INDUCTIVE LOGIC

The idea of inductive logic as providing a general, quantitative way of evaluating arguments is a relatively modern one. Aristotle's conception of 'induction' (επαγωγή)—which he contrasted with 'reasoning' (συλλογισμός)—involved moving only from particulars to universals (Kneale and Kneale 1962, 36). This rather narrow way of thinking about inductive reasoning seems to have held sway through the Middle Ages and into the seventeenth century, when Francis Bacon (1620) developed an elaborate account of such reasoning. During the eighteenth and nineteenth centuries, the scope of thinking about induction began to broaden considerably with the description 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 nondemonstrative inference. In particular, the pioneering work in probability theory by Bayes (1764), Laplace (1812), Boole (1854), and many others in the eighteenth and nineteenth centuries laid the groundwork for a much more general framework for inductive reasoning. (Philosophical thinking about the possibility of inductive knowledge was most famously articulated by David Hume 1739–1740 and 1758) (See Problem of Induction).

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 nineteenth and early twentieth centuries. Some of the most eloquent articulations of the basic ideas behind inductive logic in this modern sense appear in John Maynard Keynes's *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 Carnap, Rudolf; Confirmation Theory).

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 ["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 as desiderata 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 nondeductive relations between propositions that are characterized by inductive logic) should be objective and logical.

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

> That c is an objective concept means this: if a certain 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

[deductive and inductive logic] is their independence of the contingency of facts [of nature]. This characteristic justifies the application of the common term 'logic' to both fields. [200]

This entry 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.

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 argument: 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 are 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 the contrast between quantitative and comparative/qualitative notions. For simplicity, the focus here 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 $c(C, \{P_1, \ldots, P_n\})$ 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 $P(\bullet)$ and $P(\bullet|\bullet)$ will denote unconditional and conditional probability functions, respectively. Informally (and roughly), "$P(p)$" can be read "the probability that proposition p is true," and "$P(p|q)$" 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.

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 ("if and only if") the material conditional $(P_1 \wedge, \ldots, \wedge P_n) \rightarrow C$ 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 $(P_1 \wedge \ldots \wedge P_n) \rightarrow C$ is (logically?) probably true. More formally, one can express this naive inductive logic (NIL) proposal as follows:

$$c(C, \{P_1, \ldots, P_n\}) \text{ is high iff}$$
$$P((P_1 \wedge, \ldots, \wedge P_n) \rightarrow C) \text{ is high.}$$

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 $(P_1 \wedge, \ldots, \wedge P_n) \rightarrow C$ is probable but the argument from $\{P_1, \ldots, P_n\}$ to C is not a strong one. Skyrms (21) gives the following example:

(**P**) There is a man in Cleveland who is 1,999 years and 11 months old and in good health.
(**C**) No man will live to be 2,000 years old.

Skyrms argues that $P(\mathbf{P} \rightarrow \mathbf{C})$ is high, simply because $P(\mathbf{C})$ 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 countersupport (**C**). Thus, $P((P_1 \wedge \ldots \wedge P_n) \rightarrow C)$ being high is not sufficient for $c(C, \{P_1, \ldots, P_n\})$ being high. Note also that $P((P_1 \wedge \ldots \wedge P_n) \rightarrow C)$ cannot serve as $c(C, \{P_1, \ldots, P_n\})$, since it violates desideratum 1. If $\{P_1, \ldots, P_n\}$ refutes C, then $Pr((P_1 \wedge \ldots \wedge P_n) \rightarrow C) = Pr(\neg(P_1 \wedge \ldots \wedge P_n))$, which is not minimal, since the conjunction of the premises of an argument need not have probability one.

Skyrms suggests that the mistake that NIL makes is one of conflating the probability of the material conditional $Pr((P_1 \wedge \ldots \wedge P_n) \rightarrow C)$ with the conditional probability of C, given $P_1 \wedge \ldots \wedge P_n$, that is, $P(C|P_1 \wedge \ldots \wedge P_n)$.. According to Skyrms, it is the latter that should be used as a definition of $\mathsf{c}(C, \{P_1, \ldots, P_n\})$. The reason for this preference is that $P((P_1 \wedge \ldots \wedge P_n) \rightarrow C)$ fails to capture the evidential relation between the premises and conclusion, since $P((P_1 \wedge \ldots \wedge P_n) \rightarrow C)$ can be high solely in virtue of the unconditional probability of (C) being high or solely in virtue of the unconditional probability of $P_1 \wedge \ldots \wedge P_n$ being low. As Skyrms (20) stresses, $\mathsf{c}(C, \{P_1, \ldots, P_n\})$ 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 (RV) about inductive logic:

$$\mathsf{c}(C, \{P_1, \ldots, P_n\}) = Pr(C|P_1 \wedge \ldots \wedge P_n).$$

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

Probability: Its Interpretation and Role in Traditional Inductive Logic

The Mathematical Theory of Probability

For present purposes, assume that a probability function $P(\bullet)$ is a finitely additive measure function over a Boolean algebra of propositions (or sentences in some formal language). That is, assume that $P(\bullet)$ 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. $P(X) \geq 0$.
ii. If X is a (logically) necessary truth, then $P(X) = 1$.
iii. If X and Y are mutually exclusive, then $P(X \vee Y) = Pr(X) + Pr(Y)$.

Following Kolmogorov, define conditional probability $P(\bullet|\bullet)$ in terms of unconditional probability $P(\bullet)$, as follows:

$$Pr(X|Y) = Pr(X \wedge Y)/Pr(Y),$$
provided that $Pr(Y) \neq 0$.

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

These axioms (and the definition of conditional probability) 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 controversial than the former (but 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 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.

Epistemic Interpretations of Probability

In epistemic interpretations of probability, $P_M(H)$ is (roughly) the degree of belief that 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. $P_M(H|E)$ 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, chap. 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) 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 [$P(a|h) = x$]. (4)

It is not obvious that the RV can satisfy desideratum 3—that c be logical and objective—if the probability function P that is used to explicate c in the RV 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 can be determined a priori (on the basis of a and h alone). As Keynes (1921) explains, his probability function $P(a|h)$ is not subjective, since "once the facts are given which determine our knowledge [background and h], what is probable or improbable [*viz*., a] in these circumstances has been fixed objectively, and is independent of our opinion" (4). But he later suggests that the function is contingent on what the agent's background knowledge K is, in the sense that $P(a|h)$ can vary "depending upon the knowledge to which it is related."

Carnap (1950, §45B) is keenly aware of this problem. He suggests that Keynes should have characterized $P(a|h)$ 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's 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's 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 one must settle for qualitative or comparative judgments.) The problem here is that "$P(X|K)$" ("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 $P_M(X)$, 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, chap. 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 discussing the role of the principle of indifference, logical interpretations of probability require a brief discussion.

Logical Interpretations of Probability

Philosophers who accepted the RV and 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. In such interpretations, conditional probabilities $P(X \mid Y)$ are themselves understood as quantitative generalizations of a logical entailment (or deducibility) relation between propositions Y and X. The motivation for this should be clear—it seems like the most direct way to guarantee that an RV-type theory of inductive logic will satisfy desideratum 3. If $P(\bullet|\bullet)$ is itself logical, then c(•,•), which is defined by the RV as $P(\bullet|\bullet)$, should also be logical, and the satisfaction of desideratum 3 (as well as the other two) seems automatic. Below it will become clear that RV + 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. It is interesting to note that Carnap's (1950 and 1952) systems are almost identical to those described 20–30 years earlier by W. E. Johnson (1921 and 1932) (Paris 1994; Kyburg 1970, Ch. 5). His later work (Carnap 1971 and 1980) became increasingly complicated, involving two-dimensional continua, and was less tightly coupled with the syntax of L (Maher 2000 and 2001; Skyrms 1996 discusses some recent applications of Carnapian techniques to Bayesian statistical models involving continuous random variables; Glaister 2001 and Festa 1993 provide broad surveys of Carnapian theories of logical probability and inductive logic).

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 $P(\bullet)$ over the sentences of L. Finally, following the standard Kolmogorovian approach, construct a conditional probability function $P(\bullet|\bullet)$ over pairs of sentences of L, using the ratio definition of conditional probability given 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 sixteen possible states of the world that can be described. These sixteen maximally specific descriptions are called the *state descriptions* of L, and they are as follows:

$Fa \wedge Ga \wedge Fb \wedge Gb$	$\neg Fa \wedge Ga \wedge Fb \wedge Gb$
$Fa \wedge Ga \wedge Fb \wedge \neg Gb$	$\neg Fa \wedge Ga \wedge Fb \wedge \neg Gb$
$Fa \wedge Ga \wedge \neg Fb \wedge Gb$	$\neg Fa \wedge Ga \wedge \neg Fb \wedge Gb$
$Fa \wedge Ga \wedge \neg Fb \wedge \neg Gb$	$\neg Fa \wedge Ga \wedge \neg Fb \wedge \neg Gb$
$Fa \wedge \neg Ga \wedge Fb \wedge Gb$	$\neg Fa \wedge \neg Ga \wedge Fb \wedge Gb$
$Fa \wedge \neg Ga \wedge Fb \wedge \neg Gb$	$\neg Fa \wedge \neg Ga \wedge Fb \wedge \neg Gb$
$Fa \wedge \neg Ga \wedge \neg Fb \wedge Gb$	$\neg Fa \wedge \neg Ga \wedge \neg Fb \wedge Gb$
$Fa \wedge \neg Ga \wedge \neg Fb \wedge \neg Gb$	$\neg Fa \wedge \neg Ga \wedge \neg Fb \wedge \neg Gb$

Two state descriptions S_1 and S_2 are said to be *permutations* of each other if S_1 can be obtained from S_2 by some permutation of the individual constants. For instance, $Fa \wedge \neg Ga \wedge \neg Fb \wedge Gb$ can be obtained from $\neg Fa \wedge Ga \wedge Fb \wedge \neg Gb$ by permuting a and b. Thus, $Fa \wedge \neg Ga \wedge \neg Fb \wedge Gb$ and $\neg Fa \wedge Ga \wedge Fb \wedge \neg Gb$ 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 the toy language L, there are the following ten structure descriptions:

$Fa \wedge Ga \wedge Fb \wedge Gb$	$(Fa \wedge \neg Ga \wedge \neg Fb \wedge Gb) \vee (\neg Fa \wedge Ga \wedge Fb \wedge \neg Gb)$
$(Fa \wedge Ga \wedge Fb \wedge \neg Gb) \vee (Fa \wedge \neg Ga \wedge Fb \wedge Gb)$	$(Fa \wedge \neg Ga \wedge \neg Fb \wedge \neg Gb) \vee (\neg Fa \wedge \neg Ga \wedge Fb \wedge \neg Gb)$
$(Fa \wedge Ga \wedge \neg Fb \wedge Gb) \vee (\neg Fa \wedge Ga \wedge Fb \wedge Gb)$	$\neg Fa \wedge Ga \wedge \neg Fb \wedge Gb$
$(Fa \wedge Ga \wedge \neg Fb \wedge \neg Gb) \vee (\neg Fa \wedge \neg Ga \wedge Fb \wedge Gb)$	$(\neg Fa \wedge Ga \wedge \neg Fb \wedge \neg Gb) \vee (\neg Fa \wedge \neg Ga \wedge \neg Fb \wedge Gb)$
$Fa \wedge \neg Ga \wedge Fb \wedge \neg Gb$	$\neg Fa \wedge \neg Ga \wedge \neg Fb \wedge \neg Gb$

Now assign nonnegative real numbers to the state descriptions, so that these sixteen numbers sum to 1. Any such assignment will constitute an unconditional probability function $P(\bullet)$ over the state descriptions of L. To extend $P(\bullet)$ 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. 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 $P(\bullet)$ over L. For instance, since $Fa \wedge Ga \wedge \neg Gb$ is equivalent to the disjunction $(Fa \wedge Ga \wedge Fb \wedge \neg Gb) \vee (Fa \wedge Ga \wedge \neg Fb \wedge \neg Gb)$, one will have:

$$\begin{aligned} &Pr(Fa \wedge Ga \wedge \neg Gb) \\ &\quad = Pr((Fa \wedge Ga \wedge Fb \wedge \neg Gb) \vee (Fa \wedge Ga \wedge \neg Fb \wedge \neg Gb)) \\ &\quad = Pr(Fa \wedge Ga \wedge Fb \wedge \neg Gb) + Pr(Fa \wedge Ga \wedge \neg Fb \wedge \neg Gb). \end{aligned}$$

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

$$P(X|Y) = P(X \wedge Y)/Pr(Y), \text{provided that } P(Y) \neq 0.$$

Thus, once the unconditional probability function $P(\bullet)$ is specified for the state descriptions of a language L, all probabilities both conditional and unconditional are thereby determined over L. And, this gives one 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 $m^\dagger$, which assumes that each of the state descriptions is equiprobable a priori: If there are N state descriptions in L, then $m^\dagger$ assigns $\frac{1}{N}$ 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), $m^\dagger$ has the consequence that the resulting probabilities cannot reflect learning from experience. Consider the following simple example. Assume that one adopts a logical probability function $P(\bullet)$ based on $m^\dagger$ as one's own a priori degree of belief (or credence) function. Then, one learns (by conditionalizing) that an object a is F, that is, Fa. Intuitively, one's conditional degree of credence $P(Fb|Fa)$ that a distinct object b also is F, given that a is F, should not always be the same as one's a priori degree of credence that b is F. That is, the fact that one has observed another F object should (at least in some cases) make it more probable (a posteriori) that b will also be F (i.e., more probable than Fb was a priori). More generally, if one observes that a large number of objects have been F, this should raise the probability that the next object one observes will also be F. Unfortunately, no a priori probability function based on $m^\dagger$ is consistent with learning from experience in either sense. To see this, consider the simple case $Pr(Fb|Fa)$:

$$\begin{aligned} P(Fb|Fa) &= m^\dagger(Fb \wedge Fa)/m^\dagger(Fa) \\ &= \frac{1}{2} = m^\dagger(Fb) = Pr(Fb). \end{aligned}$$

So, if one assumes an a priori probability function based on $m^\dagger$, 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 (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 $m^\dagger$) to the hypothesis that a distinct object will also be F.

The fact that (on the probability functions generated by the measure $m^\dagger$) 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 $m^\dagger$ (Carnap 1955). As a result, Carnap formulated an alternative measure function m^*, which is defined as follows. First, assign equal probabilities to each structure description (in the toy language above, n $\frac{1}{10}$). 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 $Fa \wedge Ga \wedge \neg Fb \wedge Gb$ gets assigned an a priori probability of $\frac{1}{20}$ ($\frac{1}{2}$ of $\frac{1}{10}$), but the state description $Fa \wedge Ga \wedge Fb \wedge Gb$ receives an a priori probability of $\frac{1}{10}$ ($\frac{1}{1}$ of $\frac{1}{10}$). To further illustrate the differences between $m^\dagger$ and m^*, here are some numerical values in the toy language L:

Measure function $m^\dagger$	Measure function m^*
$m^\dagger(Fa \wedge Ga \wedge \neg Fb \wedge Gb) = \frac{1}{16}$	$m^*(Fa \wedge Ga \wedge Fb \wedge Gb) = \frac{1}{10}$
$m^\dagger((Fa \wedge Ga \wedge \neg Fb \wedge Gb) \vee ((\neg Fa \wedge Ga \wedge Fb \wedge Gb))) = \frac{1}{8}$	$m^*(Fa \wedge Ga \wedge \neg Fb \wedge Gb) = \frac{1}{20}$
$m^\dagger(Fa) = \frac{1}{2}$	$m^*(Fa) = \frac{1}{2}$
$Pr^\dagger(Fa \mid Fb) = \frac{1}{2} = m^*(Fa) = Pr^\dagger(Fa)$	$Pr^*(Fa \mid Fb) = \frac{3}{5} > \frac{1}{2} = m^*(Fa) = Pr^*(Fa)$

Unlike $m^\dagger$, m^* can model learning from experience, since in the simple language

$$P(Fa \mid Fb) = \frac{3}{5} > \frac{1}{2} = Pr(Fa)$$

if the probability function P is defined in terms of the logical measure function m^*. Although m^* does have some advantages over $m^\dagger$, even m^* can give counterintuitive results in more complex languages (Carnap 1952).

Carnap (1952) presents a more complicated framework, which describes a more general class (or "continuum") of conditional probability functions, from which the definitions of $P(\bullet \mid \bullet)$ in terms of m^* and $m^\dagger$ fall out as special cases. This 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 $m^\dagger$ (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 m^*.

But even this λ–continuum has problems. First, none of the Carnapian systems allow universal generalizations to have nonzero probability. This problem was addressed by Hintikka (1966) and Hintikka and Niiniluoto (1980), who provided various alterations of the Carnapian framework that allow for nonzero probabilities of universal generalizations. Moreover, Carnap's early systems did not allow for the probabilistic modeling of analogical effects. That is, in his 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 more 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 (Maher 2000 and 2001). 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 (Putnam 1963; Festa 1993; Maher 2001).

This discussion has glossed over technical details in the development of (Carnapian) logical interpretations or theories of probability since 1950. To recapitulate, what is important for present purposes is that Carnap (along with the other advocates of logical probability) was an RV theorist about inductive logic. He identified the concept c(•,•) of inductive strength (or inductive support) with the concept of conditional probability $P(\bullet|\bullet)$. And he thought (partly because of the problems he saw with epistemic interpretations) that in order for an RV account to satisfy desideratum 3, it needed to presuppose a logical interpretation (or theory) of probability. This led him, initially, to develop various logical measures (e.g., the a priori logical probability functions $m^{\dagger}$ and m^{*}), and then to define conditional logical probability $Pr(\bullet|\bullet)$ 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 $P(\bullet|\bullet)$ to inductive logic. These difficulties mainly had to do with the ability of Carnap's $P(\bullet|\bullet)$ 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 $P(\bullet|\bullet)$. In 1952, he moved to a parameterized definition of $P(\bullet|\bullet)$, which contained an "index of inductive caution" (λ) that was supposed to regulate the speed with which learning from experience is reflected by $P(\bullet|\bullet)$. Later, Carnap (1971, 1980) added γ and η to the definition of $P(\bullet|\bullet)$, as noted above, in an attempt to further generalize the theory and 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 an RV conception of inductive logic.

At this point, it is important to ask, In what sense are Carnap's theories of logical probability (especially his later ones) *logical*? His early theories (based on the measure functions $m^{\dagger}$ and m^{*}) applied something like the principle of indifference to the state and/or structure descriptions of the formal language L in order to determine the logical probabilities $P(\bullet|\bullet)$. 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 one which statements are equiprobable a priori? Carnap (1955) 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. (22)

Carnap seems to be saying that the principle of indifference is to be applied only to possible events that exhibit certain a priori symmetries with respect to some rational agent's background evidence. But this appears no more logical than Keynes's epistemic approach to probability. It seems that the resulting probabilities $P(\bullet|\bullet)$ will not be logical in the sense Carnap desired (at least no more so than Keynes's 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 the principle of indifference is not very useful. Recall that the goal of Carnap's 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 is one to understand what it means for evidence not to "favor any of the possible events" in a way that does not require one to already understand how to measure the degree to which the evidence confirms each of the possible events? Here, Carnap's discussion of the principle of indifference presupposes that degree of confirmation is to be identified with degree of conditional probability. In that reading, "not favoring" just means "conferring equal probability on," and Carnap's unpacking of the principle of indifference reduces directly to a mathematical truth (which, for Carnap, is good enough to render the principle *logical*). If one 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 will be seen below, there are independent reasons to doubt Carnap's presupposition here that degree of confirmation should be identified with degree of conditional probability. Without that assumption, Carnap's principle of indifference is no longer logical (by his own lights), and the problem of the contingency (nonlogicality) 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 (van Fraassen 1989).

Carnap's later theories of $P(\bullet|\bullet)$ introduce even further contingencies, in the form of adjustable parameters, the "proper values" of which do not seem to be determinable a priori (Carnap 1952, 1971, 1980). In particular, consider Carnap's (1952) λ-continuum. The parameter λ is supposed to indicate how sensitive $P(\bullet|\bullet)$ 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 (Carnap 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 noncontingently 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's epistemic probabilities (or Carnap's measure functions m^* and $m^\dagger$) have a chance of being objectively and noncontingently determined.

A pattern is developing. Both Keynes and Carnap give accounts of a priori probability functions $P(\bullet|\bullet)$ that involve certain contingencies and indeterminacies. They each feel pressure (owing to desideratum 3) to eliminate these contingencies when the time comes to use $P(\bullet|\bullet)$ as an explication of 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 to constrain the ultimate choice of model. Carnap is sensitive to the fact that the principle of indifference does not seem logical, but his attempts to render it so (and useful for grounding the choice of an 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 c from the beginning as a three-place relation that depends on premises, conclusion, and a particular probability model?

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 the RV *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 RV.

Rethinking the Received View

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) 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 the RV one should rather say:

The inductive strength of the argument from $\{P_1, \ldots, P_n\}$ to C relative to a probability model $M = \langle B, P_M \rangle$ is $P_M(C|P_1 \wedge \ldots \wedge P_n)$.

Relativizing judgments of inductive support to particular probability models fully and transparently eliminates the contingency and indeterminacy of these judgments. It is clear that the revision of RV above satisfies all three desiderata, since:

1. $P_M(C \mid P_1 \wedge \ldots \wedge P_n)$ is maximal and constant when $\{P_1, \ldots, P_n\}$ entails C, and $Pr_M(C \mid P_1 \wedge \ldots \wedge P_n)$ 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.
3. Once the conditional probability function $P_M(\bullet|\bullet)$ is specified (as it is, a fortiori, once the probability model M has been), its values are determined objectively and in a way that is contingent on only certain mathematical facts about the probability calculus. This is, the resulting c-values are determined mathematically by the specification of a particular probability model M.

One might respond at this point by asking, Where do the probability models M come from? and how does one choose an "appropriate" probability model in a given inductive logical context? These are good questions. However, it is not clear that they must be answered by the inductive logician *qua* logician. Here it is interesting to note the analogy between the P_M-relativity of inductive logical relations (in the present approach) and the language relativity of deductive logical relations in Carnap's (early) approach to deductive logic. For the early Carnap, deductive logical (or, more generally, analytic) relations obtain only 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 noncontingently, and it is up to the deductive logician to explicate these relations. In 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.*, c) are determined objectively and noncontingently. In the present approach, the duty of the inductive logician is (simply) to explicate the 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 the velocities of objects were now determined only 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 noncontingently 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 c-functions defined relative to M are determined objectively and noncontingently (in precisely the sense Carnap had in mind when he used those terms).

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 that properly incorporates one's intuitively salient background knowledge about human biology, etc.), $P_M(\mathbf{C}|\mathbf{P})$ is very high. But does one want to say that there is a strong evidential relation between **P** and **C**? According to proponents of the RV, one should say just that. This seems wrong, because intuitively $P_M(\mathbf{C} \mid \mathbf{P}) = P_M(\mathbf{C})$. That is, $P_M(\mathbf{C}|\mathbf{P})$ is high solely because $P_M(\mathbf{C})$ is high, and not because of any evidential relation between **P** and **C**. This is the same kind of criticism that 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** (Popper 1954 and 1992; Salmon 1975). This leads to the following fourth material desideratum for c:

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

In particular, desideratum 4 implies that if P_1 raises the probability of C_1, but P_2 lowers the probability of C_2, then $c(C_1, P_1) > c(C_2, P_2)$. This rules out $P(C|P_1 \wedge \ldots \wedge P_n)$ as a candidate for $c(C, \{P_1, \ldots, P_n\})$, and it is therefore inconsistent with the RV. Many nonequivalent probabilistic-relevance measures of support (or confirmation) satisfying desideratum 4 have been proposed and defended in the philosophical literature (Fitelson 1999 and 2001).

One can combine desiderata 1–4 into the following single probabilistic inductive logic. This unified desideratum gives constraints on a three-place probabilistic confirmation function $c(C, \{P_1, \ldots, P_n\}, M)$, which is the degree to which $\{P_1, \ldots, P_n\}$ inductively supports C, relative to a specified probability model $M = \langle B, Pr_M \rangle$:

$$c(C, \{P_1, \ldots, P_n\}, M) \text{ is } \begin{cases} \text{maximal and } > 0 & \text{if } \{P_1, \ldots, P_n\} \text{ entails } C \\ > 0 & \text{if } P_M(C|P_1 \wedge \ldots \wedge P_n) > P_M(C) \\ 0 & \text{if } P_M(C|P_1 \wedge \ldots \wedge P_n) = P_M(C) \\ < 0 & \text{if } P_M(C|P_1 \wedge \ldots \wedge P_n) < P_M(C) \\ \text{minimal and } < 0 & \text{if } \{P_1, \ldots, P_n\} \text{ entails } \neg C \end{cases}$$

To see that any measure satisfying probabilistic inductive logic will satisfy desiderata 1–4, note that

- the cases of entailment and refutation are at the extremes of c, with intermediate values of support and countersupport in between the extremes;
- the constraints in probabilistic inductive logic can be stated purely probabilistically, and 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;

- the measure c is defined relative to a probability model, and so its values are determined objectively and noncontingently by the values in the specified model; and
- sensitivity to P-relevance is built into the desideratum (probabilistic inductive logic).

Interestingly, almost all relevance measures proposed in the confirmation theory literature fail to satisfy probabilistic inductive logic (Fitelson 2001, §3.2.3). One historical measure that does satisfy probabilistic inductive logic was independently defended by Kemeny and Oppenheim (1952) as the correct measure of confirmation (in opposition to Carnap's RV c-measures) within a Carnapian framework for logical probability:

$$\mathsf{c}(C, \{P_1, \ldots, P_n\}, M) = \frac{P_M(P_1 \wedge \ldots \wedge P_n|C) - P_M(P_1 \wedge \ldots \wedge P_n|\neg C)}{P_M(P_1 \wedge \ldots \wedge P_n|C) + P_M(P_1 \wedge \ldots \wedge P_n|\neg C)}.$$

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

Historical Epilogue on the Relevance of Relevance

In the second edition of *Logical Foundations of Probability*, Carnap (1962) 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) (and rewriting much of the first edition of *Logical Foundations of Probability*), Carnap (1962) simply postulates an ambiguity in the term "confirmation." He now argues that there are two kinds of confirmation: confirmation as firmness and confirmation as increase in firmness, where the former is properly explicated using just conditional probability (à la the RV) 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 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, he suggests for that purpose a relevance measure that 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 RV.

BRANDEN FITELSON

References

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

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

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

Carnap, R. (1950), *Logical Foundations of Probability*. Chicago: University of Chicago Press.

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

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

——— (1962), *Logical Foundations of Probability*, 2nd ed. Chicago: University of Chicago Press.

——— (1971), "A Basic System of Inductive Logic, I," in R. Carnap and R. Jeffrey (eds.), *Studies in Inductive Logic and Probability*, vol. 1. Berkeley and Los Angeles: University of California Press, 33–165.

——— (1980), "A Basic System of Inductive Logic, II," in R. Jeffrey (ed.), *Studies in Inductive Logic and Probability*, vol. 2. Berkeley and Los Angeles: University of California Press, 7–155.

Copi, I., and C. Cohen (2001), *Introduction To Logic*, 11th ed. New York: Prentice Hall.

Dale, A. (1999), *A History of Inverse Probability: From Thomas Bayes To Karl Pearson*, 2nd ed. New York: Springer-Verlag.

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

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

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

Glaister, S. (2001), "Inductive Logic," in D. Jacquette (ed.), *A Companion to Philosophical Logic*. London: Blackwell.

Good, I. J. (1985), "Weight of Evidence: A Brief Survey," in J. Bernardo, M. DeGroot, D. Lindley, and A. Smith (eds.), *Bayesian Statistics, 2*. Amsterdam: North-Holland, 249–269.
Hájek, A. (2003), "What Conditional Probabilities Could Not Be," *Synthese* 137: 273–323.
Heckerman, D. (1988), "An Axiomatic Framework for Belief Updates," in L. Kanal and J. Lemmer (eds.), *Uncertainty in Artificial Intelligence 2*. New York: Elsevier Science Publishers, 11–22.
Hempel, C. (1945), "Studies in the Logic of Confirmation," parts I and II, *Mind* 54: 1–26 and 97–121.
Hintikka, J. (1966), "A Two-Dimensional Continuum of Inductive Methods," in J. Hintikka and P. Suppes (eds.), *Aspects of Inductive Logic*. Amsterdam: North-Holland.
Hintikka, J., and I. Niiniluoto (1980), "An Axiomatic Foundation for the Logic of Inductive Generalization," in R. Jeffrey, *Studies in Inductive Logic and Probability*, vol. 2. Berkeley and Los Angeles: University of California Press.
Hume, D. (1739–1740), *A Treatise of Human Nature: Being an Attempt to Introduce the Experimental Method of Reasoning into Moral Subjects*, vols. 1–3. London: John Noon (1739), Thomas Longman (1740).
——— (1758), *An Enquiry Concerning Human Understanding in Essays and Treatises on Several Subjects*. London: A. Millar.
Hurley, P. (2003), *A Concise Introduction to Logic*, 8th ed. Melbourne, Australia, and Belmont, CA: Wadsworth/Thomson Learning.
Johnson, W. E. (1921), *Logic*. Cambridge: Cambridge University Press.
——— (1932), "Probability: The Deductive and Inductive Problems," *Mind* 49: 409–423.
Joyce, J. (1999), *The Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press.
Kemeny, J., and P. Oppenheim (1952), "Degrees of Factual Support," *Philosophy of Science* 19: 307–324.
Keynes, J. (1921), *A Treatise on Probability*. London: Macmillan.
Kneale, W., and M. Kneale (1962), *The Development of Logic*. Oxford: Clarendon Press.
Kolmogorov, A. (1950), *Foundations of the Theory of Probability*. New York: Chelsea.
Kyburg, H. E. (1970), *Probability and Inductive Logic*. London: Macmillan.
Laplace, P. S. M. d. (1812), *Théorie Analytique des Probabilités*.. Paris: Ve. Courcier.
Layman, C. S. (2002), *The Power of Logic*, 2nd ed. New York: McGraw-Hill.
Maher, P. (1996), "Subjective and Objective Confirmation," *Philosophy of Science* 63: 149–174.
——— (2000), "Probabilities for Two Properties," *Erkenntnis* 52: 63–91.
——— (2001), "Probabilities for Multiple Properties: The Models of Hesse and Carnap and Kemeny," *Erkenntnis* 55: 183–216.
Michalos, A. (1971), *The Popper–Carnap Controversy*. The Hague: Martinus Nijhoff.
Mill, J. (1843), *A System of Logic, Ratiocinative and Inductive, Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*. London: Parker.
Paris, J. (1994), *The Uncertain Reasoner's Companion: A Mathematical Perspective*. Cambridge: Cambridge University Press, chap. 12.
Popper, K. (1954), "Degree of Confirmation," *British Journal for the Philosophy of Science* 5: 143–149.
——— (1992), *The Logic of Scientific Discovery*. London: Routledge.
Putnam, H. (1963), "'Degree of Confirmation' and Inductive Logic," in P. A. Schilpp (ed.), *The Philosophy of Rudolf Carnap*. La Salle, IL: Open Court Publishing, 761–784.
Roeper, P., and H. Leblanc (1999), *Probability Theory and Probability Logic*. Toronto: University of Toronto Press.
Salmon, W. C. (1975), "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, 3–36.
Schum, D. (1994), *The Evidential Foundations of Probabilistic Reasoning*. New York: John Wiley & Sons.
Skyrms, B. (1996), "Carnapian Inductive Logic and Bayesian Statistics," in *Statistics, Probability and Game Theory: Papers in Honor of David Blackwell. IMS Lecture Notes, Monograph Series* 30: 321–336.
——— (2000), *Choice and Chance*. Melbourne, Australia, and Belmont, CA: Wadsworth/Thomson Learning.
Stigler, S. (1986), *The History of Statistics*. Cambridge, MA: Harvard University Press.
van Fraassen, B. (1989), *Laws and Symmetry*. Oxford: Oxford University Press, chap. 12.

INNATE/ACQUIRED DISTINCTION

Arguments about innateness center on two distinct but overlapping theoretical issues. One concerns the explanation of the origin of ideas in the human mind. This ancient question was famously introduced in Plato's *Meno* and it took center stage in seventeenth- and eighteenth-century debates between rationalists and empiricists. More recently, it has seen a sophisticated revival in arguments about