sets of propositions in cases where it is not possible to argue demonstratively from one to another." Nearly thirty years later, Rudolf Carnap (1950) published his encyclopedic work *Logical Foundations of Probability* in which he very clearly explicates the idea of an inductive-logical relation called "confirmation" which is a quantitative generalization of deductive entailment. (See also CONFIRMATION THEORY.) The following quote from Carnap (1950) gives some insight into the modern project of inductive logic and its relation to classical deductive logic:

> Deductive logic may be regarded as the theory of the relation of logical consequence, and inductive logic as the theory of another concept [$\mathfrak{c}$] which is likewise objective and logical, viz., … degree of confirmation. (43)

More precisely, the following three fundamental tenets have been accepted by the vast majority of proponents of modern inductive logic

1. Inductive logic should provide a quantitative generalization of (classical) deductive logic. That is, the relations of deductive entailment and deductive refutation should be captured as limiting (extreme) cases with cases of 'partial entailment' and 'partial refutation' lying somewhere on a continuum (or range) between these extremes.

2. Inductive logic should use probability (in its modern sense) as its central conceptual building block.
3. Inductive logic (*i.e.*, the non-deductive relations between propositions that are characterized by inductive logic) should be objective and logical.

(See Skyrms (2000: chapter 2) for a contemporary overview.) In other words, the aim of inductive logic is to characterize a quantitative relation (of inductive strength or confirmation) $\mathfrak{c}$ which satisfies desiderata (1)–(3) above. The first two of these desiderata are relatively clear (or will quickly become clear below). The third desideratum is less clear. What does it mean for the quantitative relation $\mathfrak{c}$ to be objective and logical? Carnap (1950) explains his understanding of (3) as follows (brackets added):

> That $\mathfrak{c}$ is an objective concept means this: if a certain $\mathfrak{c}$ value holds for a certain hypothesis with respect to a certain evidence, then this value is entirely independent of what any person may happen to think about these sentences, just as the relation of logical consequence is independent in this respect. (43)
>
> The principal common characteristic of the statements in both fields [*viz.*, deductive and inductive logic] is their independence of the contingency of facts [*viz.*, facts of nature]. This characteristic justifies the application of the common term 'logic' to both fields. (200)

The remaining sections of this article will examine a few of the prevailing modern theories of inductive logic, and discuss how they fare with respect to these three central desiderata. The meaning and significance of these desiderata will be clarified, and the received view about inductive logic critically evaluated.

2. Inductive Logic — the Basic Ideas

2.1. Some Basic Terminology and Machinery for Inductive Logic

It is often said (e.g., in many contemporary introductory logic texts) that there are two kinds of arguments: deductive and inductive, where the premises of deductive arguments are intended to guarantee the truth of their conclusions, while inductive arguments involve some risk of their conclusions being false even if all of their premises are true (see, e.g., Hurley (2003)). It seems better to say that there is just one kind of argument: an argument is a set of propositions, one of which is the conclusion, the rest premises. There are many ways of evaluating arguments. Deductive logic offers strict, qualitative standards of evaluation—the conclusion either follows from the premises or it does not; whereas, inductive logic provides a finer-grained (and thereby more liberal) quantitative range of evaluation standards for arguments. (One can also define comparative and/or qualitative notions of inductive support or confirmation. Carnap (1950: §8) and Hempel (1945) both provide penetrating discussions of quantitative vs. comparative/qualitative notions of confirmation and/or inductive support. For simplicity, our focus will be on quantitative approaches to inductive logic, but most of the main issues and arguments discussed below can be recast in comparative or qualitative terms.)

Let $\{P_1, \ldots, P_n\}$ be a finite set of propositions constituting the premises of an (arbitrary) argument, and let C be its conclusion. Deductive logic aims to explicate the concept of *validity* (*i.e.*, 'deductive-logical goodness') of arguments. Inductive logic aims to explicate a quantitative generalization of this deductive concept. This generalization is often called the "inductive strength" of an argument. (Carnap (1950) uses the word "confirmation" here.) Following Carnap, the notation $\mathfrak{c}(C, \{P_1, \ldots, P_n\})$ will denote the degree to which $\{P_1, \ldots, P_n\}$ jointly inductively support (or "confirm") C.

As desideratum (2) indicates, the concept of probability is central to the modern project of inductive logic. The notation Pr(•) and Pr(• | •) will denote unconditional and conditional probability functions, respectively. Informally (and roughly), "Pr(p)" can be read "the probability that proposition p is true", and "Pr(p | q)" can be read "the probability that proposition p is true, given that proposition q is true". The nature of probability functions and their relation to the project of inductive logic will be a central theme in what follows.

2.2. A Naive Version of Basic Inductive Logic, and The Received View

According to classical deductive propositional logic, the argument from $\{P_1, \ldots, P_n\}$ to C is *valid* iff the material conditional $(P_1 \;\&\; \ldots \;\&\; P_n) \supset C$ is (logically) necessarily true. Naively, one might try to define "inductively strong" as follows: the argument from $\{P_1, \ldots, P_n\}$ to C is *inductively strong* iff the material conditional $(P_1 \;\&\; \ldots \;\&\; P_n) \supset C$ is (logically?) probably true. More formally, one can express this Naive Inductive Logic (NIL) proposal as follows:

(NIL) $\mathfrak{c}(C, \{P_1, \ldots, P_n\})$ is high iff $\Pr((P_1 \;\&\; \ldots \;\&\; P_n) \supset C)$ is high.

There are problems with this first, naive attempt to use probability to generalize deductive validity quantitatively. As Skyrms (2000: 19–22) points out, there are (intuitively) cases in which the material conditional $(P_1 \;\&\; \ldots \;\&\; P_n) \supset C$ is probable but the argument from $\{P_1, \ldots, P_n\}$ to C is not a strong one. Skyrms (2000: 21) gives the following example:

(P) There is a man in Cleveland who is 1999 years and 11-months-old and in good health.

(C) No man will live to be 2000 years old.

Skyrms argues that $\Pr(P \supset C)$ is high, simply because Pr(C) is high, and not because there is any evidential relation between P and C. Indeed, intuitively, the argument from (P) to (C) is not strong, since (P) seems to disconfirm or counter-support (C). Thus, $\Pr((P_1 \;\&\; \ldots \;\&\; P_n) \supset C)$ being high is not sufficient for $\mathfrak{c}(C, \{P_1, \ldots, P_n\})$ being high. Note, also, that $\Pr((P_1 \;\&\; \ldots \;\&\; P_n) \supset C)$ cannot serve as $\mathfrak{c}(C, \{P_1, \ldots, P_n\})$, since it violates desideratum (1). If $\{P_1, \ldots, P_n\}$ refutes C, then $\Pr((P_1 \;\&\; \ldots \;\&\; P_n) \supset C) = \Pr({\sim}(P_1 \;\&\; \ldots \;\&\; P_n))$, which is not minimal, since the conjunction of the premises of an argument need not have probability one.

Skyrms suggests that the mistake (NIL) makes is one of conflating the probability of the material conditional: $\Pr((P_1 \;\&\; \ldots \;\&\; P_n) \supset C)$ with the conditional probability of C, given $P_1 \;\&\; \ldots \;\&\; P_n$: $\Pr(C \mid P_1 \;\&\; \ldots \;\&\; P_n)$. And, according to Skyrms, it is the latter which should be used as a definition of $\mathfrak{c}(C, \{P_1, \ldots, P_n\})$. The reason for this preference is that $\Pr((P_1 \;\&\; \ldots \;\&\; P_n) \supset C)$ fails to capture the evidential relation between the premises and conclusion, since $\Pr((P_1 \;\&\; \ldots \;\&\; P_n) \supset C)$ can be high solely in virtue of the unconditional probability of (C) being high or solely in virtue of the unconditional probability of $P_1 \;\&\; \ldots \;\&\; P_n$ being low. As Skyrms (2000: 20) stresses, $\mathfrak{c}(C, \{P_1, \ldots, P_n\})$ should measure the "evidential relation between the premises and the conclusion." This leads Skyrms (and many others) to defend the following account, which might be called "The Received View" (TRV) about inductive logic:

(TRV) $\mathfrak{c}(C, \{P_1, \ldots, P_n\}) = \Pr(C \mid P_1 \,\&\, \ldots \,\&\, P_n)$

The idea that $\mathfrak{c}(C, \{P_1, \ldots, P_n\})$ should be identified with the conditional probability of C, given $P_1 \,\&\, \ldots \,\&\, P_n$ has been nearly universally accepted by inductive logicians since the inception of the contemporary discipline (recent pedagogical advocates of (TRV) include Copi and Cohen (2001), Hurley (2003), and Layman (2002), and historical champions of various versions of (TRV) include Keynes (1921), Carnap (1950), Kyburg (1970), Skyrms (2000), and many others). There are nevertheless some compelling reasons to doubt the correctness of (TRV) These reasons, which are analogous to Skyrms's reasons for rejecting (NIL), will be discussed below. But, before one can adequately assess the merits of (NIL), (TRV), and other proposals concerning inductive logic, one needs to say more about probability models and their relation to inductive logic. (For greater detail, see PROBABILITY.)

3 Probability, its Interpretation, and its Role in Traditional Inductive Logic

3.1 The Mathematical Theory of Probability

For present purposes, assume that a probability function Pr(•) is a finitely additive measure function over a Boolean algebra of propositions (or sentences in some formal language). That is, assume that Pr(•) is a function from a Boolean algebra B of propositions (or sentences) to the unit interval [0,1] satisfying the following three axioms (this is Kolmogorov's (1950) Axiomatization), for all propositions X and Y in B:

i. $\Pr(X) \geq 0$.
ii. If X is a (logically) necessary truth, then $\Pr(X) = 1$.
iii. If X and Y are mutually exclusive, then $\Pr(X \vee Y) = \Pr(X) + \Pr(Y)$.

And, following Kolmogorov, define conditional probability Pr(• | •) in terms of unconditional probability Pr(•), as follows:

Definition. $\Pr(X \mid Y) = \Pr(X \,\&\, Y) / \Pr(Y)$, provided that $\Pr(Y) \neq 0$.

A *probability model* $M = \langle B, \Pr_M \rangle$ consists of a Boolean algebra B of propositions (or sentences in some language), together with a particular probability function $\Pr_M(\bullet)$ over the elements of B.

These axioms (and definition) say what the mathematical properties of probability models are, but they do not say anything about the interpretation or application of such models. The latter issue is philosophically more central, and more controversial than the former (although, see Popper (1992: appendix *iv), Roeper and Leblanc (1999), and Hájek (2003) for dissenting views on the formal theory of conditional probability). There are various ways in which one can interpret or understand probabilities. (See Hájek (2002) and PROBABILITY for a thorough discussion.) The two interpretations that are most commonly encountered in the context of applications to inductive logic are the so-called "epistemic" and "logical" interpretations of probability.

3.2 Epistemic Interpretations of Probability

On epistemic interpretations of probability, $\Pr_M(H)$ is (roughly) the degree of belief an epistemically rational agent assigns to H, according to a probability model M of the agent's epistemic state. A rational agent's background knowledge K is assumed (in orthodox theories of epistemic probability) to be "included" in any epistemic probability model M, and therefore K is assumed to have an unconditional probability of 1 in M. $\Pr_M(H \mid E)$ is the degree of belief an epistemically rational agent assigns to H upon learning that E is true (or on the supposition that E is true – see Joyce (1999: Chapter 6) for discussion), according to a probability model M of the agent's epistemic state. According to standard theories of epistemic probability, agents learn by conditionalizing on evidence. So, roughly speaking, (the probabilistic structure of) a rational agent's epistemic state evolves (in time t) through a series of probability models $\{M_t\}$, where evidence learned at time t has probability 1 in all subsequent models $\{M_{t'}\}$, $t' > t$.

Keynes (1921: 4) seems to be employing an epistemic interpretation of probability in his inductive logic when he says

> Let our premises consist of any set of propositions *h*, and our conclusion consist of any set of propositions *a*, then, if a knowledge of *h* justifies a rational degree of belief in *a* of degree *x*, we say that there is a *probability-relation* of degree *x* between *a* and *h* [$\Pr(a \mid h) = x$].

It is not obvious that (TRV) can satisfy desideratum (3) — that $\mathfrak{c}$ be logical and objective — if the probability function Pr that is used to explicate $\mathfrak{c}$ in (TRV) is given an epistemic interpretation of this kind. After all, whether "a knowledge of *h* justifies a rational degree of belief in *a* of degree *x*" seems to depend on what one's background knowledge K is. And, while this is arguably an objective fact, it also seems to be a contingent fact, and not something that could be determined a priori (on the basis of *a* and *h* alone). As Keynes (1921: 4) explains, his probability function $\Pr(a \mid h)$ is not subjective, since "once the facts are given which determine our knowledge [*viz*., background + *h*], what is probable or improbable [*viz*., *a*] in these circumstances has been fixed objectively, and is independent of our opinion." But, he later suggests that the function is contingent on what the agent's background knowledge K is, in the sense that $\Pr(a \mid h)$ can vary "depending upon the knowledge to which it is related." Carnap (1950: §45B) is keenly aware of this problem. Carnap suggests that Keynes should have characterized $\Pr(a \mid h)$ as the degree of belief in *a* that is justified by knowledge of *h — and nothing else* (the reader may want to ponder what it might mean for an agent to "know *h and nothing else*"). As Keynes' remarks suggest (and as Maher (1996) explains), the problem is even deeper than this, since even a complete specification of an agent's background knowledge K may not be sufficient to pick out a unique (rational) epistemic probability model M for an agent (Keynes' reaction to this was to conclude that sometimes quantitative judgments of inductive strength or degree of conditional probability are not possible, and that in these cases we must settle for qualitative or comparative judgments). The problem here is that "$\Pr(X \mid K)$" ("the probability of X, given background knowledge K") will not (in general) be determined, unless an epistemic probability model M is specified, which (*a fortiori*) gives $\Pr_M(X)$, for each X in M. And, without a determination of these fundamental or "a priori" probabilities $\Pr_M(X)$, a general (quantitative) theory of inductive logic based on epistemic probabilities seems all but hopeless. This raises the problem of specifying an appropriate "a priori" probability model M. Keynes

(1921: chapter 4) and Carnap (see below) both look to the Principle of Indifference at this point, as a guide to choosing "a priori" probability models. Before we discuss the role of the Principle of Indifference, we must first address logical interpretations of probability.

3.3 Logical Interpretations of Probability

Philosophers who accepted (TRV) and who were concerned about the inductive-logical ramifications (mainly, regarding the satisfaction of desideratum (3)) of interpreting probabilities epistemically began to formulate logical interpretations of probability. On logical interpretations of probability, conditional probabilities Pr(X | Y) are themselves understood as quantitative generalizations of an (logical) entailment (or deducibility) relation between propositions Y and X. The motivation for this should be clear. This seems like the most direct way to guarantee that a (TRV)-type theory of inductive logic will satisfy desideratum (3). If Pr(• | •) is itself logical, then c(•,•) — which, according to (TRV), is defined as Pr(• | •) — should also be logical, and the satisfaction of desideratum (3) seems automatic (and, of course, the other two desiderata would automatically be satisfied by this move as well). Below it will become clear that "(TRV) + logical probability" is not the only way (and not necessarily the best way) to satisfy the three desiderata for providing an adequate account of the logical relation of inductive support. In preparation, the notion of logical probability must be examined in some detail.

Typically, logical interpretations of probability attempt to define Pr(*q* | *p*) — where *p* and *q* are sentences in some formal first-order language L — in terms of the syntactical features of *p* and *q* (in L). The most famous logical interpretations of probability are those of Carnap (1950, 1952, 1971, 1980). A sketch of the simplest of Carnap's early (1950) constructions will give the main ideas. (It is interesting to note that Carnap's (1950) and (1952) systems are almost identical to systems described 20–30 years earlier by W.E. Johnson (1921) and (1932), respectively. See Paris (1994: chapter 12) and Kyburg (1970: chapter 5) for a rigorous technical development of Carnap's (1952) — and W.E. Johnson's (1932) — continuum; see Maher (2000, 2001) for a detailed critical discussion of Carnap's more recent (1971, 1980) two-dimensional continua, which are increasingly complicated, and less tightly coupled with the syntax of L; see Glaister (2001) and Festa (1993) for broader surveys of Carnapian theories of logical probability and inductive logic; and, see Skyrms (1996) for discussion of some recent applications of "Carnapian" techniques to Bayesian statistical models involving continuous random variables.)

Begin with a standard first-order logical language L containing a finite number of monadic predicates: F, G, H, …, and a finite or denumerable number of individual constants *a*, *b*, *c*, …. Define an unconditional probability function Pr(•) over the sentences of L. Finally, following the standard Kolmogorovian approach, construct a conditional probability function Pr(• | •) over pairs of sentences of L, using the ratio definition of conditional probability, above. To fix ideas, consider a very simple toy language L with only two monadic predicates 'F' and 'G' and only two individual constants 'a' and 'b'. In this language, there are only 16 possible states of the world that can be described. These 16 maximally specific descriptions are called the *state descriptions* of L, and they are as follows:

Fa & Ga & Fb & Gb	~Fa & Ga & Fb & Gb
Fa & Ga & Fb & ~Gb	~Fa & Ga & Fb & ~Gb
Fa & Ga & ~Fb & Gb	~Fa & Ga & ~Fb & Gb
Fa & Ga & ~Fb & ~Gb	~Fa & Ga & ~Fb & ~Gb

Fa & ~Ga & Fb & Gb
Fa & ~Ga & Fb & ~Gb
Fa & ~Ga & ~Fb & Gb
Fa & ~Ga & ~Fb & ~Gb
~Fa & ~Ga & Fb & Gb
~Fa & ~Ga & Fb & ~Gb
~Fa & ~Ga & ~Fb & Gb
~Fa & ~Ga & ~Fb & ~Gb

Two state descriptions S1 and S2 are said to be *permutations* of each other if S1 can be obtained from S2 by some permutation of the individual constants. For instance, 'Fa & ~Ga & ~Fb & Gb' can be obtained from '~Fa & Ga & Fb & ~Gb' by permuting 'a' and 'b'. Thus, 'Fa & ~Ga & ~Fb & Gb' and '~Fa & Ga & Fb & ~Gb' are permutations of each other (in L). A *structure description* in L is a disjunction of state descriptions, each of which is a permutation of the others. In our toy language L, we have the following 10 structure descriptions:

Fa & Ga & Fb & Gb
(Fa & Ga & Fb & ~Gb) ∨ (Fa & ~Ga & Fb & Gb)
(Fa & Ga & ~Fb & Gb) ∨ (~Fa & Ga & Fb & Gb)
(Fa & Ga & ~Fb & ~Gb) ∨ (~Fa & ~Ga & Fb & Gb)
Fa & ~Ga & Fb & ~Gb
(Fa & ~Ga & ~Fb & Gb) ∨ (~Fa & Ga & Fb & ~Gb)
(Fa & ~Ga & ~Fb & ~Gb) ∨ (~Fa & ~Ga & Fb & ~Gb)
~Fa & Ga & ~Fb & Gb
(~Fa & Ga & ~Fb & ~Gb) ∨ (~Fa & ~Ga & ~Fb & Gb)
~Fa & ~Ga & ~Fb & ~Gb

Now assign non-negative real numbers to the state descriptions, so that these 16 numbers sum to one. Any such assignment will constitute an unconditional probability function Pr(•) over the state descriptions of L. To extend Pr(•) to the entire language L, stipulate that the probability of a disjunction of mutually exclusive sentences is the sum of the probabilities of its disjuncts. And, since every sentence in L is equivalent to some disjunction of state descriptions, and every pair of state descriptions is mutually exclusive, this gives a complete unconditional probability function Pr(•) over L. For instance, since 'Fa & Ga & ~Gb' is equivalent to the disjunction '(Fa & Ga & Fb & ~Gb) ∨ (Fa & Ga & ~Fb & ~Gb)' we will have:

$$\begin{aligned}\Pr(Fa \,\&\, Ga \,\&\, {\sim}Gb) &= \Pr((Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb) \vee (Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb)) \\ &= \Pr(Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb) + \Pr(Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb)\end{aligned}$$

Now, it is only a brief step to the definition of the conditional probability function Pr(• | •) over pairs of sentences in L. Using the standard, Kolmogorovian ratio definition of conditional probability, we have, for all pairs of sentences X, Y in L:

$$\Pr(X \mid Y) = \Pr(X \,\&\, Y) / \Pr(Y), \text{ provided that } \Pr(Y) \neq 0$$

Thus, once the unconditional probability function Pr(•) is specified for the state descriptions of a language L, all probabilities both conditional and unconditional are thereby determined over L. And, this gives us a "logical probability model" M over the language L. The unconditional, logical probability functions so defined are typically called *measure functions*. Carnap (1950) discusses two "natural" measure functions.

The first Carnapian measure function is $\mathfrak{m}^\dagger$, which assumes that each of the state descriptions is equiprobable a priori: if there are N state descriptions in L, then $\mathfrak{m}^\dagger$ assigns 1/N to each state description. While this may seem like a very natural measure function, since it applies something like the Principle of Indifference to the state descriptions of L (see below for discussion), $\mathfrak{m}^\dagger$ has the consequence that the resulting probabilities cannot reflect learning from experience. Consider the following simple example. Assume that you adopt a logical probability function Pr(•) based on $\mathfrak{m}^\dagger$ as your own *a priori* degree of belief (or credence) function. Then, you learn (by conditionalizing) that an object a is F [*i.e.*, you learn that Fa]. Intuitively, your conditional degree of credence Pr(Fb | Fa) that a distinct object b also has F, given that a has F should not always be the same as your *a priori* degree of credence that b is F. That is, the fact that you have observed another F object should be capable (at least in some cases) of making it more probable (*a posteriori*) that b will also have F (*i.e.*, more probable than Fb was *a priori*). More generally, if you observe that a large number of objects have been F, this should be capable of raising the probability that the next object you observe will also be F. Unfortunately, no *a priori* probability function based on $\mathfrak{m}^\dagger$ is consistent with learning from experience in either sense. To see this, consider the simple case Pr(Fb | Fa):

$$\Pr(Fb \mid Fa) = \mathfrak{m}^\dagger(Fb \;\&\; Fa) \,/\, \mathfrak{m}^\dagger(Fa) = 1/2 = \mathfrak{m}^\dagger(Fb) = \Pr(Fb)$$

So, if one assumes an *a priori* probability function based on $\mathfrak{m}^\dagger$, the fact that one object has property F cannot affect the probability that any other object will also have property F. Indeed, it can be shown (see Kyburg (1970: 58-59)) that no matter how many objects are assumed to be F, this will be irrelevant (according to probability functions based on $\mathfrak{m}^\dagger$) to the hypothesis that a distinct object will also be F.

The fact that (on the probability functions generated by the measure $\mathfrak{m}^\dagger$) no object's having certain properties can be informative about other objects also having those same properties has been viewed as a serious shortcoming of $\mathfrak{m}^\dagger$. (See Carnap (1955) for discussion.) As a result, Carnap formulated an alternative measure function $\mathfrak{m}^*$, which is defined as follows. First, assign equal probabilities to each structure description (in the toy language above, 1/10). Then, each state description belonging to a given structure description is assigned an equal portion of the probability assigned to that structure description). For instance, in the toy language, the state description 'Fa & Ga & ~Fb & Gb' gets assigned an *a priori* probability of 1/20 (1/2 of 1/10), but the state description 'Fa & Ga & Fb & Gb' receives an *a priori* probability of 1/10 (1/1 of 1/10). To further illustrate the differences between $\mathfrak{m}^\dagger$ and $\mathfrak{m}^*$, here are some numerical values in the toy language L:

Measure Function $\mathfrak{m}^\dagger$	Measure Function $\mathfrak{m}^*$
$\mathfrak{m}^\dagger$(Fa & Ga & ~Fb & Gb) = 1/16	$\mathfrak{m}^*$(Fa & Ga & Fb & Gb) = 1/10
$\mathfrak{m}^\dagger$((Fa & Ga & ~Fb & Gb) ∨ ((~Fa & Ga & Fb & Gb))) = 1/8	$\mathfrak{m}^*$(Fa & Ga & ~Fb & Gb) = 1/20
$\mathfrak{m}^\dagger$(Fa) = 1/2	$\mathfrak{m}^*$(Fa) = 1/2
$\Pr^\dagger(Fa \mid Fb) = 1/2 = \mathfrak{m}^\dagger(Fa) = \Pr^\dagger(Fa)$	$\Pr^*(Fa \mid Fb) = 3/5 > 1/2 = \mathfrak{m}^*(Fa) = \Pr^*(Fa)$

So, unlike $\mathfrak{m}^\dagger$, $\mathfrak{m}^*$ can model "learning from experience," since in the simple language $\Pr(Fa \mid Fb) = 3/5 > 1/2 = \Pr(Fa)$, if the probability function Pr is defined in terms of the logical measure function $\mathfrak{m}^*$. Although $\mathfrak{m}^*$ does have some advantages over $\mathfrak{m}^\dagger$, even $\mathfrak{m}^*$ can give counterintuitive results in more complex languages. (See Carnap (1952) for discussion.)

Carnap (1952) presents a more complicated framework (one very much like that reported by W.E. Johnson (1932) 20 years earlier), which describes a more general class (or "continuum") of conditional probability functions (from which the definitions of Pr(• | •) in terms of $\mathfrak{m}^*$ and $\mathfrak{m}^\dagger$ fall out as special cases). Carnap's (1952) continuum of conditional probability functions depends on a parameter λ which is supposed to reflect the "speed" with which learning from experience is possible. In this continuum, $\lambda = 0$ corresponds to the "straight rule" of induction, which says that the probability that the next object observed will be F, conditional upon a sequence of past observations is simply the frequency with which F objects have been observed in the past sequence; $\lambda = +\infty$ yields a conditional probability function much like that given above by assuming the underlying logical measure $\mathfrak{m}^{\dagger}$ (*i.e.*, $\lambda = +\infty$ implies that there is no learning from experience); and, setting $\lambda = \kappa$ (where κ is the number of independent families of predicates in Carnap's more elaborate 1952 linguistic framework) yields a conditional probability function equivalent to that generated by the measure function $\mathfrak{m}^{*}$.

Even Carnap's more elaborate (1952) λ–continuum has problems. First, none of the Carnapian systems allow universal generalizations to have non-zero probability. This problem was addressed by Hintikka (1966) and Hintikka and Niiniluoto (1980) who provided various alterations of the Carnapian framework that allow for non-zero probabilities of universal generalizations. Moreover, Carnap's early systems did not allow for the probabilistic modeling of analogical effects. That is, in Carnap's (1950, 1952) systems, the fact that two objects share several properties in common is always irrelevant to whether they share any other properties in common. Carnap's most recent (and most complex) theories of logical probability (1971, 1980) include two additional adjustable parameters (γ and η) designed to provide the theory with enough flexibility to overcome these (and other) limitations. Unfortunately, no Carnapian logical theory of probability to date has successfully dealt with the problem of analogical effects. (See Maher (2000, 2001) for extended discussion.) Moreover, as Putnam (1963) explains, there are further (and some say deeper) problems with Carnapian (or, more generally, syntactical) approaches to logical probability, if they are to be applied to inductive inference generally. The consensus now seems to be that the Carnapian project of characterizing an adequate logical theory of probability is (by his own standards and lights) not very promising. (See Putnam (1963), Festa (1993), and Maher (2001) for discussion.)

The present discussion glosses over technical details in the development of (Carnapian) logical interpretations or theories of probability since 1950 because the present article is about inductive logic, rather than logical probability. To recap: what is important for present purposes is that Carnap (along with the other advocates of "logical probability") was a (TRV)-theorist about inductive logic. He identified the concept $\mathfrak{c}(\bullet,\bullet)$ of inductive strength (or inductive support) with the concept of conditional probability $\Pr(\bullet \mid \bullet)$. And, he thought (partly, because of the problems he saw with epistemic interpretations) that in order for a (TRV) account to satisfy desideratum (3), it needed to presuppose a logical interpretation (or theory) of probability. This lead him, initially, to develop various logical measures (*e.g.*, the *a priori* logical probability functions $\mathfrak{m}^{\dagger}$ and $\mathfrak{m}^{*}$), and then to define conditional logical probability $\Pr(\bullet \mid \bullet)$ in terms of these underlying *a priori* logical measures, using the standard ratio definition. This approach ran into various problems when it came to the application of $\Pr(\bullet \mid \bullet)$ to inductive logic. These difficulties mainly had to do with the ability of Carnap's $\Pr(\bullet \mid \bullet)$ to undergird learning from experience and/or certain kinds of analogical reasoning (for other philosophical objections to Carnap's logical probability project, see Putnam (1963)). In response to these difficulties, Carnap began to fiddle directly with the definition of $\Pr(\bullet \mid \bullet)$. In 1952 Carnap moved to a parameterized definition of $\Pr(\bullet \mid \bullet)$, which contained an "index of inductive caution" (λ) that was supposed to regulate the "speed" with which learning from experience is reflected by $\Pr(\bullet \mid \bullet)$. Later, Carnap (1971, 1980) added further parameters (γ and η) to the definition of $\Pr(\bullet \mid \bullet)$ in an attempt to further generalize the theory, and to allow for sensitivity to certain kinds of analogical effects.

Ultimately, no such theory was ever viewed by Carnap (or others) as fully adequate for the purposes of grounding a (TRV) conception of inductive logic.

At this point, it is important to ask the following question: in what sense are Carnap's theories (especially his later ones) of logical probability *logical*? His early theories (based on the measure functions $\mathfrak{m}^{\dagger}$ and $\mathfrak{m}^{*}$) applied something like the Principle of Indifference (PI) to the state and/or structure descriptions of the formal language L in order to determine the logical probabilities Pr(• | •). In this sense, these early theories assume that certain sentences of L are equiprobable *a priori*. Why is such an assumption *logical*? Or, more to the point, how is *logic* supposed to tell us which statements are equiprobable *a priori*? Carnap (1955: 22) explains that:

> …the statement of equiprobability to which the principle of indifference leads is, like all other statements of inductive probability, not a factual but a logical statement. If the knowledge of the observer does not favor any of the possible events, then with respect to this knowledge as evidence they *are* equiprobable. The statement assigning equal probabilities in this case does not assert anything about the facts, but merely the logical relations between the given evidence and each of the hypotheses; namely, that these relations are logically alike. These relations are obviously alike if the evidence has a symmetrical structure with respect to their possible events. The statement of equiprobability asserts nothing more than the symmetry.

Carnap seems to be saying that (PI) is only to be applied to possible events which exhibit certain *a priori* symmetries with respect to some rational agent's background evidence. But, this appears no more logical than Keynes' epistemic approach to probability. It seems that the resulting probabilities Pr(• | •) will not be logical in the sense Carnap desired (at least, no more so than Keynes' epistemic probabilities were), unless Carnap can motivate — on logical grounds — the choice of an a priori probability model. To that end, Carnap's application of (PI) is not very useful. Recall that the goal of Carnap's project (the project of inductive logic) was to explicate the confirmation relation, which is itself supposed to reflect the evidential relation between premises and conclusions (Carnap (1950) uses the locutions "degree of confirmation" and "weight of evidence" synonymously). How are we to understand what it means for evidence not to "favor any of the possible events" in a way that does not require us to already understand how to measure the degree to which the evidence confirms each of the "possible events?" Here, Carnap's discussion of the (PI) presupposes that degree of confirmation is to be identified with degree of conditional probability. On that reading, "not favoring" just means "conferring equal probability on", and Carnap's unpacking of the (PI) reduces directly to a mathematical truth (which, for Carnap, is good enough to render the (PI) a *logical* principle). If we had independent grounds for thinking that conditional probabilities were the right way to measure confirmation (or weight of evidence), then Carnap would have a rather clever (albeit not terribly informative) way to (logically) ground his choice of a priori probability models. Unfortunately, as we will see below, there are independent reasons to doubt Carnap's presupposition here that degree of confirmation should be identified with degree of conditional probability. And, without that assumption, Carnap's (PI) is no longer a logical principle (by his own lights), and the problem of the contingency (non-logicality) of the ultimate inductive-logical probability assignments returns with a vengeance. (There are independent and deep problems with any attempt to consistently apply the Principle of Indifference to contexts in which hypotheses and/or evidence involve continuous magnitudes. See Van Fraassen (1989: chapter 12) for extended discussion.)

Carnap's later theories (1952, 1971, 1980) of Pr(• | •) introduce even further contingencies, in the form of adjustable parameters, the "proper values" of which do not seem to be determinable *a priori*. In particular, consider Carnap's (1952) λ-continuum. The parameter λ is supposed to indicate how sensitive Pr(• | •) is to "learning from experience." A higher value of λ indicates "slower learning," and a lower λ indicates "faster learning." As Carnap (1952) concedes, no one value of λ is "best *a priori*". Presumably, different values of λ are appropriate for different contexts in which confirmational judgments are made. (See Festa (1993) for a contextual Carnapian approach to confirmation.) It seems that the same must be said for the additional parameters γ and η (added to further generalize Pr(• | •), and to provide it with the ability to cope with analogical and other effects) that appear in Carnap's later systems (1971, 1980). The moral here seems to be that it is only relative to a particular assignment of values to λ, γ, and η that probabilistic (and/or confirmational) judgments are objectively and non-contingently determined in Carnap's later systems. This is analogous to the fact that it is only relative to a (probabilistic) characterization of the agent's background knowledge and complete epistemic state — in the form of a specific epistemic probability model M — that Keynes' epistemic probabilities (or Carnap's measure functions $\mathfrak{m}^*$ and $\mathfrak{m}^\dagger$) have a chance of being objectively and non-contingently determined.

A pattern is developing. Both Keynes and Carnap give accounts of "a priori" probability functions Pr(• | •) which involve certain contingencies and indeterminacies. They each feel pressure (owing to desideratum (3)) to eliminate these contingencies when the time comes to use Pr(• | •) as an explication of $\mathfrak{c}$(•,•). The general strategy for rendering these probabilities "logical" is to choose some privileged, "a priori" probability model. Here, both Keynes and Carnap appeal to the Principle of Indifference (PI) to constrain the ultimate choice of model. Carnap is sensitive to the fact that the (PI) doesn't seem like a logical principle, but his attempts to render (PI) logical (and useful for grounding the choice of a priori probability model) are both unconvincing and uninformative. There is a much easier and more direct way to guarantee the satisfaction of desideratum (3). Why not just define $\mathfrak{c}$ as a three-place relation — depending on premises, conclusion, *and a particular probability model* — from the beginning?

The next section describes a simple, general recipe (along the lines suggested by the preceding considerations) for formulating probabilistic inductive logics in such a way that they transparently satisfy desiderata (1)–(3). This section will also address the following question: Is (TRV) *materially* adequate as an account of inductive strength or inductive support? This will lead to a fourth material desideratum for measures of inductive support, and ultimately to a concrete alternative to The Received View.

4. Rethinking The Received View (TRV)

4.1 How to Ensure the Transparent Satisfaction of Desideratum (3)

The existing attempts to use the notion of probability to explicate the concept of inductive support (or inductive strength) $\mathfrak{c}$ have foundered on the question of their contingency (which threatened violation of desideratum (3)). It may be that these contingencies can be eliminated (in general) only by making the notion of inductive support explicitly relational. To follow such a plan, in the case of (TRV) one should rather say:

(TRVr) The inductive strength of the argument from $\{P_1, \ldots, P_n\}$ to C — relative to a probability model $M = \langle B, Pr_M \rangle$ — is $Pr_M(C \mid P_1 \& \ldots \& P_n)$.

Relativizing judgments of inductive support to particular probability models fully and transparently eliminates the contingency and indeterminacy of these judgments. It is clear that (TRVr) satisfies all three desiderata, since (1) $Pr_M(C \mid P_1 \& \ldots \& P_n)$ is maximal and constant when $\{P_1, \ldots, P_n\}$ entails C, and $Pr_M(C \mid P_1 \& \ldots \& P_n)$ is minimal and constant when $\{P_1, \ldots, P_n\}$ refutes C; (2) the relation of inductive support is defined in terms of the notion of probability; and, (3) once the conditional probability function $Pr_M(\bullet \mid \bullet)$ is specified (as it is, *a fortiori*, once the probability model M is specified), its values are determined objectively and in a way that is only contingent on certain mathematical facts about the probability calculus. That is, the resulting $\mathfrak{c}$-values are determined mathematically (and, for someone with logicist leanings like Carnap, perhaps even *logically*) by the specification of a particular probability model M.

One might respond at this point by asking: "Okay, but, where do the probability models M come from? And, how do we choose an 'appropriate' probability model in a given inductive-logical context?" These are good questions. However, it is not clear that these are questions the inductive logician *qua* logician must answer. Here, it interesting to note the analogy between the Pr_M-relativity of inductive logical relations (on the present approach) and the language relativity of deductive logical relations on Carnap's (early) approach to deductive logic. For the early Carnap, deductive logical (or, more generally, analytic) relations only obtain between sentences in a formal language. The deductive logician is not in the business of telling people which languages they should use, since this (presumably, pragmatic) question is "external" to deductive logic. However, once a language has been specified, the deductive relations among sentences in that language are determined objectively and non-contingently, and it is up to the deductive logician to explicate these relations. On the approach to inductive logic just described, the same sort of thing can be said for the inductive logician. It is not the business of the inductive logician to tell people which probability models they should use (presumably, that is an epistemic or pragmatic question), but, once a probability model is specified, the inductive logical relations in that model (*viz*., $\mathfrak{c}$) are determined objectively and non-contingently. On the present approach, the duty of the inductive logician is (simply) to explicate the $\mathfrak{c}$-function — not to decide which probability models should be used in which contexts.

One last analogy might be useful here. When the theory of special relativity came along, some people were afraid that it might introduce an element of "subjectivity" into physics, since (*e.g.*) the velocities of objects were now only determined relative to a frame of reference. There was no physical ether with respect to which objects received their absolute velocities. However, the velocities and other values were determined objectively and non-contingently once the frame of reference was specified, which is the reason Einstein originally intended to call his theory the theory of invariants. Similarly, it seems that there may be no *logical ether* with respect to which pairs of propositions (or sentences) obtain their *a priori* relations of inductive support. But, once a probability model M is specified (and independently of how that model is interpreted), the values of $\mathfrak{c}$ functions defined relative to M are determined objectively and non-contingently (in precisely the sense Carnap had in mind when he used those terms).

4.2 A Fourth Material Desideratum: Relevance

Consider the following argument:

(P) Fred Fox (who is a male) has been taking birth control pills for the past year.

(C) Fred Fox is not pregnant.

Intuitively (*i.e.*, assuming a probability model M which properly incorporates our intuitively salient background knowledge about human biology, *etc.*), $\Pr_M(C \mid P)$ is very high. But, do we want to say that there is a strong evidential relation between P and C? According to proponents of (TRV), we should say just that. This seems wrong, because, intuitively, $\Pr_M(C \mid P) = \Pr_M(C)$. That is, $\Pr_M(C \mid P)$ is high solely because $\Pr_M(C)$ is high, and not because of any evidential relation between P and C. This is the same kind of criticism Skyrms (2000) made against the (NIL) proposal. And, it is just as compelling here. The problem here is that P is irrelevant to C. Plausibly, it seems that if P is going to be counted as providing evidence in favor of C, then P should raise the probability of C. (See Popper (1954, 1992) and Salmon (1975) for more on this point.) This leads to the following fourth material desideratum for $\mathfrak{c}$.

(4) $\mathfrak{c}(C, \{P_1, \ldots, P_n\})$ should be sensitive to the probabilistic relevance of $P_1 \,\&\, \ldots \,\&\, P_n$ to C.

In particular, (4) implies that if P_1 raises the probability of C_1, but P_2 lowers the probability of C_2, then $\mathfrak{c}(C_1, P_1) > \mathfrak{c}(C_2, P_2)$. This rules-out $\Pr(C \mid P_1 \,\&\, \ldots \,\&\, P_n)$ as a candidate for $\mathfrak{c}(C, \{P_1, \ldots, P_n\})$, and it is therefore inconsistent with (TRV). Many non-equivalent probabilistic-relevance measures of support (or confirmation) satisfying (4) have been proposed and defended in the philosophical literature. (See Fitelson (1999, 2001) for surveys and discussion.)

We can combine desiderata (1)–(4) into the following single desideratum (which will be called "PIL" for "Probabilistic Inductive Logic"). This unified desideratum gives constraints on a three-place probabilistic confirmation function $\mathfrak{c}(C, \{P_1, \ldots, P_n\}, M)$, which is the degree to which $\{P_1, \ldots, P_n\}$ inductively supports C, relative to a specified probability model $M = \langle B, \Pr_M \rangle$.

$$\text{(PIL)}\quad \mathfrak{c}(C, \{P_1, \ldots, P_n\}, M) \text{ is } \begin{cases} \text{maximal and } > 0 & \text{if } \{P_1, \ldots, P_n\} \text{ entails } C \\ > 0 & \text{if } \Pr_M(C \mid P_1 \,\&\, \ldots \,\&\, P_n) > \Pr_M(C) \\ 0 & \text{if } \Pr_M(C \mid P_1 \,\&\, \ldots \,\&\, P_n) = \Pr_M(C) \\ < 0 & \text{if } \Pr_M(C \mid P_1 \,\&\, \ldots \,\&\, P_n) < \Pr_M(C) \\ \text{minimal and } < 0 & \text{if } \{P_1, \ldots, P_n\} \text{ entails } {\sim} C \end{cases}$$

To see that any measure satisfying (PIL) will satisfy (1)–(4), note that: (1) the cases of entailment and refutation are at the extremes of $\mathfrak{c}$, with intermediate values of support and counter-support in between the extremes; (2) the constraints in (PIL) can be stated purely probabilistically, and $\mathfrak{c}$'s values must be determined relative to a probability model M, so any measure satisfying it must use probability as a central concept in its definition; (3) the measure $\mathfrak{c}$ is defined relative to a probability model, and so its values are determined objectively and non-contingently by the values in the specified model; and, (4) sensitivity to Pr-relevance is built-in to the desideratum (PIL).

Interestingly, almost all relevance measures proposed in the confirmation theory literature fail to satisfy (PIL). (See Fitelson (2001: §3.2.3) for discussion.) One historical measure that does satisfy (PIL) was independently proposed and defended by Kemeny and

Oppenheim (1952) as the correct measure of confirmation (in opposition to Carnap's (TRV) $\mathfrak{c}$–measures), within a Carnapian framework for logical probability.

$$\mathfrak{c}(C, \{P_1, \ldots, P_n\}, M) = \frac{Pr_M(P_1 \,\&\, \ldots \,\&\, P_n \mid C) - Pr_M(P_1 \,\&\, \ldots \,\&\, P_n \mid \sim C)}{Pr_M(P_1 \,\&\, \ldots \,\&\, P_n \mid C) + Pr_M(P_1 \,\&\, \ldots \,\&\, P_n \mid \sim C)}$$

Indeed, of all the historically proposed (probabilistic) measures of degree of confirmation (and there have been dozens), the above measure is the only measure (up to ordinal equivalence) which satisfies all four of our material desiderata (there are other measures in the literature that differ conventionally from, but are ordinally equivalent to, the above measure, e.g., the log-likelihood ratio — see Fitelson (2001), Good (1985), Heckerman (1988), Kemeny and Oppenheim (1952), and Schum (1994) for various other virtues of measures in this family). As such, our four simple desiderata are sufficient to (nearly uniquely) determine the desired explicandum $\mathfrak{c}$ — the degree of inductive strength of an argument.

5 Historical Epilogue on the Relevance of Relevance

In the second edition of *Logical Foundations of Probability*, Carnap (1962, new preface) acknowledges that probabilistic relevance is an intuitively compelling desideratum for measures of inductive support. This acknowledgement was in response to the trenchant criticisms of Popper (1954), who was one of the first to urge relevance as a desideratum in this context (see Michalos (1971) for a thorough discussion of this important debate between Popper and Carnap). But, instead of embracing relevance measures like Kemeny and Oppenheim's (1952) measure (and rewriting much of the first edition of *Logical Foundations of Probability*), Carnap simply postulated an ambiguity in the term "confirmation". According to Carnap (1962, new preface), there are two kinds of confirmation: (i) confirmation as firmness and (ii) confirmation as increase in firmness, where the former is properly explicated using just conditional probability (a là TRV) and does not require relevance of the premises to the conclusion, while the latter presupposes that the premises are probabilistically relevant to the conclusion. Strangely, Carnap (1962, new preface) does not even mention Kemeny and Oppenheim's measure (of which he was aware) as a proper measure of "confirmation as increase in firmness". Instead, Carnap suggests for that purpose a relevance measure which does not satisfy desideratum (1), and so is not even a proper generalization of deductive entailment. This puzzling but crucial sequence of events in the history of inductive logic may explain why relevance-based approaches (like that of Kemeny and Oppenheim) have never enjoyed as many proponents as The Received View.

Branden Fitelson
Department of Philosophy, University of California, Berkeley, CA 94720-2390, USA.

References

Bacon, F. *The Novum Organon*. Oxford: The University Press, 1620.

Bayes, T. "An Essay Towards Solving a Problem in the Doctrine of Chances." *Philosophical Transactions of the Royal Society of London* 53 (1764).

Boole, G. *An Investigation of the Laws of Thought, on Which Are Founded the Mathematical Theories of Logic and Probabilities*. London: Walton & Maberly, 1854.

Carnap, R. “A Basic System of Inductive Logic I.” In *Studies in Inductive Logic and Probability, Volume I*, eds. Carnap, R. and Jeffrey, R. 33-165. Berkeley: University of California Press, 1971.

———. “A Basic System of Inductive Logic II.” In *Studies in Inductive Logic and Probability, Volume II*, ed. Jeffrey, R. 7-155. Berkeley: University of California Press, 1980.

———. “Statistical and Inductive Probability” and “Inductive Logic and Science” (leaflet). Brooklyn: The Galois Institute of Mathematics and Art, 1955.

———. *Logical Foundations of Probability*, Second Edition. Chicago: University of Chicago Press, 1962.

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

———. *The Continuum of Inductive Methods*. Chicago: The University of Chicago Press, 1952.

Copi, I., and C. Cohen. *Introduction To Logic*, Eleventh Edition. New York: Prentice Hall, 2001.

Dale, A. *A History of Inverse Probability: From Thomas Bayes To Karl Pearson*, Second Edition. New York: Springer-Verlag, 1999.

Festa, R. *Optimum Inductive Methods*. Dordrecht: Kluwer Academic Publishers, 1993.

Fitelson, B. “The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity.” *Philosophy of Science* 66 (1999): S362-S378.

———. *Studies in Bayesian Confirmation Theory*. PhD. Dissertation, University of Wisconsin–Madison (Philosophy), 2001.

Glaister, S. “Inductive Logic.” In *A Companion to Philosophical Logic*, ed. Jacquette, D. London: Blackwell, 2001.

Good, I. J. “Weight of Evidence: a Brief Survey.” In *Bayesian Statistics, 2*, eds. Bernardo, J., M. DeGroot, D. Lindley, and A. Smith. 249–69. Amsterdam: North-Holland, 1985.

Hájek, A. “Probability: Interpretations Of.” *Stanford Encyclopedia of Philosophy*, (Winter 2002 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/win2002/entries/probability-interpret/.

———. “What Conditional Probabilities Could Not Be.” *Synthese* (2003), forthcoming.

Heckerman, D. “An Axiomatic Framework for Belief Updates.” In *Uncertainty in Artificial Intelligence 2*, eds. Kanal, L., and J. Lemmer. 11–22. New York: Elsevier Science Publishers, 1988.

Hempel, C. “Studies in the Logic of Confirmation I & II.” *Mind 54* (1945): 1–26 & 97–121.

Hintikka, J. “A Two-Dimensional Continuum of Inductive Methods.” In *Aspects of Inductive Logic*, eds. Hintikka, J., and P. Suppes. Amsterdam: North Holland, 1966.

Hintikka, J., and I. Niiniluoto. "An Axiomatic Foundation for the Logic of Inductive Generalization." In *Studies in inductive logic and probability, Volume II*, ed. Jeffrey, R. Berkeley: University of California Press, 1980.

Hume, D. *A Treatise of Human Nature: Being an Attempt to Introduce the Experimental Method of Reasoning into Moral Subjects*, 3 volumes; volumes 1 and 2 (London: John Noon, 1739); volume 3 (London: Thomas Longman, 1740).

———. *An Enquiry Concerning Human Understanding in Essays and Treatises on Several Subjects.* London: A. Millar, 1758.

Hurley, P. *A Concise Introduction to Logic*, Eighth Edition. Australia & Belmont, CA: Wadsworth/Thomson Learning, 2003.

Johnson, W.E. "Probability: The Deductive and Inductive Problems." *Mind* 49 (1932): 409–423.

———. *Logic*, Cambridge: Cambridge University Press, 1921.

Joyce, J. *The Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press, 1999.

Kemeny, J., and P. Oppenheim. "Degrees of Factual Support." *Philosophy of Science* 19 (1952): 307-24.

Keynes, J. *A Treatise on Probability*. London: Macmillan, 1921.

Kneale, W., and M. Kneale. *The Development of Logic*. Oxford: Clarendon Press, 1962.

Kolmogorov, A. *Foundations of the Theory of Probability.* New York: Chelsea, 1950.

Kyburg, H. E. *Probability and Inductive Logic*, London: Macmillan, 1970.

Laplace, P. S. M. d. *Théorie Analytique des Probabilités*. Paris: Ve. Courcier, 1812.

Layman, C.S. *The Power of Logic*, Second Edition. New York: McGraw Hill, 2002.

Maher, P. "Probabilities for Multiple Properties: the Models of Hesse and Carnap and Kemeny." *Erkenntnis* 55 (2001): 183-216.

———. "Probabilities for Two Properties." *Erkenntnis* 52 (2000): 63-91.

———. "Subjective and Objective Confirmation" *Philosophy of Science* 63 (1996): 149–174.

Michalos, A. *The Popper-Carnap Controversy*. The Hague: Martinus Nijhoff, 1971.

Mill, J. A System of Logic, Ratiocinative and Inductive, Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation. London: Parker, 1843.

Paris, J. *The Uncertain Reasoner's Companion: A Mathematical Perspective*. Cambridge: Cambridge University Press, 1994.

Popper, K. "Degree of Confirmation." *The British Journal for the Philosophy of Science* 5 (1954): 143-49.

———. R. *The Logic of Scientific Discovery*. London: Routledge, 1992.

Putnam, H. "'Degree of Confirmation' and Inductive Logic", in P.A. Schilpp (ed.), *The Philosophy of Rudolf Carnap*, La Salle, Ill.: The Open Curt Publishing Co., 1963: 761–784.

Roeper, P., and H. Leblanc. *Probability Theory and Probability Logic*. Toronto: University of Toronto Press, 1999.

Salmon, W. C. "Confirmation and Relevance." In G. Maxwell and R. M. Anderson. Jr., (eds.), *Induction, Probability, and Confirmation*: *Minnesota Studies in the Philosophy of Science, vol. 6*. Minneapolis: University of Minnesota Press, 1975, 3-36.

Schum, D. *The Evidential Foundations of Probabilistic Reasoning*. New York: John Wiley & Sons, 1994.

Skyrms, B. "Carnapian Inductive Logic and Bayesian Statistics." 30 (1996): 321-36.

———. *Choice and Chance*. Australia & Belmont, CA: Wadsworth/Thomson Learning, 2000.

Stigler, S. *The History of Statistics*. Cambridge, MA: Harvard University Press, 1986.

Van Fraassen, B. *Laws and Symmetry*. Oxford: Oxford University Press, 1989.