Epistemology: Five Questions
John L. Pollock

Regents' Professor and Philosophy and Research Professor of Cognitive Science
University of Arizona

1. Why were you initially drawn to epistemology (and what keeps you interested)?

As a high school student, I rediscovered Hume's problem of induction on my own. For a while, I was horrified. I thought, "We cannot know anything!" After a couple of weeks I calmed down and reasoned that there had to be something wrong with my thinking, and that led me quickly to the realization that good reasons need not be deductive, and to the discovery of defeasible reasoning. From there it was a short jump to a more general interest in how rational cognition works.

I am interested in rational cognition in general. Epistemology is one constituent of rational cognition, practical cognition (rational decision making) another. Much of the work on rational cognition begins with the supposition that only ideal agents can be truly rational. Real agents have limited powers of reasoning and limited memory capacity. It is often supposed that such resource-bounded agents can only approximate rationality, and that as philosophers we should confine our attention to ideal agents. If one wishes, one can of course define "rationality" in this way, but this has never been what interested me. We come to philosophy wondering what *we* should believe, what *we* should do, and how *we* should go about deciding these matters. These are questions about ourselves, with all of our cognitive limitations. For example, it is often claimed that ideal agents, with unlimited cognitive powers, should believe all of the logical consequences of their beliefs. But we, as real resource-bounded agents, cannot do that, so that is not something we *should* do. What I want to know is how I, as a real agent, should go about deciding what to believe and what to do. Thus my topic is *real rationality* as opposed to *ideal rationality*. In the realm of practical decision making, I have explored this distinction at great length in my recent book (2006). Here I will focus on its implications for epistemology.

For many years epistemology was derailed by the Gettier problem. That distracted philosophers from what I regard as the more interesting questions of how specific kinds of epistemic cognition work. How does perception apprise us of the state of our surroundings? How do induction and abduction lead to the discovery of general truths? How is it possible to know the mental states of others? These were once among the central problems of epistemology, but most of the energy that would otherwise have been directed at these problems was drained off by the Gettier problem. The Gettier problem is just an interesting puzzle, and solving it is unlikely to throw much light on how rational cognition works.

2. What do you see as being your main contributions to epistemology?

Defeasible Reasoning

My most important contribution to epistemology was to be one of the discoverers of defeasible reasoning and its central role in rational cognition. Today's younger epistemologists find it incredible that anyone ever thought that all good reasoning had to be deductive, even though that was an almost universally shared opinion prior to the 1960's. I first wrote about defeasible reasoning in an appendix to my PhD dissertation, in 1965, and my first paper on the topic was my (1967). A few other philosophers, most notably Stephen Toulmin (1950), Roderick Chisholm (1966) and Nicholas Rescher (1967), were also starting to think about defeasible reasoning at that time. Insofar as my early views were infuenced by other philosophers, the main influence was probably Wittgenstein's remarks about criteria in *Philosophical Investigations* (1953)

Initially, epistemologists writing about defeasible reasoning used the concept as a tool for the analysis of specific kinds of epistemic cognition (and as a tool for attempts to solve the Gettier problem). They did not give too much thought to precisely how defeasible reasoning works from a logical point of view. More than others, I investigated some of the logical details because that was necessary for understanding how the reasoning was used in specific epistemological problems. In 1978, I produced what was probably the world's first formal semantics for defeasible reasoning. However, I did not publish it until (Pollock 1986, 1987) because I doubted that other philosophers would be interested in the question. I did not realize then that there was a developing interest among artificial intelligence researchers in the same question.

Direct Realism and Non-Doxastic Theories

My initial application of defeasible reasoning was to perceptual knowledge, where I argued (1967, 1968, 1971, 1974) for the then-radical thesis that "*x* looks red to me" gives me a defeasible reasoning for believing "*x* is red", where this reason does not derive from anything else more fundamental, and most importantly, not from a definition of (logically necessary and sufficient conditions for) "red".

Historically, most epistemological theories were *doxastic theories*, in the sense that they endorsed the *doxastic assumption*. That is the assumption that the justifiability of a cognizer's belief is a function exclusively of what beliefs he or she holds. Perceptual beliefs — the first beliefs formed on the basis of perception — are by their very nature not obtained by inference from previously held beliefs. But on a doxastic theory, the justification of a belief cannot depend on anything other than the cognizer's beliefs. Thus perceptual beliefs must be *self-justified* in the sense that they are justified (at least defeasibly) by the mere fact that the cognizer holds them.

Historical foundations theories tried to make this plausible by taking perceptual beliefs to be about the cognizer's perceptual experience. The trouble is, perceptual beliefs, as the first beliefs the agent forms on the basis of perception, are not generally about appearances. It is rare to have any beliefs at all about how things look to you. You normally just form beliefs about ordinary physical objects. You look at your dinner table and judge perceptually that your cat is sitting on the table licking the dirty plates. It never occurs to you to form a belief like, "It appears to me that there is a fuzzy brown blob atop an oval surface and a strangely shaped pink object next to it is moving in the vicinity of another smaller oval shape". You do not have the latter belief (although you could form it with a shift of attention), and the belief about the cat cannot be self-justified because the very same belief can be held for non-perceptual reasons or for inadequate reasons. Still, such beliefs can be justified by perception. What is it about my perceptual experience that justifies me in believing, for example, that the cat is licking the plates? It seems clear that the belief is justified by *the fact that* it looks to me that way. In general, there are various states of affairs *P* for which visual experience gives us direct evidence. Let us say that the relevant visual experience is that of *being appeared to as if P*. Then *direct realism* is the following principle:

(DR) For appropriate *P*'s, if *S* believes *P* on the basis of being appeared to as if *P*, *S* is defeasibly justified in doing so.

Direct realism is "direct" in the sense that our beliefs about our physical surroundings are the first beliefs produced by cognition in response to perceptual input, and they are not inferred from lower-level beliefs about the perceptual input itself. But, according to direct realism, these beliefs are not self-justified either. Their justification depends upon having the appropriate perceptual experiences. Thus the doxastic assumption is false.

Direct realism has had occasional supporters in the history of philosophy, perhaps most notably Peter John Olivi in the 13th century and Thomas Reid in the 18th century. But the theory was largely ignored by contemporary epistemologists until I resurrected it (1971, 1974, 1986) on the basis of the preceding argument.

Procedural Justification

Non-doxastic theories of epistemic justification cry out for an explanation of how beliefs can be justified by non-beliefs, i.e., by perceptual states. Investigating this question led me to what I regard as a fundamental insight that profoundly changed my perspective on epistemology. Thirty years ago, a philosopher could write a paper on epistemology, making free and frequent use of the term "epistemic justification", without further explanation and confident that his audience would understand him. That is no longer a reasonable way to proceed. It has become increasingly apparent that there is more than one important concept that might reasonably be called "epistemic justification", and it seems clear that these different concepts have often been confused with one another in the epistemological literature. The literature on the Gettier problem has highlighted one concept, dear to the hearts of many epistemologists, which is something like "what turns true belief into knowledge". The reliabilist literature, for example, might best be read as pertaining to some such concept as this. But there is at least one other important concept that pertains to "the directing of one's own cognition". This is an essentially first person concept. In my book *Contemporary Theories of Knowledge* (Pollock 1986, Pollock and Cruz 2000), I proposed that this concept could be understood as descriptive of our procedural knowledge for "how to cognize". That part of epistemology that is more concerned with the procedural aspects of rationality than with the analysis of "*S* knows that *P*" can, I urged, be viewed as pursuing a competence theory of cognition,

in the same way that theories of grammar in Linguistics are competence theories of language. On this view, the normative language employed in the formulation of both theories of grammar and epistemological theories is a reflection, in part, of the competence/performance distinction that can be drawn in connection with any procedural knowledge. More recently (Pollock 2008), I have come to realize that the normativity also reflects the fact that the competence theory of cognition is actually embedded in the performance theory of human cognition. Human cognition is organized so that when we violate the norms of the competence theory, and realize that we have done so, we are moved to attempt to correct our cognition and bring it into conformance with the norms. This is a purely descriptive fact about human cognition, and amounts to saying that we treat the norms of competent cognition normatively (not just that we *should* treat them that way).

This leads to a shift of perspective. We can think of epistemology as being part of the enterprise of designing a rational cognizer. That is the traditional task of artificial intelligence, and epistemology provides part of the analysis of rationality driving the design. What I will call *procedural epistemology* is directed at how to build the system of cognition. My interest can then be described as one in procedural epistemology. I want to know how the rules of epistemic cognition should be formulated in order for cognition to conform to our pre-analytic judgments of rationality.

The OSCAR Project

Procedural epistemology is a theory about how something works — rational epistemic cognition. It became clear to me quite early that the standard methodology of armchair philosophy is not well suited to getting such theories right. Armchair epistemology can go a long way to getting us started in constructing theories of procedural epistemology, but human beings are quite limited in the complexity of examples they can construct and evaluate from the armchair. Rules of cognition must be applicable to cases of arbitrary complexity, and my experience has revealed many cases in which rules that appear to work in simple cases produce absurd results in some complex cases. Part of the difficulty is that it can be hard to tell what a theory even implies about a sufficiently complex case. A philosopher may think that it has one implication, and deem that correct, when in fact the subtle complexities of the case yield a different and unacceptable implication. I became convinced that the only way to avoid this problem is to mechanize the production of test cases and the application of epistemological theories to them. To do that we must build a system that models proposed accounts of cognition, so that we can see what the system really does in complex cases. The best way to do that is with a computer model, and the resulting system becomes an AI system . Given a precise implementation of a procedural epistemological theory, we can apply it automatically to complex cases and see what it does, and then evaluate whether that is what it rationally should do. (See Pollock (1998) for an extended example dealing with perceptual, temporal, and causal reasoning.) It is noteworthy that such systems almost always misbehave initially in complex cases. We can then investigate why we got the unexpected result, and modify the cognitive system being modeled to try to get it to produce more congenial results.

Thus the OSCAR Project was born in 1985. OSCAR is an AI system that seeks to model various aspects of human rational cognition. OSCAR has been generously supported by the National Science Foundation for a number of years. In its current state, OSCAR is a general architecture for rational cognition based on a system of defeasible reasoning and natural deduction, and implements rules for many specific kinds of cognition including perceptual reasoning, temporal reasoning, causal reasoning, and various kinds of probabilistic reasoning (Pollock 2008c). OSCAR performs both epistemic cognition and rational decision making, the latter driven by systems of decision-theoretic planning that are currently under development (Pollock 2006).

The core of OSCAR is a system of defeasible reasoning. This system has changed in important ways over the years in response to failures revealed by actually running the system and applying it to interesting examples. The early system of my (1986, 1987) proved inadequate for dealing with "self-defeating arguments" — arguments that support defeaters for some of their own steps. Such arguments are produced with surprising frequency in the course of reasoning about matters that would not naturally be expected to produce self-defeating arguments. In my (1994, 1995) I produced a new semantics that seemed to get most of the problematic cases right. More recently, however, I have become convinced that the new semantics is not quite right either, and I proposed a newer semantics (Pollock 2002). I am still working on modification to that semantics, and it has not yet been implemented in OSCAR.

Epistemology and Probability

I have long been interested in the role of probability in epistemology. Most probabilists working in epistemology are Bayesians who think that all of epistemology is reducible to subjective probability and the probability calculus. I have repeatedly argued (e.g., in my 2006a) that the logical structure of probabilities

does not mimic that of epistemic justification. For example, it follows from the probability calculus that all necessary truths have probability 1, but clearly we are not automatically justified in believing necessary truths. For most of them we only become justified as a result of giving an argument, but giving an argument cannot affect their probabilities. They had probability 1 all along — we just did not know that they did. So their having probability 1 did not make us justified in believing them.

Bayesians sometimes reply that they are only interested in ideal agents, unconstrained by the realistic limits on memory and reasoning ability that plague real agents. But my interest is in real agents. I want to know how *I* should solve cognitive problems, not how a mythical ideal agent should do it. So I do not think Bayesians have much to tell us about epistemology.

Still, probabilities are important. Most of our general knowledge of the world is probabilistic. These cannot be subjective probabilities, so I have spent a lot of time trying to make sense of objective and partly objective probabilities (Pollock 1990, 2006). At the time subjective probabilities were becoming popular, objective probabilities were being roundly criticized because their defenders were unable to define them (state necessary and sufficient conditions) in terms of philosophically simpler notions. But that was in the middle of the twentieth century when people still thought that most concepts had definitions and good philosophical analyses consisted of giving definitions. If we learned anything from the 1960's and 1970's, it is that this is a bad theory of philosophical analysis. Few if any philosophically interesting concepts have definitions. You cannot define "person" in terms of behavior, "red" in terms of "looks red", or concepts like "time" or "physical object" in terms of anything. Concepts just do not work that way. The best way to clarify such concepts philosophically is instead by constructing precise theories of how to use them in reasoning. Thus direct realism clarifies concepts like "red" by explaining how to reason defeasibly about the colors of things on the basis of perception.

Although nobody believes any more that concepts are somehow illegitimate or philosophically confused if we cannot define them, objective probability was tarred with that brush and somehow the smear has stuck in many minds. This is a completely illegitimate charge. We should no more expect "probability" to have a definition than "red" or "person". To clarify the concept of objective probability, we need an epistemological theory of how to reason about probabilities, and that has been the goal of my work.

There is a nowadays often overlooked distinction between *generic probabilities*, which relate properties, and *singular probabilities*, which attach to propositions. For example, on the basis of statistical evidence the medical community may estimate that the probability of an adult male of Slavic descent being lactose intolerant is .6. This is not about any particular adult male — it relates the *property* of being an adult male of Slavic descent to the *property* of being lactose intolerant. A doctor may go on to inquire about the probability that Boris, a particular Slavic male, is lactose intolerant. This is a singular probability, about the proposition that Boris is lactose intolerant, and that need not be the same as the generic probability because we may have other relevant information about Boris. Historically, theories of objective probability have tended to be theories of generic probability, and then theories of "direct inference" were proposed for how to infer the values of singular probabilities from knowledge of collections of relevant generic probabilities (Reichenbach 1949, Kyburg 1974, Pollock 1990, Bacchus 1990, Halpern 1990, Bacchus et al 1996).

I regard one of my most important accomplishments to be the theory of nomic probability (Pollock 1990). Nomic probabilities are generic probabilities glossed informally by taking prob(*Fx/Gx*) to be a measure of the proportion of physically possible *G*'s that would be *F*'s. I showed that if we make a small number of seemingly obvious assumptions about the proportion function used in this formulation, all of which are trivial theorems of set theory for proportions among finite sets, we can generate a very rich calculus of nomic probabilities. This is coupled with a single pair of epistemological principles, *the statistical syllogism* and *subproperty defeat*, formulated as follows:

Statistical Syllogism:

If F is projectible with respect to G and $r > 0.5$, then $\ulcorner Gc$ & $\text{prob}(Fx/Gx) \geq r \urcorner$ is a defeasible reason for $\ulcorner Fc \urcorner$, the strength of the reason being a monotonic increasing function of r.

Subproperty Defeat for the Statistical Syllogism:

If H is projectible with respect to G, then $\ulcorner Hc$ & $\text{prob}(Fx/Gx\&Hx) < \text{prob}(Fx/Gx)\urcorner$ is an undercutting defeater for the inference by the statistical syllogism from $\ulcorner Gc$ & $\text{prob}(Fx/Gx) \geq r\urcorner$ to $\ulcorner Fc\urcorner$.

In my (1990) I was able to show that from this parsimonious set of assumptions we can derive a rich theory of probabilistic reasoning including a theory of direct inference and principles of statistical induction.

However, the theory of my (1990) left some important issues unresolved. One of the most important is that although direct inference is occasionally useful, very often we know too much to be able to use it. Suppose the generic probability of a person with Boris' symptoms being lactose intolerant is .6. Suppose we have two seemingly unrelated diagnostic tests for a disease, and Boris tests positive on both tests. We know that the probability of a person with his symptoms having the disease if he tests positive on the first test is .7, and the probability if he tests positive on the second test is .75. What should we conclude about the *joint probability* of his having the disease if he tests positive on both tests? The probability calculus gives us no guidance here. It is consistent with the probability calculus for the joint probability to be anything from 0 to 1. Nor does direct inference help. Direct inference gives us one reason for thinking that the probability of Boris having the disease is .7, and it gives us a different reason for drawing the conflicting conclusion that the probability is .75. It gives us no way to combine the information. Intuitively, it seems that the probability of his having the disease should be higher if he tests positive on both tests. But how can we justify this?

In my (2008a, 2008b), employing the same assumptions about the proportion function ρ, I proved the following very fundamental theorem about both proportions among finite sets and generic probabilities:

Expectable Probabilities Principle:

Let $U,X_1,\ldots,X_n$ be a set of variables ranging over sets, and consider a finite set LC of linear constraints on proportions between Boolean compounds of those variables. Then for any pair of Boolean compounds P,Q of $U,X_1,\ldots,X_n$ there is a real number r between 0 and 1 such that for every $\varepsilon,\delta > 0$, there is an N such that if U is finite and $\#U > N$, then

$$\rho\left(\rho(P,Q) \underset{\delta}{\approx} r \;/\; LC \;\&\; X_1,\ldots,X_n \subseteq U\right) \geq 1-\varepsilon .$$

For any properties $X_1,\ldots,X_n$, if P and Q are the corresponding compound properties, it is defeasibly reasonable to expect that $\text{prob}(P/Q) = r$.

Furthermore, there is an algorithm for computing r. Applying this to the preceding problem, let us define:

$$\Upsilon(r,s\,|\,a) = \frac{rs(1-a)}{a(1-r-s)+rs}$$

I then proved:

Y-Principle:

Given background information U, if B,C include (i.e., nomically imply) U, $\text{prob}(A/B) = r$, $\text{prob}(A/C) = s$, and $\text{prob}(A/U) = a$, then it is defeasibly reasonable to expect that $\text{prob}(A/B \;\&\; C) = \Upsilon(r,s\,|\,a)$.

The Y-principle makes knowledge of generic probabilities useful in ways it was never previously useful. It tells us how to combine different probabilities that would lead to conflicting direct inferences and still arrive at a univocal value. Consider Boris again, who has symptoms suggesting a particular disease, and tests positive on two independent tests for the disease. Suppose again that the probability of a person with those symptoms having the disease is .6. Suppose the probability of such a person having the disease if they test positive on the first test is .7, and the probability of their having the disease if they test positive on the second test is .75. What is the probability of their having the disease if they test positive on both tests? We can infer defeasibly that it is $\Upsilon(.7,.75\,|\,.6) = .875$. We can then apply direct inference to conclude that the probability of

Boris' having the disease is .875. This is a result that we could not have gotten from the probability calculus alone or from direct inference alone. Similar reasoning will have significant practical applications in many realms.

The Y-principle is just the tip of the iceberg. The expectable probabilities principle justifies an immense number of often previously unknown principles for reasoning defeasibly about probability, and promises to make probability practically useful in ways it was never previously useful because no one has known how to make reasonable defeasible estimates for probabilities whose values are uncomputable in the probability calculus. We can now solve analytically many problems that could previously only be solved using Monte Carlo methods.

3. What do you think is the proper role of epistemology in relation to other areas of philosophy and other academic disciplines?

Procedural epistemology is about how rational cognition works, and as I have urged, it makes up a large part of a purely psychological theory of how human cognition works in general. Not all cognition is rational, but you cannot understand cognition in general without understanding the central role rational cognition plays in it and the way in which general cognition tries to move human cognizers in the direction of rationality. As such, procedural epistemology is a subtheory of a psychological theory of human cognition. However, this does not mean that it is best studied by psychologists employing currently available psychological methodology. That methodology is well-suited for studying human epistemological performance, but as I urged above, the theory of rational cognition is a theory of cognitive competence, not a theory of cognitive performance (just as linguistic theories of grammar purport to be competence theories, while psycholinguistic theories are performance theories). At this time psychologists do not have good tools for studying competence theories experimentally. You cannot do it just by seeing how subjects handle epistemic problems at a surface level, because what is sought is how they would handle the problems given adequate time to fully deliberate, often with considerable help from their peers who can often point out errors in their reasoning that the subjects will ultimately recognize as errors. Take the Wason selection task (Wason 1966). Subjects are presented with four cards. One card shows a red face, one a black face, one a nine, and one a ten. The hypothesis is that all cards with one red face have an even number on the other side. Subjects are asked which cards they must turn over to determine whether that is true. Overwhelmingly, most subjects judge that they must turn over the red card and the ten. But the right answer is the red card and the nine. To confirm the hypothesis, they must verify that the nine is not red. Knowing whether the ten is red is irrelevant. Although people almost invariably get this wrong the first time, they also have no trouble seeing what the right answer is when it is explained to them. So a performance theory will describe them as turning over the red card and the ten, but a competence theory will have them turn over the red card and the nine. It is the latter kind of theory that procedural epistemology seeks, but an avenue to it through contemporary experimental psychology is obscure. This seems to be an error that is pervasive in much of the currently popular philosophical work on experimental epistemology.

If procedural epistemology is at heart a psychological theory about important aspects of human cognition, but it cannot be studied using standardly available psychological methodology, how can it be studied? Armchair epistemologists propound thought experiments, use their "philosophical intuitions" to judge whether certain reasoning would be rationally correct, and then search for general theories to accommodate their judgments. I agree that that is the right way to start, providing you are careful to rely only on very firm intuitions. The explanation for why this works is that the human cognitive architecture makes it possible for us to cognize irrationally (see my 2008 for a lengthy discussion of this), but also provides an important feedback mechanism for recognizing divergences from our built-in epistemic norms and a built-in conative disposition to correct our cognitive performance when we find ourselves making cognitive mistakes. The output of this feedback mechanism is our philosophical intuitions. Similar feedback mechanisms are ubiquitous in procedural knowledge. When linguists appeal to their linguistic intuitions to judge that certain utterances are ungrammatical, they are employing a similar feedback mechanism (Pollock 1986, 2008, Pollock and Cruz 2000). The philosophical intuitions thus produced are about single cases. We have no direct intuitive access to our built-in epistemic norms themselves. So our task becomes a standard problem in scientific reasoning to discover a general account of rational cognition that accommodates and explains our singular judgments.

Armchair epistemologists stop here. This is where I diverge from them in an important way. The complexity of the examples we can consider from the armchair is very limited . To test armchair-generated theories on real-world problems, we must mechanize the process by building a model that implements the

theory. The result is an AI system, and once we have the implementation we can apply it automatically to complex problems and see how the resulting system reasons its way through the problems. 25 years of experience doing this in the OSCAR Project reveals two facts that philosophers find startling but scientists who build models in other disciplines will find commonplace. First, our armchair theories are never formulated precisely enough to be implemented. They tend to have huge holes in them that we never noticed until we try to implement them, but this has the consequence that the theories do not actually have any implications about interesting cases. So the implementation forces us to be much more careful in the formulation of our theories. The second observation is that once made precise, our initial theories never do what we expect them to do. In other words, they are wrong. No computer programmer will find this surprising. Basically, what we are trying to do is write a computer program for how to cognize, and nobody ever writes complex bug-free computer programs on a single pass. Once the simple programming bugs themselves are fixed, what remain are bugs in theory, and there will be lots of them that can only be found by running the resulting system and seeing what it does. Thus armchair epistemology is an appropriate starting point for procedural epistemology, but it can only be the starting point. To complete the theory in a credible way, a partial merger of philosophy and artificial intelligence is essential.

4. What do you consider to be the most neglected topics and/or contributions in contemporary epistemology?

Before the Gettier problem took the epistemological world by storm, much work was done on topics like perceptual knowledge, temporal reasoning, causal reasoning, knowledge of other minds, induction, inference to the best explanation, etc. These are topics that make up the core of procedural epistemology. I think it is time to deem the Gettier problem as what it is — an interesting puzzle, but not the central question of a mature discipline — and turn our attention back to the hard and more far-reaching issues of procedural epistemology.

5. What do you think the future of epistemology will (or should) hold?

I believe the future will hold a partial merger of epistemology with both psychology and artificial intelligence. That will never replace an appeal to philosophical intuitions. They will probably always provide the starting point for the construction of epistemological theories. But it is important to rely only upon very clear intuitions. When Nelson Goodman (1955) provided his "grue" counterexample to the Nicod principle, no one could doubt that he was right. The philosophical intuition was absolutely clear. But most philosophical intuitions are less clear. I welcome the development of psychological tools that will help clarify vague philosophical intuitions. This must be accompanied by a better understanding of how to investigate competence theories experimentally. When we have that, experimental epistemology will come of age. But even then, this provides only the single case data for theory construction. Constructing general epistemological theories to accommodate the single case data is a matter of scientific theory formation, and the resulting theories will be too complex to be tested just by looking at simple cases. They must be implemented and tested on real-world problems, and that requires a partial merger with artificial intelligence.

Bibliography

Bacchus, Fahiem
1990 "Representing and Reasoning with Probabilistic Knowledge", MIT Press.
Bacchus, Fahiem, Adam J. Grove, Joseph Y. Halpern, Daphne Koller
1996 "From statistical knowledge bases to degrees of belief", *Artificial Intelligence* **87**, 75-143.
Chisholm, Roderick
1966 *Theory of Knowledge*. Englewood Cliffs, NJ: Prentice-Hall.
Halpern, J. Y.
1990 "An analysis of first-order logics of probability", *Artificial Intelligence* **46**, 311-350.
Kyburg, Henry, Jr.
1974 *The Logical Foundations of Statistical Inference*. Dordrecht: Reidel.
Pollock, John

1967 "Criteria and our Knowledge of the Material World", *The Philosophical Review*, **76**, 28-60.
1968 What is an epistemological problem? *American Philosophical Quarterly* 5: 183-90.
1971 "Perceptual Knowledge", *Philosophical Review*, **80**, 287-319.
1974 *Knowledge and Justification*, Princeton University Press.
1986 *Contemporary Theories of Knowledge*, Rowman and Littlefield.
Pollock, John
1987 "Defeasible reasoning", *Cognitive Science* **11**, 481-518.
1987a "Epistemic Norms", *Synthese* **71**, 61-96.
1990 Nomic Probability and the Foundations of Induction, Oxford University Press.
1994 "Justification and defeat", *Artificial Intelligence* 67: 377-408.
1995 *Cognitive Carpentry*, MIT Press.
1998 "Perceiving and reasoning about a changing world", *Computational Intelligence*. **14**, 498-562.
2002 "Defeasible reasoning with variable degrees of justification", *Artificial Intelligence* **133**, 233-282.
2006 *Thinking about Acting: Logical Foundations for Rational Decision Making*, New York: Oxford University Press.
2008 "Irrationality and cognition", in *Epistemology: New Philosophical Essays*, ed. Quentin Smith, New York: Oxford University Press.
2008a "Probable probabilities", http://oscarhome.soc-sci.arizona.edu/ftp/PAPERS/Probable Probabilities.pdf.
2008b "Reasoning defeasibly about probabilities", in Michael O'Rourke and Joseph Cambell (eds.), *Knowledge and Skepticism*, Cambridge, MA: MIT Press.
2008c "OSCAR: A Cognitive Architecture for Intelligent Agents", in *A Roadmap for Human-Level Intelligence*, eds. W. Duch and J. G. Taylor, Springer.Pollock, John
Pollock, John, and Joseph Cruz
1999 *Contemporary Theories of Knowledge*, 2nd edition, Lanham, Maryland: Rowman and Littlefield.
Reichenbach, Hans
1949 *A Theory of Probability*. Berkeley: University of California Press. (Original German edition 1935)
Rescher, Nicholas
1976 *Plausible Reasoning*. Amsterdam: Van Gorcum.
Toulmin, Stephen
1950 *An Examination of the Place of Reason in Ethics,* Chicago: University of Chicago Press.
Wason, P.
1966 "Reasoning". In B. Foss (ed.), *New Horizons in Psychology*. Harmondsworth, England: Penguin.
Wittgenstein, Ludwig
1953 *Philosophical Investigations.* Translated by G.E.M. Anscombe. New York: Macmillan.