Order Number 9227330

Probabilistic reasoning and teaching critical thinking

McCarthy, Christine L., Ph.D.

The Ohio State University, 1992

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

PROBABILISTIC REASONING

AND

TEACHING CRITICAL THINKING

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Christine L. McCarthy, B.S., M.A., M.A.

* * * * *

The Ohio State University

1992

Dissertation Committee:

G.M. Reagan

M. Leach

V. Rentel

Approved by

Advisor
College of Education

To My Mother,
Kathleen McCarthy

VITA

August 2, 1953................Born, New Britain, Connecticut

1975..........................B.S., Quinnipiac College, Hamden, Connecticut

1980..........................M.A., University of Colorado, Boulder, Colorado

PUBLICATIONS

"Rights, Community and Civic Education", Proceedings of the Ohio Valley Philosophy of Education Society, 1991.

"Applied Ethics: What Is It? And How Do We Teach It?", Proceedings of the Ohio Valley Philosophy of Education Society, 1990.

"The Role of Moral Intuition in Applied Ethics: A Consideration of the Positions of C.S. Peirce, and Strike and Soltis", Proceedings of the Ohio Valley Philosophy of Education Society, 1989.

FIELDS OF STUDY

Major Field: Education

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

Introduction

In this work I would like to build a case for the explicit teaching of the fundamentals of probabilistic reasoning in the context of a general education curriculum, at the upper high school and/or college level, as one part of a "critical thinking" program. There are several points that suggest that the inclusion of probabilistic reasoning in such a program would be useful.

The Ubiquity of Probabilistic Reasoning

The first reason that attention should be given to probabilistic reasoning is that this sort of reasoning is nearly ubiquitous. Probabilistic reasoning occurs in certain obvious contexts, i.e., in those scientific fields that rely on the mathematical calculation of probabilities, particularly in the social sciences. It occurs in the calculations of risk undertaken by insurance companies and by social policy makers concerned with, for instance, the siting of nuclear power plants or hazardous waste disposal sites.[1] [2]

"Risks" are an inescapable part of everyday life, and the informal evaluation of such risks, requiring the estimation of the probabilities associated with the situations one encounters, is thus a task facing every individual. As Levinson puts it, "We live in a world of chance, and if we wish to live intelligently we must know how to take chances intelligently. To do so we must know and understand the laws of chance..."[3]

Probabilistic reasoning is a significant part of one's "everyday" informal reasoning, problem-solving and decision-making. In the context of everyday problems, however, one usually does not undertake an overt, mathematical process of calculation, and in the absence of such calculation, the probabilistic element that is implicit in every-day reasoning can easily be overlooked.

Consider, for example, the following short list of common problem situations, each of which would require one to make a judgment of probability.

<u>Examples of Everyday Probability Judgments</u>

Judgments of probability must be made when one:

1. evaluates the relative risk of traveling by air, car and train;

2. considers the time, and cost, of cab-fare to the airport, and decides to take the cross-town short-cut, rather than the longer free-way route;

3. estimates the chance of being caught in a serious snowstorm when driving to Connecticut in January;

4. worries about not finding a job, considering the number of openings and the number of applicants;

5. pays for a "state of the art" crash helmet for a twelve year-old, who's received roller-blades for his birthday;

6. decides to save some money by having the wood stove chimney cleaned biennially, instead of annually;

7. considers whether the woodstove fire will eventually dry out and then ignite the stockings you've "hung by the chimney with care";

8. decides not to test a child's cholesterol level, knowing the parents' levels are normal; and then makes the same decision for the adopted child, whose family history is unknown;

9. decides to take seriously the risk of developing heart disease, given that one has high cholesterol;

10. expects to pass the midterm exam, without studying for it;

11. expects to pass the comprehensive final exam, without studying for it, after having failed the midterm;

12. decides not to annoy the fire department when the smoke alarm sounds at midnight in the empty apartment next door, because one doesn't smell any smoke;

13. buys <u>two</u> platters of shrimp for the party, instead of one;

14. decides to search for a close-to-campus parking space, fifteen minutes before class;

15. notices that, up ahead, a child on a bicycle has darted suddenly into traffic from a hidden driveway, and anticipates another;

16. predicts that the widening crack in the ceiling right above the computer will probably not leak;

17. accepts that the "positive" blood test is a sure sign of the presence of disease;

18. assumes that the women in one's philosophy class will provide the "woman's viewpoint", while the men will only be able to "speak for the men";

19. continues to speak of the "woman's viewpoint", despite noting that each woman in class argues for a different position;

20. decides not to go to any more plays at the community theatre, since the play just seen was terrible;

21. decides not to see a movie that seemed interesting, after reading a single negative review of it;

Each of these commonplace decisions or judgments requires the making of a judgment of probability. One needs, in general, to estimate the frequency of occurrence of an event or situation. In some of the examples, one attempts to estimate the probability that a particular, singular event will occur (whether the crack in the ceiling really will let go). In others, one needs to assess the degree of variability present within a reference class (e.g., "women", in #18), and use that assessment to predict the behavior or status of a "sample" from that class. In others, one needs to re-evaluate one's confidence in the truth of a hypothesis, on the basis of new evidence (19).

Fallacious reasoning with respect to even simple problems involving chance and probability is common. Levinson tells the tale, for instance, of a thoughtful midshipman, who during battle "...was prudent enough to stick his head through the first hole in the side of the ship made by an enemy cannonball, as...by a calculation...the odds were 32.647 and some decimals to boot that another ball would not come in at the same hole".[4] Paradoxical though it may seem, the probability that a ball will strike any given spot on the ship is remains exactly what it was originally--the probability that a second ball will strike that spot is the same as the probability was for the first ball (discounting the possibility of a patterned sweep of the cannons). It is only the <u>prior</u> probability that two balls will strike that same spot that is very small. The probability that two independent events will both occur is the product of the two probabilities, i.e., Pr(a & b) = Pr(a) * Pr(b), but, the probability of the second event, <u>given</u> that the first has already occurred, is the same as the probability of the second event alone. I.e., Pr(b/a) = Pr(b). This, in fact, is precisely what is meant by saying that the two events are "independent".

In the course of this work many further examples of probabilistic reasoning will be given, and the issues, difficulties, biases and outright errors that commonly arise in making probabilistic judgments will be explored in much greater detail; for now, this brief sketch should suffice to indicate the general realm to be explored.

I shall try to show that probabilistic reasoning is the sort of reasoning that is commonly required to resolve everyday problem situations.

The Aim of Critical Thinking Programs

The second point indicating the merit of probabilistic reasoning in a critical thinking program is that, in adopting a program to teach "critical thinking", what we would like to see improved is just such "everyday reasoning"[5]. On this point there seems to be widespread agreement (McPeck, Sternberg, Paul, Siegel). But, despite the general agreement about the purpose of critical thinking programs, there exists a considerable disagreement about just how one should set out to achieve the desired improvement of everyday reasoning.

The disagreement about the proper content of a critical thinking course is linked to the conceptual disagreement about the meaning of 'critical thinking'. There are currently a number of different views as to the most appropriate interpretation of the term 'critical thinking'. And, accordingly, there exists considerable variation in the emphasis of the programs that set out to teach critical thinking. Some approaches construe "critical thinking" as logical thinking, and hence stress the teaching of the rules of logic, either formal, informal, or both, and emphasize the central importance of the "giving of reasons" for the decisions one makes, as well as the formulation and evaluation of arguments.

Other programs are based on the notion that 'critical thinking' is best understood as the close examination of one's own thinking process, i.e., it is "thinking about thinking", and hence emphasize the development of the student's "meta-cognition".

Still other programs start with the assumption that "critical thinking" is "skillful thinking", and set out to explicitly teach what are taken to be the various component skills of thinking, e.g., analyzing, making comparisons, making metaphors, etc..

And, in sharp contrast to all the other approaches, some theorists, principally McPeck,[6] maintain that "critical thinking", properly understood, is that thinking which conforms to the standards of reasoning peculiar to particular, discrete disciplines. Given this interpretation, "critical thinking" cannot be separately taught; it can only be taught within the context of an education in the various disciplines.

Nevertheless, despite the general agreement on the broad goals of a critical thinking program, and despite the diversity of approaches to teaching critical thinking, and despite the importance of the notion of probability to everyday reasoning, most critical thinking programs currently give little or no explicit attention to teaching the norms of probabilistic reasoning.

But, although approaches to the teaching of critical thinking are quite diverse, a brief survey of the approaches to teaching critical thinking reveals no approach that at present gives explicit and extended attention to teaching the norms of probabilistic reasoning.[7]

This is not to assert, of course, that no programs give any attention to this sort of reasoning. There may well exist programs that do undertake to fully incorporate probabilistic reasoning in the curriculum. The thesis being advanced here is that such programs, if any, are meeting a significant educational need.

Pratte, in his analysis of the concept of "need", notes that the concept incorporates two necessary aspects, the empirical and the normative.[8] Empirically, whenever one has a "need", one must be lacking something, i.e., something is not present or does not exist. But, the mere fact that one lacks "x" is not sufficient for the claim that one *needs* "x". One might, for instance, lack AIDS, but would not on that account be said to "need" AIDS. If the lack of x is claimed to constitute a "need" for x, one must also make the normative claim, that x is desirable.

It is this second, normative aspect of the claim that "a need exists for attention to the teaching of probabilistic reasoning in the context of critical thinking programs" that I am most concerned to address here.

In support of this normative claim, I shall examine several of the most prominent philosophical interpretations of critical thinking, and argue that attention to probabilistic reasoning would be appropriate, given many of these conceptual interpretations. In other words, in this section I shall develop the following argument:

(1) given that either "a" or "b" or "c" is the best interpretation of the term 'critical thinking'; then,
(2) if "a" implies probabilistic reasoning; and,
(3) if "b" implies probabilistic reasoning; and,
(4) if "c" implies probabilistic reasoning,

one may conclude that

(5) the teaching of probabilistic reasoning is appropriate in a critical thinking program.

Probabilistic Reasoning Norms Are Not Intuitively Obvious

Third, I shall argue that the short shrift given to probabilistic reasoning in critical thinking programs is problematic, since there is a considerable body of research in sociology and cognitive psychology which suggests that individuals make systematic errors when facing problems that require probabilistic reasoning. Studies by Kahneman, Tversky, Nisbett, Ross and others indicate that individuals develop from their experience sets of judgmental heuristics, or rules-of-thumb, that are employed as guides in the making of probabilistic inferences, often at an intuitive level.[9] These researchers suggest that, although the "intuitive" strategies commonly employed may have been adequate for the limited circumstances in which they arose, generating "successful" judgments in those particular circumstances, those strategies often come to be applied widely and indiscriminately, in circumstances in which they are not at all appropriate. This, it is claimed, leads to frequent errors in judgment.

Further, the sorts of reasoning errors that are described in the psychological literature, it is claimed, indicate the existence of systematic biases, rather than random errors. For example, Kahneman and Tversky maintain many persons combine a systematic overestimation of the variability that exists among persons in their own "in-group", whatever that might be, with a considerable underestimation of the variability that exists among persons in other groups.[10] That is, one might simultaneously reject the stereotyping of one's own ethnic group, recognizing that such sweeping generalizations could never be valid across such a widely variable group, but nevertheless remain willing to make just such sweeping generalizations about other ethnic groups. E.g., One might argue:

> (1) given the range of variability along every dimension among racial groups, racial prejudice of any sort is absurd; therefore
>
> (2) the prejudice whites harbor towards blacks is absurd.

The argument of course would be sound if in the second premise one intends to refer to only those who are indeed racially prejudiced; otherwise the argument illustrates the error Kahneman and Tversky describe, since it would involve failing to recognize that the variability recognized in the first premise applies to <u>every</u> racial group, not merely to that with which one is most familiar.

Such reasoning would not generally take on the structure of a formal argument; the same faulty reasoning might be expressed informally, as in the comment heard on a recent radio talk-show, "The ______ are as narrow-minded and racist as they ever were."

As another example of systematic errors in reasoning, Kahneman and Tversky, among others, argue that individuals commonly display a consistent tendency to give far too much weight to a single, vivid piece of evidence, while giving far too little weight to a large body of statistical evidence, simply because the statistical evidence is more abstract, more remote from the individual's personal experience. For example, a person may place considerable value on a study showing the nation-wide repair rates of an automobile, but nevertheless reject that data on discovering a single piece of countervailing evidence in the experience of a friend.

A third example of claimed systematic bias is what Kahneman and Tversky term the "representativeness" heuristic. This is the induced "rule of thumb" which allows one to infer that a sample drawn from a population will be "very much like" the population as a whole, and hence to infer that a sample has a very high probability of belonging to the population it closely resembles. This seems like a plausible enough rule, but it only "works" when certain other, fairly uncommon, conditions are met. The traits in question must be unique to the population, or at least highly <u>unlikely</u> to be found outside the population, i.e., the traits must be truly diagnostic of that group. For example, if an animal on close examination "looks like a bird", in that it seems to be feathered, it very likely <u>is</u> a bird. But the fact that Judy "looks like a model" does not entitle one to infer that Judy probably <u>is</u> a model. The error lies in the assumption that

> (a) the probability that one <u>is</u> a model, given that one looks like a model

is <u>equal to</u>

> (b) the probability that one looks like a model given that one <u>is</u> a model.

The problem with this rule of thumb arises from the fact that, in truth, there is no necessary relation whatsoever between these two probabilities. This may seem counter-intuitive, but, consider the two probabilities:

(c) the probability that "X" is a male, given that "X" is a past President of the U.S.; and,

(d) the probability that "X" is a past President of the U.S., given that "X" is male.

The correct probabilistic rule is that the two probabilities, Pr(A/B) and Pr(B/A), are completely independent of one another (although, if additional information is available, each can be calculated from the other).

<u>Probabilistic Reasoning May Be Improved via Instruction</u>

The fourth point which suggests that probabilistic reasoning ought to be included in critical thinking programs is that there is evidence to suggest that the incidence of errors in probabilistic reasoning can be reduced when explicit instruction in the norms of probabilistic reasoning is given.

Fischbein[11], for instance, has studied the early development of intuitive beliefs about probability, and noted both positive and the negative effects of instruction and schooling. Holland et al. argue that in reasoning about everyday problems people make use of a set of abstract inferential rules, and that these rules, "...in addition to being induced by people in the course of ordinary daily existence, can also be taught."[12]

In the course of this work I shall argue, by elaborating upon and attempting to establish these four basic points, that "critical thinking" programs would be improved by incorporating attention to "probabilistic reasoning".

<u>Chapter Overview</u>

The overall structure of the argument in this work, then, is this: In chapter II, I shall discuss "critical thinking"; in chapters III and IV, I shall discuss "probabilistic reasoning"; and, in chapter V, I shall discuss issues and research related to the teaching of probabilistic reasoning. I shall conclude in chapter VI by summarizing the points previously raised which suggest that probabilistic reasoning ought to be incorporated in programs that teach critical thinking. I shall consider as well the possible objections to that proposal.

More specifically, in chapter II, I will set out and consider several different interpretations of "critical thinking", since it is necessary to clarify what is <u>meant</u> by the term 'critical thinking' before deciding what would be appropriately included in programs intended to teach it. For instance, one might ask, Is "critical thinking" merely synonymous with *good* thinking? Is it simply "rational" thinking? Or *autonomous* thinking? Or, do particular "*thinking skills*" exist, which can be taught, and, if so, are these skills general, or are the skills specific to particular domains or disciplines? Are there certain *attitudes* or *dispositions* that are necessary for critical thinking? And, if so, can these be taught?

In this chapter, I will set out a brief overview of several prominent approaches to the teaching of critical thinking. It should be noted, though, that there is an enormous body of literature dealing with critical thinking and critical thinking programs, and it is beyond the scope of this work to provide a detailed review of that literature. Moreover, there are a number of excellent reviews of particular programs currently available, most notable that of Nickerson, Perkins and Smith, The Teaching of Thinking, 1985[13], Developing Minds: A Resource Book for Teaching Thinking, edited by A. Costa,[14] and the two volume work, Thinking and Learning Skills, edited by Segal, Chipman and Glaser[15]. In this chapter, then, I will set out several quite different representative interpretations of 'critical thinking', consider the curricular implications of those interpretations, note the paucity of explicit attention to probabilistic reasoning, and indicate points at which attention to probabilistic would seem appropriate.

In this chapter, I shall argue that one undisputed goal of education is the development of rationality, and that at least one aspect of that concept is the adherence to a set of norms for inductive reasoning. I shall further argue that any such norms of inductive reasoning must involve an understanding of the concept of probability, and a knowledge of the norms of "probabilistic reasoning".

I will argue that, under any of the major interpretations of the term 'critical thinking', programs that teach critical thinking would be enhanced by the inclusion of explicit attention to the norms of probabilistic reasoning.

I will also consider in this chapter the issue of "domain specificity" that often arises with respect to critical thinking programs. I will argue that given the ubiquity of probability, instruction in probabilistic reasoning can be expected to lead to exactly the sort of knowledge that would be "transferrable", useful, across wide variety of domains.

In chapter III, I will examine in greater detail the nature of probabilistic reasoning, and the importance of such reasoning, given the common need for probabilistic judgments in practical affairs, e.g., in the processes of estimating, judging, decision-making, learning from experience, and concept formation.

In chapter III, I will also examine in detail some of the claims current in the psychological literature, focusing particularly on the work of Kahneman and Tversky. One claim commonly made in this literature is that people tend to do poorly when faced with probabilistic information because of a reliance on sets of inadequate, mostly intuitive heuristics. For example, it is claimed that people fail to correctly intuit the norms of Bayesian reasoning, and hence fail to arrive at correct revised estimates of probability after having been given new information. I will critically examine several of what have become "classic" examples of such errors and biases, and suggest that in some cases, the researchers themselves have adopted a somewhat less than critical approach to probabilistic reasoning.

In Chapter IV, I will discuss the normative issues, i.e., the controversies that exist as to what counts as "correct" probabilistic reasoning. I will discuss "Bayesian reasoning", and examine the current controversy surrounding the use of Bayes' theorem as a norm for probabilistic reasoning. I will examine several of the well-known "paradigm problems", giving particular attention to the "cab problem" originally set out by Kahneman and Tversky. Finally I will discuss the implications of the existence of this sort of controversy for programs in critical thinking that set out to incorporate attention to probabilistic reasoning.

In chapter V, I will consider the issues related to the teaching of probabilistic reasoning. I will discuss the research into the origins of the ideas of chance and probability in young children, and the early development of "intuitions" with respect to probability. I will discuss the currently popular notion that the mind in some way functions as an "intuitive statistician". This model of the mind, which has been dominant in psychology for many years, has recently been seriously questioned[16]. I will set out some of the empirical work relevant to the evaluation this model, including some of the work on the development of the child's ideas of chance and probability, and discuss the educational implications of this issue. I will argue that the <u>lack</u> of an innate, intuitive understanding of probability provides all the more reason to explicitly <u>teach</u> the norms of such reasoning to students.

I will also examine current approaches to the teaching of probabilistic reasoning, including curricula that aim to introduce students of various ages to the basic concepts of chance and probability, "thinking under uncertainty" or to the axioms of the probability calculus.

In chapter VI, I will summarize the points made, draw out some of the implications for educational programs in critical probabilistic reasoning, and consider some possible objections.

It should be noted that it is not my intention to suggest that probabilistic reasoning would be in any way a panacea to the problems of teaching critical thinking, but merely to argue that it is a valuable addition to the individual's knowledge--an often overlooked but useful element in the evaluation of arguments <u>and</u> in the evaluation of information. In addition, since the lessons learned from an examination of the norms of probabilistic reasoning would often be surprising and counterintuitive, the instruction in probabilistic reasoning would serve to encourage students to develop a critical attitude toward their own "intuitively obvious" beliefs.

1. Paul Slovic, Baruch Fischoff, and Sarah Lichtenstein, "Facts versus fears: Understanding perceived risk," in *Judgment under uncertainty: Heuristics and biases*, ed. by Daniel Kahneman, Paul Slovic and Amos Tversky (Cambridge: Cambridge University Press, 1982), pp. 463-489.

2. Gerd Gigerenzer *et al.*, *The Empire of Chance: How probability changed science and everyday life* (Cambridge: Cambridge University Press, 1989).

3. Horace C. Levinson, *Chance, luck and statistics* (New York: Dover Publications, 1963), p. 3.

4. Ibid., p. 29.

5. John McPeck, *Teaching Critical Thinking: Dialogue and Dialectic* (New York: Routledge, Chapman and Hall, 1990), p.3.

6. Ibid.

John E. McPeck, *Critical Thinking and Education* (Oxford: Martin Robertson & Co. Ltd., 1981).

7. Marcia Heiman and Joshua Slomianko, eds., *Thinking Skills Instruction: Concepts and Techniques* (Washington, D.C.: NEA Professional Library, 1987).

Arthur L. Costa, ed., *Developing Minds: A Resource Book for Teaching Thinking* (Association for Supervision and Curriculum Development, 1985).

Marilyn Burns, *The Book of Think (Or, How to Solve a Problem Twice Your Size* (Boston: Little, Brown and Company, 1976).

Barry K. Beyer, *Teaching Thinking Skills: A Handbook for Secondary School Teachers* (Boston: Allyn and Bacon, 1991).

Barry K. Beyer, *Developing a Thinking Skills Program* (Boston: Allyn and Bacon, 1988).

Ronald R. Edwards and Wanda D. Cook, *Problem Solving through Critical Thinking, Grades 5-8* (New Rochelle, N.Y.: Cuisenaire Company of America, Inc., 1990).

Dale Seymour and Ed Beardslee, Critical Thinking Activities in Patterns, Imagery and Logic (Palo Alto, CA: Dale Seymour Publications, 1990).

Arthur A. Hyde and Marilyn Bizar, Thinking in Context: Teaching Cognitive Processes across the Elementary School Curriculum (New York: Longman Inc., 1989).

Vincent Ryan Ruggiero, The Art of Thinking: A Guide to Critical and Creative Thought (New York: Harper & Row, Publishers, 1984).

Vincent Ryan Ruggiero, Beyond Feelings: A Guide to Critical Thinking (Palo Alto, CA.: Mayfield Publishing Company, 1984).

Matthew Lipman, Thinking in Education (Cambridge: Cambridge University Press, 1991).

Robert E. Young, ed., Fostering Critical Thinking: New Directions for Teaching and Learning (San Francisco: Jossey-Bass Inc., Publishers, 1980).

Edith D. Neimark, Adventures in Thinking (San Diego: Harcourt Brace Jovanovich, Publishers, 1987).

Edward De Bono, Teaching Thinking (New York: Penguin Books, 1978).

Raymond S. Nickerson, David N. Perkins and Edward E. Smith, The Teaching of Thinking (Hillsdale, N.J.: Lawrence Erlbaum Associates, Publishers, 1985).

8. Richard Pratte, Philosophy of Education: Two Traditions (1992).

9. John H. Holland, Keith J. Holyoak, Richard E. Nisbett, and Paul R. Thagard, Induction: Processes of Inference, Learning, and Discovery (Cambridge, Massachusetts: The MIT Press, 1989).

Daniel Kahneman, Paul Slovic, and Amos Tversky, (ed.), Judgment Under Uncertainty: Heuristics and Biases (Cambridge: Cambridge University Press, 1982).

Richard Nisbett, and Lee Ross, Human Inference: Strategies and Shortcomings of Social Judgment (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1980).

10. Richard E. Nisbett, D.H. Krants, D. Jepson, and Z. Kunda, "The use of statistical heuristics in everyday inductive reasoning". *Psychological Review* 84 (1983), pp. 339-363.

G. Quattrone and E. Jones, "The perception of variability within in-groups and out-groups: Implications for the law of large numbers". *Journal of Personality and Social Psychology* 38, pp. 141-152.

11. E. Fischbein, *The Intuitive Sources of Probabilistic Thinking in Children* (Boston: D.Reidel Publishing Company, 1975).

12. Holland et al., *Induction: Processes of Inference, Learning and Discovery*, p. 255.

13. Raymond S. Nickerson, David N. Perkins and Edward E. Smith, *The Teaching of Thinking*.

14. A. Costa, ed., *Developing Minds: A Resource Book for Teaching Thinking*.

15. Judith W. Segal, Susan F. Chipman, and Robert Glaser, eds., *Thinking and Learning Skills*, vols. 1 and 2 (Hillsdale, N.J.: Lawrence Erlbaum Associates, Publishers, 1985).

16. Gerd Gigerenzer and David J. Murray, *Cognition as Intuitive Statistics* (Hillsdale, N.J.: Lawrence Erlbaum Associates, Publishers, 1987).

CHAPTER II

INTERPRETATIONS OF 'CRITICAL THINKING'

Introduction

Clearly, the assessment of the success of any particular approach to the teaching of critical thinking will depend on the conception of "critical thinking" that is adopted. Yet, currently there exists considerable disagreement among educational philosophers as to the most useful or appropriate interpretation of the term 'critical thinking'.

"Critical thinking" is variously interpreted as:

1. simply, "good thinking";
2. thinking with a particular attitude or disposition;
3. "skillful thinking";
4. "logical thinking";
5. "knowledgeable thinking"; or
6. "rational thinking".

Most theorists, though, would agree that, in some sense, "critical thinking" is closely related to "rational thinking" (although this scarcely helps, since the term 'rational thinking' is itself variously interpreted--a purely verbal agreement); most would also agree that "critical thinking" is the sort of thinking that would be useful in "everyday" reasoning.

In this chapter I will consider several philosophical interpretations of the concept of critical thinking, focusing largely on the work of McPeck and Siegel. McPeck and Siegel are prominant theorists in the critical thinking movement whose views are particularly interesting in the present inquiry because, in virtually every aspect of the conceptualizations of "critical thinking", McPeck's and Siegel's views differ radically. I will point out what I take to be flaws in each interpretation, develop what I take to be the strongest possible version of each major interpretation, and discuss the implications of each for the teaching of critical thinking. I will then indicate the relevance of probabilistic reasoning for critical thinking programs, given either of these different views.

It is not my intention here to try to resolve the underlying question, i.e., "what is the "best" interpretation of the term 'critical thinking'?". I will argue, though, that given the most plausible version of each of these interpretations, attention to probabilistic reasoning would be appropriate and advantageous.

I should note at the outset that most writers on critical thinking employ the term, at various points, in several of the senses listed above. Clarity is served, however, by sharply distinguishing the different possible senses.

1. Critical Thinking as Good Thinking

The term 'critical thinking' is used at times as if it were simply a synonym for "good thinking" or "intelligent thinking". This seems to be the broadest possible construction to place on the term--given this interpretation, any technique that improves thinking, in any way, would be appropriately taught in a critical thinking course.

McPeck's Usage

McPeck, at times, uses the term 'critical thinking' in this very broad sense (although this, it should be noted, is not his principal thesis). McPeck writes, for instance, that "intelligent thinking" is precisely what most people mean by critical thinking--he argues that "...if the disciplines are properly taught, we will get the kind of intelligent thought from students that we normally associate with the phrase *critical thinking*."[1] Elsewhere he writes, "...our public schools would like to prepare people for making *intelligent* decisions..."[2], and that this goal is accomplished when students become critical thinkers.

Used in this way, the term 'critical', applied to thinking, is simply a sign of acclamation, or commendation. It is used to confer value, but it gives no indication as to what it _is_ about the thinking that is considered praiseworthy. Any thinking that is approved of would thus "count" as critical thinking. This interpretation of the term is too broad to be of any practical use in designing a curriculum, since one is immediately faced with the question, What is it about a particular instance of thinking that makes it count as "good" or "intelligent"? And to this question, no answer is given.

2. Critical Thinking as an Attitude

Critical thinking is sometimes interpreted as requiring, at least in part, a particular attitude or disposition toward the activity of thinking. This emphasis on encouraging thinking "with a particular attitude" is of course compatible with other interpretations, e.g., those that place an emphasis on teaching logical norms of thinking or the norms of particular disciplines. The critical thinker needs not only to be aware of the norms, and to _recognize_ when particular norms are applicable, but also to actively appreciate the _importance_ of those norms, whether of logic or of the particular subject area. He or she ought to see conforming to such norms as a goal clearly worth pursuing, not as just another burdensome set of rules set out by an authority. It would seem that only with such an appreciation would the student ultimately develop a disposition to actually engage in critical thinking.

Both McPeck and Siegel would require a certain attitude of a person if that person is to "count" as a critical thinker, and presumably, if the person's action is to count as an act of critical thinking.

It should perhaps be noted, though, that while the "disposition to act" in certain ways may be considered a necessary condition of a person's being a critical thinker, that disposition ought not to be considered a necessary condition for "critical thinking" per se. One could, I take it, conceive of a single, isolated and never repeated act of critical thinking. But, this would not be possible were the disposition to think critically to be taken as part of the meaning of 'critical thinking', since the single instance can occur without the disposition.

A. McPeck's Usage

According to McPeck, the particular attitude required for critical thinking is skepticism, of an appropriate sort.

One's skepticism must not be "pernicious"; rather, it must be "judicious", or "healthy", or alternatively, "reflective".[3] The critical thinker must, moreover, have a disposition to be judiciously skeptical. The interpretation of 'skepticism' McPeck adopts seems clear--to be skeptical one is to suspend belief that available evidence (and by this it seems that McPeck means that evidence which is generally *accepted*) is sufficient to warrant belief.[4]

Unfortunately, the notion of a "judicious" or "reflective" skepticism is not explicated, except by the stipulation that reflective skepticism is skepticism "intended to advance progress toward the resolution of a problem"[5]. McPeck does not explain how one could know in any particular case that one's skepticism would count as "judicious", rather than, say "foolish". It seems that the "judicious" skeptic in McPeck's writing can be understood merely as one who is skeptical at the right time, i.e., just when he or she ought to be skeptical. Yet if this interpretation is correct, the claim that one should be judiciously skeptical would be necessarily true, but also tautological and hence empty.

According to McPeck, one engages in this reflective skepticism only when one becomes doubtful about the validity of the accepted criteria of particular field. "Normal", everyday reasoning within the field would not require such an attitude, since in such reasoning one would simply employ the usually accepted criteria without question. Hence "critical thinking" for McPeck is sharply distinguished from "... 'normal correct thinking' (or standard *disciplinary* thinking)..."[6]

Since, through the disciplines, we have "...developed entire networks of concepts, methods, and procedures for dealing with an enormous spectrum of life's familiar (and unfamiliar) problems..."[7], the best way, indeed, the "rational" way of approaching any particular problem, according to McPeck, is simply to use those established procedures. Uncritically? Evidently, yes. "Critical thinking" on this view only occurs when "...we have reason to suspect that the normal procedures, or beliefs, leave something to be desired...on those comparatively rare occasions where we suspect something is amiss. On such occasions it is right and proper to start questioning some of our fundamental assumptions, or beliefs, and to try alternatives..."[8]

On this account, it seems that "critical thinkers" would be only those working on the "cutting edge" of a field of inquiry, those who are in a position to advance human knowledge by challenging and changing the established norms of "rational" thought as understood within a particular discipline. Einstein, for example, would count as a critical thinker. Most of the rest of us would have to settle for a mere "rationality".

One problem with this account is that, in limiting "critical thinking" to those who are engaged in challenging the fundamental criteria of judgment in a discipline, this account does not seem to capture what is generally understood by the term 'critical thinking'. If this is the meaning of critical thinking, there would be little reason to require that all individuals learn to think critically.

B. Siegel's Usage

Siegel, like McPeck, includes as a necessary condition of critical thinking what he terms the "critical attitude or critical spirit component of CT."[9] ("CT" is Siegel's standard abbreviation of 'critical thinking'.) This attitudinal component is "the willingness, desire, and disposition to base one's actions and beliefs on reasons, that is, to do reason assessment and to be guided by the results of such assessment."[10]

Here, it should be noted, Siegel brings in two different dispositions, which for clarity should be distinguished.

First, Siegel requires that one have a disposition to _do_ reason assessment, i.e., to engage in critical thinking. But, as was argued above, it is difficult to see how the "tendency to _do_ X" can be part of the _meaning_ of X.

The second disposition Siegel requires is the tendency _to be guided by_, or to actually _act_ on, the results of one's critical thinking. That is, a purported instance of "critical thinking" must actually have issue in action (or belief), or it would not count as "critical". But, if this disposition is accepted as part of the _meaning_ of critical thinking, certain problems arise. Suppose that one encounters a doubtful, problem situation, and begins to think. Suppose, further, that one formulates hypotheses about the situation, and makes observations to confirm or deny the hypotheses, and that, in so doing so, one remembers and applies the norms of logic, of informal reasoning, and of probabilistic reasoning. Suppose that one even considers the particular standards of evidence extant in the particular disciplines relevant to the problem. Given Siegel's conception of critical thinking, we could not yet say that this person has been engaged in critical thinking. That assessment could only be made after we observe the thinker actually acting, or failing to act, on the basis of the reasoning. In some cases, e.g., in long-range planning, it might well be years before one could decide whether the thinking that took place was, or was not, "critical". Moreover, suppose the person decides to act on the reasoning, and does act on it, but then later in the same circumstances, recalls that earlier thinking, but decides not to act upon it? Given Siegel's account, it would seem that the same instance of reasoning, if acted upon, would count as an instance of "critical thinking", but later on, when not acted upon, would not count as such. But this seems an unacceptable consequence of the requirement that "willingness to act" be part of the meaning of 'critical thinking'.

There are several other problems with this explication of 'critical thinking'. First, both of these dispositions, to do critical thinking, and to act upon that thinking, would seem to be characteristics of the person doing the thinking, not of the thinking itself. But unless we can accurately characterize a single instance of critical thinking, we cannot tell whether a person has indeed engaged in a long string of such thinking, and so we cannot tell whether the person has or has not developed a disposition to think critically.

Siegel seems to treat "CT" as a sort of lifestyle that one is to adopt, rather than as particularly way of thinking that would be useful if it were generally followed. Siegel argues that it is only possible to characterize "critical thinking" in the full sense by setting out the characteristic traits of the person who _is_ a critical thinker.

While it is possible to require that each act of thinking be undertaken with a certain attitude in order to count as "critical thinking", the requirement that a certain _disposition_ exist could only apply to the _person_ doing the thinking.

I would agree that the notion of "critical thinking" does indeed imply an attitude, viz., critical; the question is, critical of _what_? And in what sense?

The answer, I take it, is critical of one's _own_ thinking processes, i.e., one must be "self-critical". The intended sense is not the sense of "harsh or negative assessment" of one's self. Rather, what is required is the application of a set of norms to one's thinking, by means of which one may judge the quality of that thinking. "Critical thinking", in other words, involves the _critiquing_ of one's own thinking. It is this sort of critical attitude that would be necessary to lead one, at any particular juncture, to recall or to seek out appropriate criteria and standards of judgment, whether from the field of logic or from the specific "subject-area", and to apply those standards to the problem at hand. And, if one were to habitually maintain that attitude, and persistently engage in this sort of critique, one would then "count" as a "critical thinker", i.e., one who has the disposition to think critically.

C. Probabilistic Reasoning and the "Critical Attitude"

It is the absence of just such a critical attitude that seems particularly evident whenever "numbers" creep into an argument, e.g., when one must estimate the chances or probability of an event. The same lack of self-critique seems to arise regularly in the making of predictions and class attributions. This would not, perhaps, be thought problematic, were it not for the rather large body of evidence brought out in the psychological literature indicating persistant cognitive illusions, biases, flawed heuristics and incorrect "intuitions" with respect to probability. An uncritical reliance on one's own existing beliefs can thus easily lead one into error. But, the deficiency in critical attitude cannot be remedied unless one possesses a set of norms by which to assess one's thinking. To engender the desired attitude in probabilistic problems, students must be taught the norms of probabilistic reasoning.

The study of probabilistic reasoning would constitute a particularly fertile field for exploration in a critical thinking classroom, because, while the <u>basic</u> norms of probabilistic reasoning are simple and well understood, the application of these norms to particular problems is often highly counterintuitive and seemingly paradoxical.

The student, in studying probabilistic reasoning, is thus in a position to learn: a) that general norms of reasoning do exist, and should be applied; b) that one cannot simply memorize a list of rules of reasoning to be mechanically applied--one must continue to ask how and whether a particular standard would be applied to a particular situation or problem; and c) that one's <u>own</u> initial judgments, and a *fortiori*, one's "intuitions" are subject to critical evaluation, and to change as the result of that evaluation. This is precisely the explication given by McPeck (quoted above) of the "critical attitude".

<u>3. Critical Thinking as Skilled Thinking</u>

This interpretation is built on what seems to have become a dead metaphor, that "thinking" is a *mental skill*, comparable in significant ways to ordinary physical skills. Like other skills, thinking would simply be an activity at which one becomes more proficient over time, given training and practice.[11]

The "thinking skill" metaphor may sometimes taken quite literally. Thinking is then understood as an ability to perform certain cognitive functions, and it is the student's ability to think that is to be improved, via exercises in thinking. The primary focus of a course in thinking would be on the improvement of the thinking process itself, construed as a set of physical functions. This end, it seems, would be accomplished by improving the efficacy of brain functions, e.g., improvement of memory, short and long-term recall, or the ability to perform analogical thinking.

A. Norris' Usage

Just such an interpretation is set out by Steven Norris. Norris writes: "To say someone has 'critical thinking ability' is to make a claim about a mental power which that person possesses. Mental powers, in turn, arise from mental structures and processes..."[12] Hence, according to Norris, to effectively teach "thinking", for instance, mathematical thinking ability, we would first need to discover "the mental processes which the child uses to solve... problems..."[13]

Some of the specific "thinking skills" that might be taught, for example, would be the skill of analogical reasoning, the skill of problem-structuring, analysis, classification, categorization, and synthesis. Just such a skill-oriented program is described by Charlton, in "The Direct Teaching of Analysis"[14]. In this program each particular "cognitive skill" is itself explicitly taught before any attempt is made to actually apply the skill to any subject matter. For instance, the student would memorize the "steps" in analysis, and would only then practice the skill of analyzing in various contexts.

The "skills" paradigm was introduced by Bartlett in 1958.[15] Although seldom taken as literally as it seems to be by Norris, the skills paradigm arises frequently. This model is strongly critiqued by McPeck.

B. McPeck's Critique of Thinking Skills

1. General Thinking Skills

McPeck rejects the notion that critical thinking is a single mysterious "general skill", similar to general intelligence (despite his occasional equating, noted above, of critical thinking and intelligent thinking). McPeck argues that, just as one would not set out to teach a "general athletic skill" dissociated from any particular skills in some particular sport, similarly, one should not set out to teach "general thinking skill", dissociated from any particular "skill" in reasoning within the context of some particular subject areas, viz., the disciplines.[16]

McPeck rejects the belief that one can be taught "thinking-in-general", because, he writes, whenever one thinks, one must be thinking "about some particular thing or subject".[17] He concludes that, since "...specific subject content determines the required ingredients of thinking critically...the notion of 'general critical thinking skills' is largely meaningless."[18]

McPeck argues that the term 'reasoning ability' itself is misleading and should be eschewed, since it subtly suggests the existence of some sort of "single underlying capacity"; he argues that "reasoning ability covers all manner of cognitive phenomena, scarcely any cluster of which resembles another".[19] Hence no single generalized program could be effective in improving a person's performance of all those various specialized cognitive skills. McPeck concludes that the term 'general reasoning ability' itself is incoherent.

2. Specific Skills

McPeck also rejects the notion that critical thinking is to be equated with the use of a single set of what he terms "specialized skills", viz., the "specific skills" of logic and argument analysis. He continues, though, at times to speak of critical thinking as a set of mental "capacities". For instance, he tells us that conscientious teachers try to "improve the various thinking capacities of students".[20]

Nevertheless, despite the rejection of both the "specific skills" and "general skill" models of critical thinking, McPeck does seem to implicitly accept the notion that critical thinking is some sort of a skill. Like Norris, McPeck speaks of critical thinking as an "ability" that must be developed, or a "capacity" that must be fulfilled. "Reasoning ability", he writes, is a matter of the interaction of numerous "cognitive phenomena". The mind, it would seem, can be made to function more efficaciously if given the proper instruction, and when this occurs, better thinking will result. Presumably, from an improved thinking process, a better quality of thoughts will emerge.

3. Logical Skills

According to McPeck there is a sharp, qualitative difference, a difference in kind, between: a) the teaching of "logic", which he interprets as the teaching of a set of "specific thinking skills"; and b) the teaching of "the disciplines", which he interprets as the attempt to teach "information", or a body of knowledge. But this distinction breaks down when the teaching of logical reasoning "skill" is understood as the teaching of a set of norms by which lines of reasoning are to be judged, rather than as the teaching of a set of "skills" in the ordinary sense. Logic can be considered one of the disciplines to be taught, a particular body of normative knowledge which differs from other disciplines only in the broad, indeed universal, applicability of its subject matter.

McPeck does argue that "classical logic" simply does not have this broad applicability across the traditional disciplines[21]--a point to be considered in detail later.

C. A Critique of the Thinking Skills Paradigm

But despite the common use of the term 'thinking skills', it is not clear that this is a useful metaphor, nor that in teaching "critical thinking" we aim to improve a certain set of desirable skills or cognitive capabilities. The questions that must be addressed are, Is "thinking" most usefully construed as a skill? How much similarity exists between the putative mental skill of thinking and an undisputed "model case" of a skill, a physical skill? In what sense, if at all, is it useful to construe thinking as a sort of skill, and "critical thinking" as skillful thinking?

As a model case of a skill that a student might acquire, we might take "typing". In teaching this physical skill, one would try to improve upon, or to fine-tune, certain neuromuscular phenomena--e.g., eye-hand coordination, rhythm, speed, agility and strength. With practice, these phenomena would become automatic, habitual; the physical activity would no longer require a conscious effort. Nor would the performance of the activity require any <u>thought</u>. The skilled typist could initiate the activity at will, and could perform it proficiently without needing to devote any particular attention to it.

Yet, in this respect the putative mental skill differs from the physical. One could not say, for instance, that one has developed one's "skill in analysis" to such a high level that one is now able to perform the most complex of analyses, without even thinking about it.

With respect to the physical skill of typing, one would expect to see a considerable amount of "transfer" of training--e.g., the student who was an accomplished pianist would probably do much better at typing, at least in the beginning, than the student who had no prior opportunity to develop manual dexterity.

In teaching a (primarily) physical skill such as typing, one would rely largely on an extensive repetition of various exercises of increasing difficulty. Unfortunately, while we have a good idea of the sort of exercises that might improve this sort of physical skill, that would facilitate a particular neuromuscular performance, we seem to have no comparable knowledge about how to improve the "cognitive phenomena", i.e., the mental functions that presumably are involved in performing a mental task. The development of a set of exercises that would actually improve the functioning of the mind/brain, and that would in this sense "improve thinking", seems a remote possibility (though such claims are often made, for instance, for various meditation exercises said to improve concentration, or biofeedback devices to enable one to generate at will particularly propitious brain-wave patterns, etc.).

It seems true that, if one were to set out to improve "thinking skills", in this sense, in which the skills in question are taken to be analogous to ordinary physical skills, one would need to focus attention on the simpler skills, rather than on some single complex, and putative "general skill". Paradigm cases of such mental skills might be: skill in observation; in memorization and recall; in rapid reading; and even, perhaps, in speaking foreign languages. In each of these cases one could readily imagine repetitive exercises that would bring the student from the initial unskilled state of clumsy ineptitude, to a state of far greater "mental proficiency". The activity, for instance, understanding Russian, would change from one requiring considerable mental effort, attention and will-power, to one that proceeds habitually and effortlessly.

But, unfortunately, it is not at all clear that the student whose "mental skills" had been so improved would, as a result of that training, be a "critical thinker" in the commonly accepted sense of being in some sense "more rational", nor would that person seem likely to be any more successful at resolving everyday problems than one whose mental functioning was less skilled.

The "skill" in the thinking skills metaphor could, however, be interpreted in another way. Just as a "skilled" carpenter is one who possesses a great deal of knowledge about his or her particular craft, a "skilled thinker" might be construed as a thinker who possesses, and is able to make appropriate use of, a body of expert knowledge about the thinking process itself. The skilled carpenter is not merely proficient at hammering nails; he or she knows when to use certain types of nails, e.g., galvanized vs ordinary, to accomplish a certain goal in certain context. The skilled carpenter has an extensive background of knowledge and experience, and a history of successful application of the knowledge base to a variety of problems.

Programs to develop "thinking skills", if based on this conception of skill, might adopt a "meta-cognitive" approach, teaching students to become conscious of their own processes of reasoning, through, for instance, "think-aloud" exercises or the development of "verbal protocols".

Or, such programs might emphasize instruction in the use of efficacious "thinking strategies". Students might be taught, for instance, specific problem-solving strategies, such as "using trial and error", or "working backward in problem-solving".[22] An example of such an approach would be

De Bono's CoRT Thinking Lessons.[23] The emphasis here is on the skillful performance of the activity of thinking.

But, the thinking skills metaphor could be taken in yet another sense. The "skilled thinker" might be one who possesses and makes use of a particular body of expert normative knowledge, viz., knowledge not merely about the process, and/or strategies of thinking, but knowledge of the criteria by which the product of that reasoning process is to be judged. These criteria may be variously interpreted--one might take these norms to be, for instance, the norms of formal and/or informal logic, (as do those in the Informal Logic Movement) or, as particular sets of norms applicable to various fields of study (as does McPeck). The critical thinker, "skilled" in this sense, is thus one who is in a position to evaluate arguments, to determine whether conclusions are warranted. The emphasis is less on production of a skillful performance than on the development of a discriminating judgment in the evaluation of the performance, i.e., of the thoughts of oneself and others.

Note that, in common cases of physical skills, the ability to perform and the ability to judge a performance are independent. One may possess either ability without possessing the other. For instance, a judge of a gymnastic competition possesses the knowledge needed to evaluate and critique the gymnast's performance, but would not personally possess the ability to skillfully perform the activity being judged.

But, this seems an odd aspect of this version of the "skills" model, when applied to thinking. The person who possesses the normative knowledge necessary to assess thinking, who can appropriately critique a line of reasoning, and who is engaged in judging his or her own thinking according to the applicable norms, is thinking critically. I.e., the one who skillfully judges gymnastics is not by virtue of that judging activity doing gymnastics, but the one who skillfully judges thinking is, in the very act of judging, doing critical thinking. The term 'skill' when used in this sense is misleading, since what both the "skilled" judge of an Olympic performance and the "skilled" judge of thinking are expected to possess is knowledge and understanding, rather than a "mental ability".

D. Probabilistic Reasoning and Thinking Skills

In many every-day problems, for instance, those sketched out in the introduction, and others to be examined in detail in later chapters, the applicable norms are the norms of probabilistic reasoning. Hence, if the "skilled thinking" interpretation of critical thinking is construed in this last sense, the successful critical thinking program should provide the student with an understanding of the norms of probabilistic reasoning. Such a program, involving the direct teaching of those norms, would be a necessary addition to critical thinking courses if a) the appropriate norms are not generally induced from experience, and b) the problems requiring probabilistic reasoning are both common and significant. The evidence set out in the next chapter suggests strongly that both (a) and (b) are the case.

4. Critical Thinking as Logical Thinking

In this interpretation, "critical thinking" is understood to be that thinking which conforms to the standards of deductive logic and/or informal logic. To become a "critical thinker" the student must acquire an understanding of the rules and norms of logic, as well as some competence in developing and critiquing arguments. Critical thinking programs based on this interpretation stress the construction of valid deductive arguments, and the recognition and avoidance of informal fallacies. Such programs have been said to constitute the "standard" approach. These programs, however, would be incomplete as introductions to logic unless explicit attention is given to the reasoning norms applicable to inductive arguments. Although informal logic courses do point out some probabilistic errors, e.g., hasty generalization, general observations about the existence of such fallacies would not teach students the factors that would make particular generalizations "hasty" or unwarranted. Without an understanding of those factors, the student is unprepared to avoid the probabilistic errors.

McPeck rejects the "analysis of arguments" interpretation of critical thinking, concluding, oddly enough, that argument analysis ought not to be regarded "as the equivalent to, or a substitute for, "*general reasoning ability*" (emphasis added)[24]. But, he has already rejected the notion that thinking critically involves the deployment of some general "thinking ability".

One should also note that, as indicated above, when one teaches argument analysis, it is not a "skill", but a set of norms and criteria for good reasoning that is being taught. It is only when the student has acquired a knowledge that those norms exist, and some knowledge of what they are, that that student is in a position to critically and effectively evaluate the quality of his or her own arguments.

A. The Problem of Transfer

McPeck also at times argues the empirical, psychological thesis that students' knowledge of the norms of deductive logic and/or informal logic, whether applicable or not, simply does not "transfer" across the boundaries of the disciplines. McPeck thus rejects the "standard" approach to the teaching of critical thinking on the grounds that it is pedagogically ineffective.

McPeck consistently contrasts what he terms the "standard thinking skills" interpretation of critical thinking, i.e., the "logical thinking" interpretation, with his own "knowledge and information" interpretation. He does, however, acknowledge that "...some kinds of specific knowledge and information will have far more transfer capacity than other kinds..."[25], that the problem is to find "...the kind of knowledge likely to be the richest or most powerful from the point of view of transfer."[26] McPeck thus rejects the "standard" approach to the teaching of critical thinking as pedagogically ineffective.

McPeck's <u>answer</u> to the question "What knowledge is richest and most powerful?" is, simply, the knowledge of the traditional disciplines that a liberal education comprises.

"...I see no competitive substitute for a liberal education."[27] Here as elsewhere McPeck ignores the obvious, that those who interpret "critical thinking" as logical thinking and recommend the explicit teaching of critical thinking do not at any point suggest this as a "substitute" for a liberal education, but simply as a significant component _of_ such an education.

While the question of "transfer", as often noted, is an empirical question, the more important question of the _applicability_ of certain sorts of knowledge is not. It would seem that one kind of knowledge that would have the desired broad applicability would be the knowledge of the norms of logic; a second sort would be knowledge of the norms of probabilistic reasoning.

But McPeck explicitly rejects the notion that logic has this broad applicability, arguing that, although "logic" is valuable, it cannot be and _ought not_ to be applied indiscriminately across disparate fields. McPeck asserts that there does not exist a single monolithic "logic" that is applicable to all fields of inquiry. Each separate discipline, he claims, has a "logic" all its own: "...there are almost as many distinguishable logics, or kinds of reasoning, as there are distinguishable kinds of fields."[28]

Hence, any attempt to teach a single logic that would be equally applicable across different disciplines is (logically?) doomed to fail.

McPeck does acknowledge that "the tools and rigor of logic" would be important in thinking critically, even within particular fields. However, he makes a number of dubious assertions about the nature of "logic". For instance, he asserts that there exist uniquely appropriate "logics" for different subjects. "Just as there are different kinds of 'language games'...so there are different rules of predication, or 'reasoning'...which govern the different kinds of thought.[29] He goes on to claim that the syntax of each logic is determined by the semantics: "A bit more formally, the actual rules for what is a 'well-formed formula' (i.e., an intelligible statement) is [sic] determined at the *semantic* level of discourse. Thus, there are almost as many distinguishable kinds of reasoning as there are distinguishable kinds of subjects."[30]

But this claim is simply untrue--correct syntax in a system of logic is not determined by the semantics of the terms used in that system. It is true, though, that semantics, the meanings of the terms, would determine which predicates are in fact generally meaningfully applied to which entities. For instance, one could not correctly assert (except metaphorically) that "sleep is green", or "snow is happy", since those particular predicates happen not to apply to those entities. But this does not establish McPeck's point--to say that some predicates are only applicable to and possibly true of certain entities is not to say that the rules of deductive logic are subject specific.

McPeck does allow that "classical logic" is not entirely irrelevant to his postulated "unique logics" of the separate disciplines. He writes that, although the rules of the logics of each field are "...not determined by classical logic, these rules will, by and large, obey classical logic."[31]

McPeck's position raises some questions. The first is, is it true that there are "special logics", applicable only to particular curricular subject areas? The implication is that there are "logics" that apply only to Art, Music, Physics, Biology, History, Literature, etc.. And if so, the "logic" applicable to, say, Modern Literature could be different from that used to understand Shakespeare, or Beowulf. If, as McPeck claims, one needs to learn a different logic for each "distinguishable field", and if there is no non-arbitrary way to limit what counts as a "distinguishable field", then one must postulate an infinite number of specific, non-generalizable "logics".

The second question is, are there aspects of deductive logic that not only would not apply to particular fields, but would be inimical to the understanding of particular fields? I.e., would the learning of logic be not only insufficient, but actually detrimental to one's effort to grasp the postulated "special logics" of the particular subject areas?

It seems McPeck would answer in the affirmative on both counts. He writes, for instance, that mathematical reasoning and moral reasoning or literary reasoning are so different that "...not only are the canons of validity different, but what might be fallacious reasoning in one context or domain, might be perfectly correct in another.[32] And this he asserts is a "fact".

If this view were correct, then the study of the norms of logic and, a *fortiori*, of probabilistic reasoning, would be of limited value. But it seems that when McPeck speaks of the "logics" of different fields, what he refers to are not norms of logic, but rather merely the conventions of methodology, and the accepted explanatory theories in different fields. The chief and definitive characteristic of a "logical truth" is that such a statement is true in every domain, given any interpretation. Its proof is domain-independent. For instance, in every context, it is true that (x)Fx --> Fa. This is simply what is meant by saying that "x" is a logical truth. Hence it is a simple error in language usage to claim that "x" is a logical truth but that the truth of "x" is domain-specific. So, McPeck's sets of "domain-specific truths" cannot be said to constitute "logics" in the strict sense.

But, having rejected the universality of logical truths, McPeck goes on to deny that critical thinking should be understood as logical thinking of any stripe. According to McPeck, "logic is to be distinguished from critical thinking precisely because it is not logic, but information which is relevant to reason assessment".[33] In the "area of complex information ... the lion's share of the difficulty comes from the intelligibility and reliability of this information".[34] McPeck's point here, that one key aspect of any critical thinking is the assessing of information, is a good one. Let us turn, then, to McPeck's positive thesis, that critical thinking ought to be understood as the successful evaluation of knowledge claims within a particular discipline.

5. Critical Thinking as the Evaluation of Knowledge Claims

A. McPeck's Interpretation

This interpretation of critical thinking is advanced by McPeck, principally in Critical Thinking and Education (1981)[35], and in Teaching Critical Thinking (1991)[36].

McPeck begins the explication of "critical thinking" by examining the characteristics of the individual who does the critical thinking--critical thinking, he writes, is "that which a critical thinker does". And in McPeck's view, the "critical thinker" is one who has "...the ability to reflect upon, to question effectively, and to suspend judgment or belief about the required knowledge composing the problem a hand."[37] A person is a critical thinker precisely when that person possesses "the disposition and the skill" to suspend his or her belief that "the available evidence from the pertinent field or problem area...[is]... sufficient to establish the truth or validity of...[a proposition within that field]."[38]

It would seem, at first glance, that, given this interpretation, critical thinking would not be much of a trick--since virtually any individual capable of thinking at all would be mentally capable of "reflecting", of "suspending belief", and of "questioning". The difficulty in achieving "crit-ical" thinking of this sort arises only from McPeck's stipulation that the critical thinker's questioning must be "effective"; his or her suspension of belief must be a "skilled" suspension. But, without an explication of what it is that makes an instance of questioning "effective", this interpretation adds little to the notion that critical thinking is "good thinking".

McPeck argues that "critical thinking" can only be taught within the confines of the traditional disciplines, since the student must first be made aware of what constitutes "accepted evidence" for propositions in a given field, before he or she is in a position to be critical of that evidence in the desired effective way.

McPeck sometimes terms his interpretation an "epistemolog-ical" or "semantic" approach. He sets out what seems an uncontroversial point, that "...a minimal condition for understanding a good reason in any field is that one under-stand the full meaning of the specialized, and often techni-cal, language in which these reasons are expressed."[39]

The claim that having an understanding of the meaning of the terms in a statement is a necessary condition for evaluating the truth of that statement would seem to be incontrovertible. It would also seem unremarkable. But McPeck places a great emphasis on this condition, writing, "...it is this straightforward semantic dimension of the assessment of statements and arguments which I wish to stress as the most important, most difficult, and the most fruitful area to pursue for the development of critical thinking in any field."[40]

Yet, when McPeck sets out examples purporting to show the importance of understanding the meanings of concepts, he seems to shift his ground, arguing that it is the truth (as opposed to the meaning) of the statements that is primarily at issue.

McPeck makes a second uncontroversial point, that an understanding of the norms of logic are not sufficient for critical thinking, since those norms will not enable one to evaluate the truth of one's premises. McPeck correctly notes (maintaining the above noted shift) that "...the truth of the premises is every bit as important as the validity of the argument"[41]. He notes that often one's difficulty in evaluating the conclusion of an argument lies in "...determining the truth, not the validity, of various statements and putative evidence."[42]

Clearly, effective reasoning requires the development of a sound argument, not merely a valid argument (though argument validity would nevertheless remain a necessary condition for critical thinking).

B. Probabilistic Reasoning and the Evaluation of Knowledge Claims

Unfortunately, knowledge of the "truth" of the premises is notoriously difficult thing to come by. Even in a very simple argument--e.g.,

> (1) The deed was done, in the house, by some person, at some time;
> (2) The butler was alone in the house when the deed was done; therefore,
>
> ---
>
> (3) The butler did it,

we have no simple way to establish the truth of the premises. We must ask ourselves, Given the evidence at hand, how likely is it that this premise is true? It is seldom possible to simply "determine", in some clear, demonstrable fashion, the truth of a premise; instead one must make a judgment of probability with respect to its truth.

The need to make judgments of probability is especially apparent in the examples that McPeck uses. For instance, McPeck writes that in evaluating the radically different economic conclusions reached by Republicans and Democrats, "what I needed to know was whether the various premises were in fact true..."[43] But even the most knowledgeable and fully informed economist would be hard pressed to establish, definitively, that the premises of one or the other argument are "in fact true". Instead, a judgment of probability would be required.

McPeck recognizes the difficulties generally attached to finding the truth of the premises. Indeed, he explicitly faults those who interpret critical thinking as logical thinking for treating "knowledge" as simple and unproblematic, "more or less unambiguous, non-controversial and conceptually simple."[44] It appears to be his view, though, there are no general principles of reasoning that would be relevant to establishing the truth of the premises. One would simply have to acquire more "information".

Yet, an understanding of the fundamentals of probabilistic reasoning would give the student a basis for critically assessing the truth of the premises, i.e., estimating the probability that each particular statement is true, given the evidence for its truth.

C. Determining the significance of evidence

According to McPeck, one must recognize that "...some data...enjoy a much higher degree of certainty and reliability than others. All so-called data is not on an equal footing."[45] McPeck seems here to conflate the quite distinct notions of the "truth" of one's data and the "significance" of that data. One might easily imagine acquiring a piece of quite uncontestedly true information which would nevertheless be irrelevant to the problem at hand. Hence there are two distinct problems that the critical thinker must resolve. He or she must determine first, whether a premise is true, requiring a judgment of probability, and second, whether it matters that it is true. And, this second problem once again requires that one make a judgment of probability, this time of *conditional* probability. That is, one must consider whether the probability of the truth of the conclusion would be higher, or lower, or exactly the same, given the truth of the premise in question.

McPeck contends that in evaluating the significance of a piece of evidence, and in incorporating new information into one's existing store of accepted information, there are no general norms that might be brought to bear. He neglects to note that the norms of probabilistic reasoning do provide one with a useful content-neutral set of necessary conditions for good judgment in this task. These norms do not, of course, provide one with a mechanical decision process, but they do enable one to recognize and avoid clearly fallacious inferences. Such a fallacious inference, for instance, would be the belief that the probability of a hypothesis' truth, given certain evidence, is equal to the probability of the evidence given the truth of the hypothesis.

D. An Example

McPeck inadvertently provides an example of an argument in which it is the significance, not the truth, of a premise that is at issue. McPeck argues that understanding of the norms of logic is neither necessary nor sufficient for critical thinking. In attempting to demonstrate that logical norms are not necessary, McPeck gives us what appears to be a straightforward empirical observation: "That certain specific skills are *not necessary* for critical thinking is evidenced by the fact that many people can and do display critical thinking who have never been directly taught...the specific [logical] skills supposedly required of critical thinkers."[46] (This begs the question, since it assumes that those unable to evince an understanding of logical inference nevertheless could count as critical thinkers, but that is not the chief problem.)

The "evidence" McPeck advances in support of this assertion is of interest: He writes that "As the Watson-Glaser norm data shows, people with conventional liberal arts educations tend to score highest on their test; and there is little reason to believe that these people have been directly trained in any of these specific skills (e.g., the informal fallacies, etc.)."[47]

But, notice that the statement given as evidence, even if true, cannot be considered by McPeck to be significant, since elsewhere (2 pages earlier) McPeck argues that the Watson-Glaser test does not in fact measure critical thinking at all. Rather, the test, he writes, measures merely "...general scholastic ability, or *intelligence*..."[48]. Whatever it is that the test measures, according to McPeck, it is not "critical thinking".

McPeck errs, then, when he uses Watson-Glaser test-score data as "evidence" in support of his claim about critical thinkers, because, given his own argument, that data is irrelevant. Note that in this example, contrary to McPeck's thesis, the "truth or falsity" of the statement offered in evidence is not the point at issue--it may very well be true that the "liberally educated" score highest on the test. Let us suppose that it is true, and that the "facts" are exactly as McPeck states them. What we would still need to know to evaluate the argument, indeed, the key question, is, What is the significance of the piece of data given us, viz., the fact that those who lack logical skills score highest on the Watson Glaser test? I.e., given that information, what change would be warranted in our degree of belief in the truth of the conclusion, viz., that logical skills are not necessary for critical thinking? And this, of course, is a probabilistic question. It is clear, if we know that the test does not measure critical thinking, that the test-scores McPeck cites must be quite irrelevant to his thesis. The evidence has no bearing on the probability that the hypothesis is true.

While an understanding of the norms and pitfalls of probabilistic reasoning would not make the answer "easy", that understanding would enable one to avoid clearly fallacious inferences.

In McPeck's view, it is information, and the assessing of information, which is of paramount importance. In reasoning about public issues, he writes, "...the lion's share of the difficulty comes from the intelligibility and reliability of this information."[49] But, in the assessing of the significance of information, and in the incorporation of new evidence into a body of old knowledge, one is essentially involved the re-evaluation of the probabilities of outcomes, or the likelihood of hypotheses, given that new information. And thus that re-assessment that McPeck requires of critical thinkers is, in effect, a requirement that involves the thinker in probabilistic reasoning.

What is it that makes a piece of evidence or information "relevant" to a situation, and therefore significant to one's reasoning? It is that, as the result of one's acquisition of that new information, there is a change in the probabilities that exist (or are believed to exist) . The probability in question in such an evaluation would be interpreted as epistemic probability, the degree of confidence one is newly warranted in having in the belief that a particular assertion or hypothesis is true (e.g., one could be warranted in believing the proposition, "The butler did it" to various degrees, depending on the evidence one has). One would attempt to act "rationally" by revising one's estimate of the degree of confidence warranted in a hypothesis in accordance with the norms of probabilistic reasoning.

It should be noted here, in anticipation of the possible objection, that often one simply has no idea what the significance of new information is, and certainly cannot begin even rough "intuitive" assessments of probability. At such times, one would have no recourse but to admit that the new information cannot be usefully incorporated with the old. But, more generally, one can profitably ask such simple questions as "How likely would it be that the observed event would have occurred if my current hypothesis were right? And, even more importantly, "How likely would I be to observe this event, if my hypothesis were incorrect?" Only if such simple questions can be given an answer, however rough, would one be in any position to make an assessment of the changed likelihood that one's hypothesis is indeed correct. And one would need to recognize the "basic rules", for instance, that the probability one is interested in, the probability of the hypothesis given the evidence, Pr(H/e), is not simply equivalent to the probability of the evidence given the hypothesis, Pr(e/H).

Without, though, any explicit understanding of the relationship between new information and old, one is left to rely on purely guesswork estimates of the significance of information.

6. Critical Thinking as Rational Thinking

A. Siegel's Interpretation of Rationality

The view that the term 'critical thinking is best understood as "rational thinking" is advanced by Harvey Siegel, in Educating Reason, Rationality, Critical Thinking and Education. Siegel argues that "critical thinking" is coextensive with rationality, that critical thinking is nothing more or less than the "educational cognate" of rationality. To reach a full understanding of critical thinking then, we would need to understand rationality itself. Siegel acknowledges that at present we lack a fully elaborated theoretical account of rationality, but offers nonetheless a brief sketch. He interprets rationality as being "'coextensive with the relevance of reasons'"[50]; to be rational is simply "to believe and act on the basis of reasons". Accordingly, Siegel terms his conception of critical thinking the "appropriately moved by reasons" conception. The critical thinker, then, is the person who is rational, i.e., who "...appreciates and accepts the importance, and convicting force, of reasons."[51]

But, this characterization seems less than helpful when one sets out to teach critical thinking. In teaching critical thinking, we need not simply to establish a motivation, a disposition, the willingness to think critically--we need to show the motivated learner why it would be "appropriate" to be moved by certain reasons and not by others. To do so, we need to focus not on the desired characteristics of the learner, but on the characteristics of those reasons that the learner ought to be moved by. More generally, we need to focus on the criteria, the norms by which reasons in general are to be judged as significant, or conclusive, or worthless.

1. Reasons and Principles

Clearly then, if one is to understand the twin concepts, "rationality and "critical thinking", one must have an understanding of the notion of "reasons". Siegel gives some indication of what would, and would not, count as a reason. According to Siegel, "rational thinking" is also equivalent to "principled thinking", since a statement does not count as a person's "reason" for an action (in the justificatory sense) unless that person is committed to consistently accepting that statement as a reason in relevantly similar circumstances. A reason is thus a particular instance of the application of a principle, according to Siegel.[52]

The term 'critical thinking' thus shifts slightly in its meaning, from (a) "being appropriately moved by reasons" to

(b) acting on the basis of general principles. Such general principles, Siegel writes, alone can provide "genuine" reasons for action. Critical thinking is thus "...principled thinking... impartial, consistent and non-arbitrary..."[53] It "...presupposes a recognition of the binding force of standards, taken to be universal and objective, in accordance with which judgments are to be made."[54]

This could be seen as a problem for philosophical pragmatists, were it not for the fact that Siegel almost immediately shifts his ground. He writes that because the critical thinker "...must be able to assess reasons and their ability to warrant beliefs...the critical thinker must have a good understanding of, and the ability to utilize, *principles governing the assessment of reasons*."[55] (emphasis added). The potential problem is thus dissolved, since even a pragmatist can agree that there are <u>some</u> universal principles applicable to reason assessment--the principles of logic, for instance, the principle of non-contradiction, and even principles of probabilistic reasoning.

But note that this second claim is a far cry from the original claim, that <u>all</u> "genuine" reasons must necessarily be derived from and "backed by" some universal principle.

2. Subject-specific Principles

Siegel recognizes two categories of "principles" by which potential reasons may be assessed: first, subject-neutral principles, e.g., logical principles, either formal or informal; and second, subject-specific principles. The first category is clear enough, but the latter category is somewhat obscure. Siegel gives as examples "...principles governing the proper assessment of works of art, or novels, or historical documents, or the design of bathroom fixtures..."[56], and asserts that "there is no a *priori* reason for regarding either of these types of principles as more basic (or relevant) to critical thinking than the other..."[57]

There is, though, one very significant difference between the two sorts of "principles". The principles of logic are conceived of as being necessarily true, axiomatic, while the subject-specific "principles" are merely substantive generalizations, drawn by quite fallible individuals from previous experience. While such generalizations are intended to guide practice, they are not held to be necessarily true.

Siegel thus allows "subject-specific principles" into the ranks of principles that could serve as "genuine reasons".

This move allows Siegel to accommodate McPeck's argument, that an exclusive focus in critical thinking programs on subject-neutral principles of logic is unjustified. Yet, while it is unarguable that subject-specific information certainly is required if one is to evaluate the probability that a premise in an argument is true, McPeck's claim that there exist "subject-specific logical principles" is extremely dubious (a point discussed above).

Moreover, although it is undeniable that subject-specific standards do exist, and that these are used both to guide practice and to evaluate evidence, those standards themselves may be arbitrarily set up and theoretically unfounded. For example, in psychology, the use of "significance tests" in experimental analysis has long been a standard requirement for publication, yet that requirement, according to Gigerenzer et al., is largely the product of "wishful thinking", and of "attractive illusions" about the meaning of 'significance'. Such illusions, according to Gigerenzer, persist because of the "...neglect of controversial issues and alternative theories and the anonymous presentation of an apparently monolithic body of statistical techniques."[58] Siegel errs, it would seem, in setting such potentially incorrect "subject-specific standards" on a par with axiomatic logical truths, and in taking the adherence to such standards as a criterion for critical thinking.

3. An Epistemological Requirement

Siegel adds a further requirement that must be met if a person is to qualify as a critical thinker. "...critical thinkers need also to have a theoretical grasp of the nature of reasons, warrant, and justification,...that is, the reason assessment component [of critical thinking] involves *epistemology*."[59] One must question, though, whether this can be considered to be a logically necessary condition for critical thinking. Consider the implications. If a theoretical grasp of the nature of reasons, warrant and justification is a necessary condition for critical thinking, then it would be logically impossible to "think critically" without first having such a grasp. But one should note, first, that there is at present substantial disagreement, even among (and perhaps especially among) epistemologists as to "the" nature of reasons, warrant and justification; there are numerous theoretical accounts, and there is no clear winner among them.

The interpretation of the requirement itself is problematic. What does it mean for one to "have a theoretical grasp of" something? It seems unlikely that Siegel intends to require merely that one hold <u>some</u> theoretical account. A more plausible interpretation is that one is required to have "the best" theoretical account, or at least a well justified, coherent account. But if so, one would need to be in a position to critically <u>evaluate</u> the strengths and weaknesses of many and various theoretical accounts that exist, in order to choose one.

But, if <u>this</u> is the meaning of 'having a grasp', one logically never could arrive at a critical evaluation of incompatible epistemological theories, given that "critical thinking" itself is held to be impossible unless one already <u>has</u> such an account. Thus, the consequence of accepting the "epistemological condition" as a necessary condition for critical thinking would be that critical thinking is logically impossible. It would seem that this requirement is too strong.

It <u>does</u> seem, though, that in the "epistemological condition" Siegel has set out a significant desideratum for the "well-educated person", viz., that the person be at least acquainted with, and one would hope, <u>critical</u> with respect to the fundamental philosophical questions about the nature of reasoning, of knowledge, justification, etc..

B. McPeck's Interpretation of Rationality

McPeck, like Siegel, interprets critical thinking as being closely related to "rationality". Indeed, Siegel refers to this coincidence as indicative of an area of fundamental agreement between himself and McPeck. However, any such agreement would is illusory, a purely verbal agreement, since Siegel and McPeck differ significantly in their interpretation of the key term here, viz., 'rationality'.

McPeck, unlike Siegel, maintains that critical thinking is a subset of rational thinking. Rational thinking, he writes, is "...the intelligent use of all available evidence for the solution of some problem." Critical thinking, by contrast, is said to occur only when one must resolve some "problem" in the use of that available evidence. McPeck gives, as examples of such problems, the need to decide what is to "count" as evidence, or the decision to "disregard" some portion of the available evidence. Critical thinking is simply the "*disposition and skill* to find such difficulties in the normal course of reasoning." [60] (emphasis added)

It is unclear here what McPeck conceives as the "normal course of reasoning". It would seem that whenever one sets out to employ available evidence, in order to act rationally one would have to evaluate the purported "evidence", determine its significance (over and above its truth), and decide whether or not, in this problem, this information should count as evidence.

"Critical" thinking may indeed equal rational thinking, but, if so, rational thinking cannot be construed merely as a simple "reliance on reasons" to establish belief; one must also rely on a set of reasons that are the "right" reasons, on reasons that do in fact lend support to the conclusion. I.e., the critical thinker is not simply one who appeals to reasons to warrant beliefs, but one whose appeal to reasons conforms to certain norms of "good" reasoning. What the critical thinker needs is some understanding of what sort of norms are to be applied, and how to go about the process of evaluating reasons (as Siegel requires) and/or evidence (as McPeck requires).

The study of probabilistic reasoning, while not by itself sufficient to make the student a "critical thinker", would give the student access to a useful set of general criteria by which he or she could begin to critically evaluate a wide range of reasons and/or evidence. *Pace* McPeck, such criteria would be applicable across a number of disciplines, as will be argued later.

C. Rationality & Induction[61]

When one asks "which reasons ought I to be moved by?", one is forced to consider the relationship of induction, and hence of probability, to rationality. For if the only reasons that are to move one are reasons that logically entail the truth of one's conclusion, i.e., reasons that make up a deductive argument, one would seem to be faced with a very limited set of legitimate "reasons", and an equally limited set of "rational" actions. One would have to forego reliance on ampliative arguments, in which the conclusion goes beyond the information in the premises, and hence be "irrational" when making predictions, generalizations, categorizations. Indeed, in most everyday problems one would then be reduced to acting "irrationally", by definition.

If, on the other hand, we begin with the observation that, whatever the full explication of rationality, we do at least intend that some actions taken on the basis of non-demonstrative arguments will be counted as rational, then inductive arguments, with conclusions based on and incorporating judgments of probability, will have to be accepted as potentially appropriate reasons. Those reasons are acceptable provided, of course, that the inductive reasoning in the argument conforms to a set of appropriate norms. And the problem becomes to determine what those norms are. As Max Black puts it, "Anybody who aspires to rationality must be guided by probabilities in the face of uncertainty: how this is to be done and with what justification are the main themes of the philosophy of probability."[62]

It is not my intention here to attempt to grapple with the philosophical problem of the general justification of induction, an issue that dates back to Hume's treatment in A Treatise on Human Nature,[63] and is still problematic today. I merely note, first, that in the context of everyday problem-solving the elements of chance and uncertainty are ineluctable, and hence the requirement that judgments of probability be made is inescapable. And, second, that it seems incontestable that the *sine qua non*, and the least possible constraint, to set upon these judgments that are to count as "good reasons" for action or belief would be that they conform to the axioms of the probability calculus.

The problem is that those axioms are neither intuitively obvious nor in any sense psychologically enscribed. We seem to lack any *lume naturale* with respect to the probability calculus; moreover, often the correct probabilistic reasoning seems strongly counter-intuitive. Hence, if the goal of a program in critical thinking is construed as the development of the students' rationality in everyday problems, that program should provide instruction in the norms of probabilistic reasoning.[64]

Summary

In this chapter I have set out and examined the principal interpretations of the term 'critical thinking'. Several main points emerged. First, the term is often used in a very vague sense, as a term of approbation, to indicate no more than that thinking which is most valued, or, "good thinking". This sense, though, is unhelpful when designing a curriculum to teach critical thinking, and a more complete explication of what counts as critical thinking is required.

Second, the term is often used to stress a particular desirable attitude and/or disposition that is to accompany thinking. Both Siegel and McPeck include the attitudinal/dispositional aspect of critical thinking. I have argued that the dispositional requirement is problematic, since it would not permit any single act of thinking to count as "critical thinking".

The attitudinal requirement, which can be linked to particular acts of critical thinking, is interpreted differently by Siegel and McPeck. According to McPeck, the required attitude is a "judicious, reflective skepticism", which impels the thinker to question the norms of thinking accepted in particular disciplines. According to Siegel, the attitude required is simply the willingness to engage in "reason assessment".

I have argued that, in "everyday reasoning", the attitude appropriately associated with critical thinking is one of self-criticism, or self-evaluation, of one's own thinking. Critical thinking thus involves the informed selection of a set of justified norms of thinking, and the conscious application of those norms to one's thinking about the problem at hand. Through this critical process one can one hope to generally avoid the systematic biases, cognitive illusions, and faulty but compelling "intuitions" that might otherwise lead one into error. I will discuss in the next chapter the rather large body of evidence in the psychological literature purporting to show the existence of precisely such biases, illusions and flawed intuitions, which, it is claimed, regularly mar probabilistic reasoning. The study of probabilistic reasoning would be necessary if the student is to develop the "critical" attitude with respect to such everyday problems.

Third, the term is sometimes used as synonymous with "skillful" thinking. I have considered several possible interpretations of the notion of "skill", and have concluded that, at least with respect to thinking, a skillful practice is a practice that demonstrates knowledge of, and application of, a set of norms appropriate to the practice. To teach critical thinking in everyday problems, then, is to teach such a set of norms.

Fourth, 'critical thinking' is sometimes equated with "logical thinking", interpreted as thinking that conforms to the norms of deductive and/or informal logic and argument analysis. McPeck has criticized this view, largely on the grounds that the rules of "classical logic" are a) not applicable across disparate fields of inquiry and b) even if applicable, not transferrable across disciplines (a psychological claim). I have argued that McPeck is incorrect in these claims, but that the interpretation of critical thinking as logical thinking is incomplete unless explicit attention is given to the norms of *inductive* logic. Though there is at present no complete explication of what those norms might be, it seems reasonable to expect that, given the probabilistic nature of inductive logic, the norms of probabilistic reasoning would be fundamental. Hence, when critical thinking is interpreted as logical thinking, it would be appropriate to teach probabilistic reasoning in the critical thinking course.

Fifth, 'critical thinking' is interpreted by McPeck as the evaluation of the truth of knowledge claims, particularly, the truth of the premises in one's argument. I have agreed, in part, with this interpretation, but have pointed out that that evaluation itself requires a judgment of probability. Moreover, I have argued that in everyday problems it is not only the truth, but the significance of the (true) premises that must be evaluated. And this also requires a judgment of probability, viz., that the probability of the truth of the conclusion is greater, given the truth of the premise, than otherwise would be the case. So, if critical thinking is interpreted as the evaluation of knowledge claims, the teaching of the norms of probabilistic reasoning would be appropriate.

And sixth, the term 'critical thinking' is often interpreted as closely related to "rational thinking". According to Siegel, the two terms are co-extensive, and to teach critical thinking is to teach students to be rational human beings. If this is true, then one must have some explication of what it means to be "rational", in order to know what is to be included in the critical thinking course. In order to count as "rational", Siegel requires that students become aware of the need in reasoning for appeals to fundamental principles; he requires that genuine reasons be "principled", though he also allows for appeals to "subject-specific principles"; and he requires that students become engaged with epistemological questions about the nature of knowledge per se. The study of the theory of probability would involve students in the sort of fundamental questions Siegel seems to require for rationality. It would also provide a set of norms which, contra McPeck, would apply broadly across the disciplines.

McPeck maintains that "critical thinking" occurs only when ordinary rational thinking (thinking in accordance with the accepted norms of the disciplines) somehow fails, and one must begin to question the accepted norms. The critical thinker thus goes beyond a mere rationality. But, as will be argued later, the norms of probabilistic reasoning are at present quite controversial. To "learn the norms of probabilistic reasoning" it would not be sufficient, nor possible, to merely memorize, accept and apply a single set of rules. Rather, the student must begin to evaluate the merits of the proposed sets of norms, and reach justified judgments about the conditions under which the norms are properly applied. Hence the study of probabilistic reasoning would be conducive to the sort of "critical thinking" that McPeck envisions.

In short, then, given any of the major interpretations of the term 'critical thinking', the study of probabilistic reasoning would be an appropriate component in the critical thinking curriculum. Throughout this discussion, so far, I have relied on the assumptions a) that probabilistic reasoning is basic to what is generally thought of as "everyday reasoning", and b) that probabilistic reasoning stands in need of improvement. In the next chapter, I take up these two points.

1. John E. McPeck, Teaching Critical Thinking: Dialogue and Dialectic (New York: Routledge, 1990), p. 34.

2. Ibid., p. 7.

3. Ibid., p. 42.

4. John E. McPeck, Critical Thinking and Education (New York: St. Martin's Press, 1981), p. 9.

5. Ibid.

6. John E. McPeck, Teaching Critical Thinking: Dialogue and Dialectic, p. 41.

7. Ibid., p. 41.

8. Ibid., p. 42.

9. Harvey Siegel, "McPeck, informal logic, and the nature of critical thinking," in McPeck, Teaching Critical Thinking: Dialogue and Dialectic, p. 79.

10. Ibid.

11. Virtually every writer on "thinking" seems at some point, to some extent, to make use of this metaphor.

12. Steven P. Norris, "Thinking about critical thinking: philosophers can't go it alone", in McPeck, Teaching Critical Thinking: Dialogue and Dialectic (New York: Routledge, 1990), p. 68.

13. Ibid., p. 69.

14. Ronald E. Charlton, "The Direct Teaching of Analysis", in Thinking Skills Instruction: Concepts and Techniques, ed. by Marcia Heiman and Joshua Slomianko (Washington, D.C.:National Education Association, 1987), p. 152.

15. Sir Frederic Bartlett, Thinking: An Experimental and Social Study (New York: Basic Books, Inc., Publishers, 1958).

16. McPeck, Teaching Critical Thinking, p. 5.

17. Ibid., p. 19.

18. Ibid., p. xiv.

19. Ibid., p. 4.

20. Ibid., p. 20.

21. Ibid., p. 36.

22. Joseph S. Karmos and Ann H. Karmos, "Strategies for active Involvement in Problem-solving, in Thinking Skills Instruction: Concepts and Techniques, ed. by Marcia Heiman and Joshua Slomianko (Washington, D.C.: National Education Association, 1987), pp. 99-110.

23. Edward De Bono, Teaching Thinking (London: Temple Smith, 1976).

24. McPeck, Teaching Critical Thinking, p. 6.

25. Ibid., p. 16.

26. Ibid., p. 16.

27. Ibid., p. 30.

28. Ibid., p. 36.

29. Ibid., p. 36.

30. Ibid., p. 36.

31. Ibid., p. 37.

32. Ibid., p. 26.

33. Ibid., p. 129.

34. Ibid., p. 9.

35. John McPeck, Critical Thinking and Education.

36. John McPeck, Teaching Critical Thinking: Dialogue and Dialectic.

37. Ibid., p. 28.

38. John McPeck, Critical Thinking and Education, p. 9.

39. John McPeck, "Critical thinking without logic: Restoring dignity to information", Proceedings of the Thirty-seventh Annual Meeting of the Philosophy of Education Society (Philosophy of Education Society, 1981), p. 220.

40. Ibid.

41. John McPeck, Teaching Critical Thinking, p. 7.

42. Ibid., p. 5.

43. Ibid., p. 10.

44. Ibid., p. 27.

45. Ibid., p. 28.

46. Ibid., p. 25.

47. Ibid.

48. Ibid., p. 23.

49. Ibid., 9.

50. Harvey Siegel, Educating Reason: Rationality, Critical Thinking and Education (New York: Routledge, 1988), p. 33, quoting Israel Scheffler, Conditions of Knowledge: An Introduction to Epistemology and Education (Chicago: The University of Chicago Press, 1965), p. 106.

51. Siegel, Educating Reason, p. 33.

52. Ibid., pp. 33, 34.

53. Ibid., p. 43.

54. Ibid., p. 34.

55. Ibid.

56. Ibid., p. 35.

57. Ibid.

58. Gerd Gigerenzer et al, The Empire of Chance: How Probability Changed Science and Everyday Life, p. 209.

59. Siegel, Educating Reason, p. 35.

60. John E. McPeck, Critical Thinking and Education, p. 12.

61. L. Jonathan Cohen, An Introduction to the Philosophy of Induction and Probability (Oxford: Clarendon Press, 1989).

62. Max Black, Margins of Precision: Essays in Logic and Language (Ithaca, N.Y.: Cornell University Press, 1970), p. 92.

63. David Hume, A Treatise on Human Nature, Book 1, Part 3.

64. There is yet another interpretation of 'rationality' that is directly related to the notion of probability. One may be considered a "rational" person provided only that one does not hold a set of beliefs that actually guarantees that one will fail to maximize one's values, i.e., one is rational if and only if "Dutch Book" cannot be made against one.

CHAPTER III

PROBABILISTIC REASONING AND EVERYDAY REASONING

1. Introduction

In this chapter I will, first, discuss the concept of probabilistic reasoning, and second, examine the use of probabilistic reasoning in problems of everyday life; e.g., in the processes of estimating, judging, concept formation, and decision-making, hypothesis evaluation. In the final section I will set out and discuss some of the evidence from social psychology concerning probabilistic thinking.

2. Probabilistic Reasoning and Probability

"Probabilistic reasoning" is the sort of reasoning required whenever a problem to be solved involves an element of uncertainty, either because of the role of chance with respect to the problem, or because the problem is such that all the necessary information either is not, or cannot be known. In such problems, the conclusions one draws cannot be certain, but only in some degree probable. Clearly, this is an extensive general category--it is not surprising that most problems one encounters would fit into this category to at least some extent.

One might take as a "model" case of probabilistic reasoning any reasoning involving statistical analysis. It is not my intention, however, to suggest that courses in statistics and data analysis should form part of the critical thinking curriculum. While statistical reasoning would seem to be a clear, undisputed case of probabilistic reasoning, such reasoning constitutes just a "sliver" of a much broader category.

A second, and perhaps even the most obvious model case of probabilistic reasoning would be reasoning which overtly involves the manipulation of numbers expressing degrees of probability. For example, in describing the likelihood of an event, we often are able to give an approximate, and sometimes an exact number, e.g., "There is one chance in a million that I will win the lottery". In this case the estimation of the chances involves a straightforward, classical or "indifference theory"[1] interpretation of probability. One simply notes the number of possible outcomes (any of one million numbers <u>could</u> win), and the number of favorable outcomes (<u>my</u> number could win), and then notes the ratio between the two numbers, (1 : 1 million).

The problem is simple since in this case it seems reasonable to assume that each number has an equal chance of coming up.[2] Each outcome is assumed to have an equal probability; one reasons either that a) there is no cause in operation that would make one outcome any more likely than the others (a realist, objective interpretation), or b) that one does not know of any cause in operation that would affect the likelihood of the outcome (an idealist interpretation).

Although this interpretation of probability seems simple and intuitively obvious, is now considered to be fatally flawed, since it relies on the Principle of Indifference, and this principle can lead to inconsistent, self-contradictory conclusions about the probability of logically equivalent statements.

One might, however, take judgments of probability to indicate the relative frequency of individuals or of characteristics within a class, rather than the ratio of favorable outcomes to the total number of possible outcomes. This difference in interpretation of the meaning of probability can have dramatic effects on one's probability judgments. For instance, suppose one wants to determine the "probability" that a student has a particular professor, Professor A, as his or her adviser. Adopting the indifference interpretation, one would count the number of advisors in the department (there are three), and compare the ratio of favorable outcomes (just one) to possible outcomes, determining the p(A) to be 1/3. If, on the other hand, one were to adopt a frequency interpretation of probability, one would look at the total number of students in the department (one thousand), and at the number of students assigned to Professor A (100), and determine the ratio of students assigned to Professor A to the total number of students in the department. The p(A), on this interpretation of probability, would be 1/10, not 1/3.

But, one might then reflect that the reference class used, "all students in the department", is not the most appropriate reference class, since the particular student in question may differ in many ways from the majority, for instance, in having an unusual set of interests, shared by only 20 students. One might then determine the number of students out of the subclass, "students having similar interests" that have been assigned to Professor A, say, 15. The probability then would be 3/4 (not 1/10, nor 1/3). But, that subclass itself could conceivably be divided, and redivided, until only the particular student in question would remain as a member of the relevant class. The probability then would be 100%, or 0%, depending on whether the student was or was not assigned to Professor A. So, even in this very simple example, it is clear that the estimation of probability is no simple matter. One's estimate is directly influenced by the theoretical, conceptual assumptions that are accepted.

But, still other interpretations of probability and hence judgments of probability, are possible. One might, in advance of a student's choice, ask, What is the probability that this unique student will <u>choose</u> Professor A? One would in this case be switching to a "propensity" interpretation of probability, asking whether there are character traits or situational factors that might exist that would <u>cause</u> this particular student to tend to make one choice rather than another. In this example, though, this interpretation would seem unhelpful, since the required causal factors would be difficult to identify even if they should exist. In other contexts, though, the propensity interpretation would be useful--e.g., knowing that high blood cholesterol causes coronary artery disease, one might conclude that all individuals with high blood cholesterol have a propensity to develop such a disease, and hence have a high probability of having done so.

Still other sorts of probability judgments are possible. One might, for instance, observe that "there is a 50% chance of rain tomorrow", or, perhaps, that there is a 10% chance that I will finish this paper tonight. In this case, the assessment of probability is not quite as simple as above, since there is no obvious, well defined set of "possible outcomes" to be compared to "favorable outcomes", nor is there a clearly defined reference class that would allow a "relative frequency" judgment. In this sort of problem, one would seem to be making a statement of <u>epistemic</u> probability. That is, in making the "judgment" of probability, one is simply noting the degree of confidence that one has, or, alternatively, that would be appropriate for one to have, in the truth of the two statements, "It will rain tomorrow" and "I will finish this paper tonight". If the statements were <u>certainly</u> true, or <u>certainly</u> false, their epistemic probabilities would be 1.0, and 0.0, respectively. Since one is not certain, one needs to make a judgment of the probability that the statements are true or false. This is a personalist, "Ramsey-type", or Bayesian interpretation of the meaning of "probability". One might take the statement of probability to indicate a) the degree of belief in the truth of a proposition that a person actually holds; b) the degree of <u>rational</u> belief actually held by a person; or, c) the degree of rational belief that a person ought to be holding.

The first interpretation of probability, the "degree of actual belief", is untenable, since actual degrees of belief would not necessarily conform to the probability calculus. The second interpretation escapes this problem, since "rational" degrees of belief are simply those that do so conform--this would be a subjective idealist interpretation of probability. The third, normative or prescriptive interpretation of probability statements would be an objective idealist interpretation.[3]

In making this sort of judgment, one might proceed simply by introspection, noting the confidence that one actually feels, and translating that feeling of certitude into a number. This would be a purely subjective probability. Such a move has an advantage in that the required number is relatively easy to come by, but it also has several disadvantages. Such a set of purely subjective probabilities may be self-contradictory, i.e., it may not conform to the strictures of the probability calculus. Moreover, if one's degrees of belief were arbitrary, they would not serve one well as a "guide to life".

One could, however, attempt to determine one's _rational_ degree of belief in the proposition, i.e., attempt to make a probability judgment that is not inconsistent with one's other judgments and beliefs about the world, and so seek to establish a coherent set of beliefs which would in fact conform to the probability calculus.

In this example, having determined one's rational degree of belief in the two propositions, one might then ask, Given that "sunshine" and "having finished the paper" are the two determining factors, how likely is it that I will be attending tomorrow's picnic? To reason successfully about such obviously quantitative probabilistic matters one requires not only a conceptual understanding of probability, but also a knowledge of the fairly simple rules of the probability calculus. For instance, in this calculation a common error would be to "split the difference", to conclude that the probability of attendance is about 30%, or perhaps to add the two probabilities, concluding that the probability is about 60%, rather than to multiply the two numbers to conclude correctly that, given these two factors the probability of my attending a picnic tomorrow is a mere 5%.

What would appear to be relatively simple, straightforward calculations of probability arise in many everyday problems-

-a particularly common instance would be in the assessment of risk. For example, one might like to travel to Europe, but wonder about deferring the trip because of the danger of terrorism. To evaluate the risk associated with that travel, one would need to know, first, the relevant numbers (e.g., in 1980, 17 deaths, out of approximately 28 million travelers), and, second, the appropriate procedures for the manipulation of the relevant probabilities.

There is, in addition, a third requirement for successful reasoning in this sort of problem. One ought to have not only an understanding of how to perform the required calculation, but also an understanding of the import of the number reached as a result of that calculation. Even supposing that the numbers reached accurately describe the probability of the event, such a number is useless as a guide to practice unless one has some notion of its significance. One might, for instance, compare the risk of dying in a terrorist attack to the risk of dying in an accident at home, or the risk of being killed in an automobile accident. The probability of either of the latter two events is much higher than that of being killed by terrorists while abroad, yet neither is considered particularly "likely". This suggests an important point in probabilistic reasoning--a mere manipulation of numbers does not exhaust the realm of probabilistic reasoning. One needs not only to assign or determine the relevant probabilities, but to make judgments about matters that are essentially probabilistic.

This sort of probabilistic reasoning, whether it is done well or poorly, can have profound social consequences. Consider, for instance, the current proposals to require the mandatory testing of all medical professionals for the HIV virus, to prevent transmission of the virus to patients during treatment. While it is indeed possible to become infected with the HIV virus in this way, the probability of such an event, given the evidence at hand, is minuscule. To date, out of all the patient/medical personnel interactions, there are only 5 known cases of HIV transmission from care-giver to patient.

While one might argue that the information at hand is inadequate, and hence that further research into such a route of transmission is needed, one ought not to adopt the "argument from ignorance", that since it hasn't been shown for sure that the risk is not significant, one may conclude that it is significant. This inference would be as fallacious in probabilistic reasoning as in any other argument.

3. Probability and Induction

As noted above, there is a close relationship between probabilistic reasoning and induction. All inductive arguments, by definition, involve an element of probability, since an inductive argument is ampliative; it is simply one whose conclusion does not follow with certainty from the premises.

It seems useful, however, to draw a distinction between the two concepts. "Probabilistic reasoning" is reasoning that attempts to evaluate either the truth of the premises in the argument, or, the strength and reliability of the inductive argument itself, i.e., the "probability" of the conclusion given the premises. That is, recognizing that the conclusion of an inductive argument is not certain, it becomes appropriate to ask, "Just <u>how</u> probable <u>is</u> it, that...?". To answer this sort of question, one needs to be aware of the factors involved in making judgments of probability, the meaning of such judgments, and the norms by which judgments of probability are evaluated.

The construction and evaluation of inductive arguments is a crucial part of "everyday reasoning". Through inductive reasoning one develops expectations about the future, and makes plans of action. Later, after having acted upon those plans, one observes the results and on the basis of that evidence re-evaluates one's initial expectations. In that re-evaluation it is critically important to appreciate that the element of chance in such an argument is inescapable, and that even the strongest of inductive arguments cannot provide one with certainty. So, simply discovering that the conclusion one reached was, in the event, false, does not mean that the reasoning that led to the conclusion was in any way flawed.

Moreover, one would like to be in a position to refine somewhat one's assessments of the probable truth of the conclusion of an inductive argument. Lacking any understanding of the estimation of conditional probabilities, one is effectively limited to simply distinguishing between the two "extreme" values, i.e., one would be limited to the dichotomy, "certainty" and "uncertainty", with "uncertainty" construed as "a 50:50 chance". For instance, one often hears "You don't <u>know</u> that that will happen--it might, it might not." One ought to go on, however, to ask, "How likely is it to happen, or not to happen?".

<u>4. Probability and Deduction</u>

The obvious difference between deductive and inductive reasoning is that in deduction the truth of one's conclusion follows with certainty from the truth of one's premises. But, it is important to note that the initial set of premises in a deductive argument is simply asserted to be true, i.e., taken to be true by assumption. But, in everyday reasoning one seldom if ever has the luxury of having a set of premises in which one is warranted in having absolute confidence. Instead, the premises in one's arguments are themselves only "probably" true.

One could, of course, insist on restricting that initial set of premises to statements that are tautologous or logically necessary. But, in that case, one's arguments could not have any real bearing on those everyday problems with which, by hypothesis, one is concerned. In order to draw conclusions that are empirically significant, one must have premises that are themselves empirically, not logically, true.

Hence, even if one were to scrupulously restrict one's reasoning in everyday problems to the development of deductively valid arguments, one nevertheless could not completely exscind the element of probability from one's arguments. One needs to assess the truth of premises which are not analytically true, and if this judgment is to count as a judgment, rather than as a mere guess, one must have an understanding of the appropriate criteria of judgment, viz., the norms of probabilistic reasoning.

5. Predictions, Generalizations and Causal Attributions

While there are a great number of applications of probabilistic reasoning in every-day decision-making, I will concentrate here on its occurrence in three inter-related areas: first, the making of predictions; second, in the making of causal attributions; and third, in the making of synchronic generalizations. The need for judgments of probability in these and other areas has been explored extensively by Kahneman and Tversky, Nisbett and Ross, and others in the literature of psychology and sociology.[4]

A. Predictions

Predictions are judgments about the likely occurrence of future events, or about the future behavior of individuals or groups, based on some evidence at hand. These judgments are essentially probabilistic. Some events may be thought to be "perfectly predictable"; that is, the probability of the event given the evidence may be judged to be 100%; other events may be random and unpredictable, i.e., occurring entirely by chance, e.g., the decay of a particular radioactive particle. Most events would fall somewhere in between these two extremes, i.e., the event would be determined <u>in part</u> by known factors or causes, and in part by the chance or by the operation of some as yet unknown cause. The degree of confidence that is warranted in a prediction depends largely on the degree to which chance plays a part in determining the event in question. A prediction, moreover, is based on one's observation of the available evidence, and chance, again, may have a significant effect in determining precisely which pieces of evidence were in fact observed.

In making a prediction, two distinct sorts of information are pertinent, viz., first, the distributional or base-rate information, and, second, the singular or individuating information.

Suppose, for example, that one would like to predict the probable success of an instructional program in a particular setting. In the absence of <u>any</u> specific information about the operation of the program in that particular circumstance, one's best choice would be to anticipate the average achievement for that sort of program, and to make the <u>same</u> prediction in each case. That is, in the absence of individuating data, a prediction of the mean value would be appropriate, since that prediction is the one that would most often be most nearly correct.

In this case, the only available basis for the prediction in this case is the distributional data, i.e., information about the distribution of the values in question within the appropriate reference class. This information gives one the "prior probability", the appropriate basis for judgment in the absence of any individuating data.

Suppose, however, that one does have some relevant information about the particular situation under consideration; this is the singular or individuating information.[5] If it were the case that the outcome were perfectly predictable on the basis of this evidence, then only this singular information need be considered; if, on the other hand, the final outcomes were perfectly <u>random</u>, then only the distributional information should be used. If, however, the outcome is partially predictable on the basis of the singular information, then <u>both</u> sorts of information should be used in one's reasoning. I.e., the conclusion that one would reach by relying solely on the singular information should be "tempered" in some way by one's knowledge of what that prediction would be if one relied solely on the distributional data.

Consider, for example, a common medical problem--making a judgment about the probability of disease, given certain evidence. Suppose it is known, for instance, that a given disease occurs in about one-tenth of one percent (.1%) of the general population. The probability that any randomly chosen person X in the population will have that disease is, simply, .1%. This is the "prior probability", the probability in the absence of any information about the particular individual. But suppose it is also known that the person had not been randomly chosen, but had been chosen from a select group, for instance, persons who have tested positive for the disease. That test result is the individuating data for person X. How likely is it then, given the individuating data, that person X has the disease?

It is clear that if the test for the disease were a perfect predictor, i.e., if the test never gave false positives, then it would be 100% probable that X has the disease, given the evidence of a positive test result. I.e., the probability of the disease given the positive test result is 100%, i.e., Pr(disease/+test) = 100%.

In such a case, one could safely rely on that data exclusively, and begin vigorous treatment of the disease.

Unfortunately, few if any tests are that conclusive--a typical test would be expected to generate a certain number of false readings, both false positives and false negatives. Suppose that, in looking back among those who are later found to have actually had the disease, one discovers that the test gave positive readings for the disease 90% of the time. That is, the probability of a positive reading, given that the subjects <u>have</u> the disease, is 90%, i.e.,

Pr(<u>+test</u>/disease) = 90%.

The test could be said to be 90% accurate, in the sense that it gives "true positives" 90% of the time. But, one needs some way to move from this evidence to the desired conclusion, viz., Pr(<u>disease</u>/+test).

In this sort of typical case, then, the probabilistic reasoner must employ <u>both</u> the distributional <u>and</u> the singular data, in combination. The question, of course, is, How? As in the very simple case of the probability of picnic attendance, there are a number of possible "moves" that might seem quite plausible and tempting, but would be incorrect. For instance, one might choose, again, to "split the difference", estimating the probability that one has the disease, given the positive test, to be a value about half-way between the distributional and the singular probabilities, i.e., about 50%. Or, one might perhaps be inclined to "weight" one sort of information over the other in some way. Or, one might simply abandon the distributional data entirely, concluding that the probability of the disease given the test result is the same as the probability of the test result given the disease, i.e., that if the

Pr(+test/disease) = 90%, then the

Pr(disease/+test) = 90%.

Each of these inferences, though, would be unwarranted; each would constitute a probabilistic reasoning fallacy. The appropriate way to estimate the probability of the disease is given in Bayes' formula. The use and misuse of Bayesian reasoning will be explored in greater detail later--it will suffice to point out here the rather surprising fact that, in this case, given the data available, no conclusion at all can be validly drawn concerning the probability that X has the disease. It would seem that, without having had some explicit instruction in the norms of probabilistic reasoning, few would recognize intuitively that this problem cannot be solved, that vital information is missing.

It should be noted that this sort of difficulty is the rule, not the exception. One could, for instance, predict the probability that a couple, each of whom is known to be heterozygous for a particular disease inducing gene, will each pass that gene to their offspring, who will be homozygous for that gene, i.e., 1/2 * 1/2 = 1/4. The probability in question here would be the straightforward "indifference theory" interpretation[6]; one would reason that no cause exists (or at least no cause is known to exist) that would make any one of the four possible outcomes more likely than any other. One could go on to assess the probability that, out of a given number of children, no child born to the couple will be homozygous for the trait, i.e., for 3 children,

Pr(no homozygous children/parents are heterozygous)

= 3/4 * 3/4 * 3/4 = 1/7.

This fairly simple sort of calculation of probability would be appropriate, indeed, would be possible, if and only if be one knows a) the number of the possible outcomes, b) the probability of each possible outcome, and c) the number of "throws of the dice", i.e., the number of independent chance occurrences, that are at issue.[7]

It is a far different and more difficult matter to assess, for instance, the probability that one who has a high level of blood cholesterol will develop heart disease, given information about the incidence of high cholesterol among those who have in fact developed heart disease. One might learn, for instance, that, of those later developing heart disease, four out of five had cholesterol levels greater than 200 mg/dl. But it does not follow from this that four out of five of those having cholesterol levels greater than 200 mg/dl will later develop heart disease.

(It should be noted here that the use of "Bayesian reasoning" is itself controversial--an issue to be explored at some length later.)

B. Attributions of Causality

Probabilistic reasoning also plays a part in everyday attributions of causality, in several ways. First, the probabilistic factors of chance and random variation significantly affect the patterns of co-variation that one actually perceives. The failure to recognize or take into account the effects of chance can lead one to faulty conclusions about the "causes" of phenomena that are more plausibly explained as the result of chance.

Kahneman and Tversky relate an example of such unwarranted casual attributions.[8] A group of flight instructors had noticed a phenomenon of pedagogical interest--that when a student was praised for having made an exceptionally fine landing, his or her subsequent landings would always be worse. And when a student was severely rebuked for an exceptionally poor performance, subsequent landings would invariably improve.

The pedagogical conclusion might seem obvious--as it did to the flight instructors--the praise given for the good performances in some way <u>caused</u> the subsequent deterioration in performance, while the rebukes were causally efficacious in bringing about an improvement. Kahneman and Tversky argue, however, that the observed phenomena--that exceptionally poor performances are generally followed by better performances, and exceptionally fine ones are generally followed by worse--can be better explained as simply the operation of chance, viz., the phenomenon of regression to the mean. <u>Any</u> "extraordinarily good" performance, one at the extreme end of a bell-shaped curve, is likely to be followed by a performance that is "less good", i.e., a performance that more closely resembles the mean value, the "average" performance. And, similarly, any particularly poor attempt is most likely to be followed by an improvement. These effects, moreover, would be observed regardless of the praise or rebukes delivered by the instructors. Had the instructors merely recognized the existence of this probabilistic effect, they might have had less confidence in the efficacy of their motivational strategy.

As the skill of the pilot improves over time, one would indeed expect the quality of the average performance to improve. However, it would still be true that some of the landings would be exceptionally fine, most would resemble the average, and some would be worse than average. One would also expect that with experience the range in quality between the best and worst performances would become more narrow, perhaps so narrow as to be imperceptible to most. One would not, however, expect every single performance to be better than the last. But whatever the range or the degree of quality of the performances, the instructors would have had reason to question their causal attributions, had they recognized the possible effects of the operation of chance.

This example illustrates a seemingly paradoxical advantage of a close attention to the problems of probabilistic reasoning. It may often be that, as the result of a careful and informed estimate of probabilities, one may become considerably less confident about one's conclusions than one might otherwise have been. This, though, would seem to be a salutary effect. When one's predictions, attributions of causality, or other judgments are uncertain, are indeed determined partly by chance, one would do well to be aware of that fact. Moreover, it seems that the willingness to recognize, and the ability to realistically evaluate the degree of uncertainty of one's own best judgments is a necessary character trait for the critical thinker. If so, then an understanding of probabilistic reasoning would be not only a useful inducement to critical thinking, but a necessary condition.

Ross and Nisbet consider the dynamics of an apparently "self-serving" causal attribution, the attributing of one's successes to personal merit, and one's failures to chance or to situational factors.[9] They argue that such biases in causal attribution result from the failure to correctly evaluate the significance of the available evidence. Alternatively, they suggest that what appears to be a self-serving "bias" is in fact not a biased judgment, but rather is an accurate assessment of the <u>total</u> body of evidence one has available, much of which is uniquely available to the individual involved. That evidence might include, for instance, a history of past successes that are strongly correlated with particularly intense efforts.

Nisbet and Ross do note, however, that motivational biases--the wishes and preferences of the individual--may serve to directly determine what that body of "available" evidence actually comprises. An individual, for instance, "...can say quite correctly that 'the people I know seem to like my work,' while being blind to the fact that it is in part for that very reason that he knows them."[10]

In either case, the accuracy or quality of the causal attribution is determined by one's evaluation of the evidence with which one is presented. And in evaluating that evidence, one needs to answer a probabilistic question, viz., Given this evidence, how likely is it that my causal hypothesis is true? It would seem clear that without an understanding of the fundamentals of probabilistic reasoning, the "assessment" of the value of the evidence can be no more than a guess.

C. Generalizations

Synchronic generalizations are judgments about the probable characteristics of groups at a given time, based on the observation of a sample of the members of the group. Such generalizations are an integral part of the process of concept-formation--they involve the recognition that characteristics are not randomly distributed among individuals, but tend to occur in predictable clusters. Similarly, judgments about the anticipated characteristics of individuals are probabilistic whenever they are based on beliefs about the characteristics of the group as a whole.

Thus, when an individual is observed to show one of a hypothesized "set" of characteristics, it will be judged likely that he or she also possesses the entire set of associated traits, although these are not observed.

To make such a generalization on the basis of an inadequate sample would be to fall into the fallacy of "hasty generalization"--and, in an informal logic course students would be introduced to this as an error to be avoided. But, before one is able to in practice "avoid" making hasty generalizations, one must know when, and under what conditions, a particular generalization is to be counted as "hasty". One needs to have an appreciation of the probabilistic factors that would make a sample adequate or inadequate.

The first and the most obvious factor is simply the number of instances observed. However, that number by itself is not sufficient to determine the adequacy of the sample. One must also take into account two additional factors: first, the variability observed within the sample; and second, one's prior knowledge of the overall degree of variability within the broader reference class.[11] When there is little variability, generalizations made on the basis of a small sample are legitimate; the greater the variability, the greater the sample size that is required. Holland et al.

present a number of empirical studies which indicate that in making such generalizations people do seem intuitively to take note of the degree of variability present in a sample.[12]

However, these studies suggest that systematic errors in the assessment of variability occur, particularly when individuals are engaged in making social generalizations. It appears that individuals consistently overestimate the degree of variability occurring in their own "in-group", and are thus reluctant to generalize from a limited sample, while consistently underestimating the degree of variability within other groups. The stereotyping that results from the lack of recognition of the variability in a reference class is essentially a failure in probabilistic reasoning.[13]

Nisbett and Ross examine the current debate among psychologists as to the basis of racial and ethnic prejudice--whether such prejudice is due primarily to motivational factors or to cognitive factors. They argue that the origin of an individual's belief in ethnic stereotypes <u>cannot</u> be explained as the result of an individual's own detecting of a covariation between traits and group membership, citing studies that indicate individuals' capacity to detect degrees of covariation is simply too limited. Hence, even if there <u>were</u> some empirical validity to a particular stereotype, "...people's covariation detection capacities are far too crude to allow any such purely 'data-based' discovery. The most unreasoning bigot would have to concede that the condition of virtually perfect covariation ...necessary for data-based covariation detection is not met in his 'data'."[14]

Nisbett and Ross conclude that, rather than being erroneously induced from experience, ethnic stereotypes are culturally transmitted--one simply believes what one is told. They suggest that the more important question with respect to probabilistic reasoning is, Why, when a stereotype is not warranted empirically, does the available evidence to that effect not quickly demolish the erroneous belief? I.e., why do people continue to believe in truth of various stereotypes, despite abundant available evidence that the belief is simply not correct?

Nisbett and Ross reject the common explanation, that motivational factors, e.g., the desire to justify abuse, or greed, or fear, or simply a fundamental malevolence, adequately account for the persistence of racial and ethnic stereotypes. They note, first, that all conceivable categories of individuals are stereotyped, e.g., librarians, joggers, people with beards, etc.. And, second, few of those stereotypes are unambiguously negative, and some can be glowingly positive.

5. Systematic Errors in Probabilistic Reasoning

A. Introduction

The considerable body of literature in this field purports to be not only descriptive but also prescriptive. Bar-Hillel, for instance, writes "The proper, normative way to combine the inferential impacts of base-rate evidence and diagnostic evidence is given by Bayes' rule."[15] I.e., while the primary, initial goal is to determine empirically how people generally do think, to discover the cognitive processes individuals actually employ, the secondary goal is to critique those strategies, to determine how well, and to what extent the inferential strategies commonly employed conform to what are taken to be the acceptable inferential norms. The problem for the reviewer of this literature is that there is significant disagreement about both the empirical observations and the normative claims. I will attempt to treat these issues separately, setting out in this chapter a brief review of the empirical findings, and of the associated normative claims; in the next chapter I will examine in greater detail some of the more controversial normative claims.

B. Heuristics for Probabilistic Reasoning

Kahneman and Tversky have concluded that there are various intuitive "heuristics", or cognitive strategies, which are used in the process of judging probabilities and in coping with probabilistic information. Among the most common of these problematic heuristics are the "availability", "salience" and "representativeness" heuristics.

1. Availability and Salience

Using the availability heuristic, individuals, it is claimed, consciously or unconsciously assume that those particular items most easily retrieved from the memory are in fact the items that have been encountered most frequently in experience. The individual then uncritically infers that those items are likely to be encountered again, invoking a frequency sense of probability. This heuristic is one that the individual is supposed to have induced from his or her own experience. According to Tversky and Kahneman, "Life-long experience has taught us that instances of large classes are recalled better and faster than instances of less frequent classes."[16]

The problem with reliance on such a heuristic is that, in any particular instance, the assumed correlation between actual frequency and "retrievability" may not exist. For instance, having met a child with an uncommon name, one might notice and readily recall each subsequent hearing of that name, while failing to note, and hence being less readily able to recall, the unremarkable names that are in fact more commonly encountered. One's memory in such a case would be an unreliable indicator of the actual relative frequency of various names. Similarly, having bought a new Mercedes, one might begin to notice them everywhere; but one's increased ability to bring to mind instances of having seen a Mercedes would not be an accurate indicator of there having been an _actual_ increase in the number of Mercedes on the road.

"Availability" per se would seem to represent a limitation on the quality of one's informal reasoning, rather than an "error" or a "bias" _in_ reasoning. In making "everyday", informal judgments of probability there would often seem to be nothing else that one _could_ use, other than whatever information one is able to bring to mind. The use of the availability heuristic would be problematic only if it leads to a too hasty conclusion of one's search for information or evidence. The effective reasoner would require an apprecia-tion of the need to gather a substantial body of evidence by some objective means. An uncritical reliance on the availability heuristic would seem to be an instance of the informal fallacy of "hasty generalization".

Cohen argues that the subjects who seem to have been employing an "availability" heuristic are doing no such thing, but have been manipulated by the experimental design to fall, temporarily, for a "cognitive illusion". He maintains that the subjects themselves would, on reflection, repudiate the supposed heuristic, and notes that there is an important difference between a) making and relying on the assumption that frequency can be reliably estimated by means of ease of recall, and b) making an estimate of frequency by reference to the only data one has, namely that which is "available".[17]

The second empirical claim made by Kahneman and Tversky is that persons tend to give a greater weight to vivid, personally experienced information, while disregarding or at least discounting other equally available information that is known in a more abstract way. This tendency Kahneman and Tversky refer to as the problem of "salience". For example, after having been personally involved in a frightening automobile accident, one might henceforth be convinced that such accidents are very common indeed. Or, having personally experienced harassment of some sort, for instance, sexual or ethnic harassment, one might immediately conclude that harassment of that sort is "ubiquitous", failing to take sufficient note of the far more common experience of <u>not</u> being harassed.

This heuristic is a variant of the first, since events that have a great emotional impact would be far more readily called to mind than similar but contrary events lacking that impact.

<u>2. Representativeness</u>

The third cognitive strategy that Kahneman and Tversky claim is commonly employed in probabilistic reasoning is the "representativeness" heuristic. This heuristic is said to be used, for instance, when making judgments about the likelihood that an individual "belongs to" a certain group. The judgment that an individual is highly likely to belong to a group is made if he or she, in some particularly salient way, resembles ("represents") the stereotype of the group. Errors in this sort of judgment are important socially, since the individual, once categorized, would be expected to share a broad range of traits believed to be characteristic of that group.

Nisbett and Ross argue that the representativeness heuristic is fundamental in all cognition, that the discovery of this heuristic is a rediscovery of the fact that in mental operations one thing can "represent", i.e., be taken as a symbol for, other things, provided only that in some salient feature the symbol resembles the thing symbolized.[18] This heuristic they maintain generally serves us well. For instance, the heuristic could not lead one astray when the members of the category in question are "characterized by properties that are both unique and invariant."[19] Any individual having such diagnostic characteristics must in fact be a member of the group.

3. The "Lawyer/Engineer" Problem

Kahneman and Tversky set out what has become a classic example of the use of the representative heuristic, in a 1973 study, "On the Psychology of Prediction".[20] In this study, subjects were asked to estimate the probability that an individual, who had been chosen randomly from a group of 100 lawyers and engineers, was in fact a lawyer. The subjects were given two sorts of information to use in making the judgment about the individual, John: first, distributional data, viz., the ratio of lawyers to engineers in the sub-population (30:70, or 70:30); and second, individualized information, viz., a personality profile.

Kahneman and Tversky report that, when given only distributional information, e.g., the information that 30% of the group are lawyers, most subjects correctly concluded that there was a 30% probability that John is a lawyer. However, when the individualized information was <u>also</u> provided, the subjects seemingly ignored the distributional data, basing their probability judgments <u>solely</u> on the similarity between John's personality and the stereotyped image of the "lawyer's personality". Kahneman and Tversky conclude that the use of the representative heuristic leads to a "neglect of the base rate", i.e., a neglect of the prior probability, a probabilistic reasoning error.

The use of this sort of heuristic would not, perhaps, be too alarming, <u>if</u> the available singular data actually were of high predictive value. However, the subjects' tendency to focus primarily on the singular information persisted even when the subjects recognized that the singular data available to them was entirely worthless. For example, when given an entirely irrelevant personality profile along with the distributional data, the subjects in the studies by Kahneman and Tversky[21] tended to conclude that each of the two occupation was equally probable, ignoring the initial 70:30, or 30:70 bias in the group. According to Kahneman and Tversky, to ignore the relevant distributional information in this way is to commit an obvious probabilistic reasoning error. Indeed, even when the individual data do provide some worthwhile evidence, the base rate data remains significant and should be incorporated into the judgment.

Kahneman and Tversky conclude that people have little or no intuitive understanding of the norms of Bayesian reasoning.

At this point it should be noted the occurrence of the phenomenon itself is in some doubt. First, as Gigerenzer[22] points out, earlier studies of "intuitive statistics" indicated a tendency not to disregard but rather to excessively weight the distributional data, a tendency termed conservatism.[23]

Second, recent studies have suggested that subjects do give attention to the base-rate data, when the task is not structured in such a way that the significance of that information is disguised.[24]

Ginosar and Trope, using this same problem scenario, found the same effect. But, they also found that when the information subjects received was modified so as to provide inconsistent occupational cues, most subjects did at that point resort to the base-rate information. Moreover, subjects who failed to do so were later found, when tested on very simple "urn-type" problems, to have no conception of the relationship between "relative frequency" and "probability". I.e., nearly all the subjects who possessed even a rudimentary understanding of the concept of probability were able to apply base-rate considerations in their assessments of the probability of group membership.[25]

In addition, Ross and Nisbett claim that in certain situations the use of an inappropriate heuristic such as representativeness is in practice appropriate, in that it would ordinarily cost the reasoner little if anything. For instance, some characteristics are "diagnostic" of certain categories, and the Pr(category /characteristic) = 100%. If so, assigning a particular individual to a class on the basis of that characteristic would work well. But to appropriately use the heuristic one must distinguish between those characteristics that are unique to a group, and which may <u>either</u> be common or uncommon among the group, and other characteristics which may be "typical" of a group, i.e., very common among members of a group, but which may <u>also</u> be possessed by individuals <u>outside</u> the group. If by 'diagnostic' one refers to the first case, then it is hardly surprising that the reliance on the diagnostic characteristic "works", because this reasoning is normatively correct, according to Bayes' theorem. In this case, the probability that S belongs to the group given that S has characteristic x is 100%, and that certainty would be diminished as the occurrence of the trait outside the group increases. It seems unlikely, though, that this is generally the way that the representativeness heuristic is used, since traits that are unique to a group need not be at all common within the group, and thus would not necessarily seem "representative" of the group. Note that the relative "commonness" of the trait within the group is logically completely irrelevant in this case.

If, on the other hand, by 'diagnostic traits' one refers to traits that are typical, i.e., commonly, or even always possessed by members of the group, then the use of the trait to assess the probability of an individual's membership in that group would be problematic. This is because, however salient and remarkable such a trait may be, information about that trait is entirely irrelevant to the problem at hand. The use of this sort of information would only be safe if the common trait also happened to be unique to the group.

The error involved in this sort of reasoning seems to me to be what might be termed the "inversion" fallacy, the common but mistaken belief that the Pr(a/b) is equal to the Pr(b/a), or that those two conditional probabilities at least have <u>some</u> rather close relationship. That is, the truth of a statement of conditional probability does not permit one to infer the truth of the inverse. For example, it would be a mistake to believe that the probability that a subject has certain personality traits, given that the subject is a lawyer, [Pr(traits/lawyer)], is <u>equal to</u> the probability that the subject is a lawyer, given that the subject does possess the representative set of traits

[Pr(lawyer/traits)].

For example, suppose for the moment that <u>all</u> lawyers wear glasses and carry briefcases. The Pr(traits/lawyer) is thus 100%. So, knowing that John is a lawyer, we may rightly infer that John wears glasses and carries a briefcase. But, suppose we know instead that a) that all lawyers wear glasses and carry briefcases, and b) John wears glasses and carries a briefcase. If asked, Given this information, how likely is it that John is a lawyer?, a person using the representativeness heuristic would, incorrectly, respond, "Quite likely, because John <u>looks</u> like a lawyer."

The only correct response, however, given only this information, is that <u>no</u> inference can be made, since there is no relationship whatsoever between the probability of (trait/lawyer) and that of (lawyer/trait), the Pr(a/b) and the Pr(b/a). Before any inference can be drawn one needs to know: first, the probability that a person wears glasses and carries a briefcase given that that person is <u>not</u> a lawyer; and second, the prior probability that the person is, or is not, a lawyer. In short, while it may be true that John "looks like a lawyer", there could be any number of non-lawyers who also look like lawyers, and that is the important information.

In short, the error made by Kahneman and Tversky's subjects is not that they incorrectly used the information given, but that they failed to recognize that no answer whatsoever was possible on the information given them, that necessary information was missing. It may be, of course, that the subjects <u>did</u> recognize this, and obligingly came up with their own rough estimates of the missing data. If so, it would be impossible to assess the quality of their reasoning, since we have no way of knowing what those critically important intuitive estimates might have been.

Notice that, in order to intuitively use Bayes' theorem, one must have some estimate of the probability of observing the evidence, given that hypothesis in question is <u>false</u>.

Kahneman and Tversky manage to arrive at the correct Bayesian answer without having or using that piece of information, but only by the (technically correct) conversion of Bayes' theorem to its "odds-likelihood ratio" form. This is done by 1) devising a second hypothesis, 2) setting the information (some of which is not actually available) in the form of Bayes' theorem, and then 3) dividing the original equation by the second, at which point the missing information "drops out". This move does eliminate the need to estimate the Pr(e/-H), but it staggers the imagination to suggest that naive subjects, in their informal estimates of probability, either can or should be doing this sort of calculation.

Note, also, that <u>if</u> Pr(trait/not lawyer) were taken to be zero, i.e., if there is <u>no</u> possibility of encountering "false positives", then one <u>could</u> validly infer that the Pr(lawyer/trait) = 100%, <u>no matter what</u> the Pr(trait/lawyer) might be, i.e., even if it were only .1%. The posterior probability sought is only diminished when the probability of a "false positive" must be taken into consideration. Perhaps what is seen in this study is not a question of "ignoring of base-rate" as much as a failure to recognize the possibility that the traits in question might well appear in the larger, non-lawyer population, and/or the failure to recognize the significance of the fact that they might.

It should perhaps also be noted that the Pr(trait/ lawyer) and the Pr(trait/not lawyer) are not reciprocals. It would be perfectly possible for both to equal 100%, or, indeed, any other combination--there is no necessary relationship at all between the two probabilities.

An alternative interpretation of Kahneman and Tversky's classic study is suggested by their own observation of the subjects' responses when given an entirely non-diagnostic personality profile ("Dick is a 30 year-old man...married, with no children...of high ability...well-liked by his colleagues..."[26]). Given this "information", Kahneman and Tversky report that subjects judged that "the likelihood that a particular description belonged to an engineer rather than to a lawyer"[27] was 50:50, i.e., the probability of either occupation was 50%.

But, notice that, with this particular locution, it does seem that Kahneman and Tversky are asking, "What is the probability of finding this profile, given that the individual is an engineer (or lawyer)?"

But, if this is the question the subjects were asked, or believed themselves to have been asked, the answer they've given is exactly correct! The "occupationally neutral" profile is equally likely for these two occupations, as it would be for any occupation. A plausible explanation of the so-called "neglect of the base-rate" is that, rather than employing a faulty intuitive heuristic, the subjects were simply failing to correctly understand the question they were asked. This would be a relatively simple conceptual problem, which could be readily addressed by instruction in the fundamentals of probability.

Further questions arise when one tries to determine what the appropriate response ought to have been. A great deal hangs on the unknown assumptions made by the subjects, and incorporated into their judgments of probability. For instance, consider the following possible reasonings.

The base-rate of engineers in the group, one knows:

the Pr(Eng) in the population = .30; hence,

the Pr(-Eng) in the population = .70.

But, in order to use Bayes' theorem, one must make a judgment about the probability of finding the character traits given in the personality profile among the engineers and the lawyers. And, to do so, one can only refer to one's own background knowledge about engineers and lawyers in general, since one knows nothing at all about the character traits of the individuals in this particular sub-population.

Suppose, for instance, that one judges that:

(1) the Pr(char./Eng) = 100% and the Pr(char/-Eng) = 50%. then, using the base-rate data given in problem, the Pr(Eng./char.) = 46%.

Or, suppose (2) that one judges instead that the Pr(char./Eng) = 20% and the Pr(char/-Eng) = 10%. Then, again, the Pr(Eng./char) = 46%.

Or, suppose (3) that one believes that the Pr(char./Eng) = 80% and the Pr(char./-Eng) = 40%. Then, once again, the Pr(Eng/char) = 46%.

These three alternatives illustrate that even radically different judgments about the probabilities of observing the character traits in question can, when used in Bayes' theorem, generate the same "correct" answer.

But, suppose (4) that one judges that the Pr(char./Eng) = 80% and the Pr(char./-Eng) = 10%. Then, the Pr(Eng./char) = 77%.

But, suppose (5) that one judges, as above, that the Pr(char./Eng) is 80%, but takes the Pr(char./-Eng) to be 0%. Then, the Pr(Eng./char) = 100%.

The last three alternatives, (3), (4) and (5), illustrate that even when one holds the Pr(char/Eng) estimate constant, while varying only the estimate of the Pr(char/-Eng), very different correct answers will be reached.

And, finally, suppose one believes that the Pr(char./Eng) = 50% and the Pr(char./-Eng) = 50%. Then, the Pr(Eng/char) = 30%, which is of course equal to the prior probability with which one started.

In short, <u>any</u> answer at all can be defended as correct, since the answer depends on the unstated assumptions that the subjects might have introduced into the problem, and indeed must introduce if the problem is to be solvable at all.

Clearly, before making any judgment, one needs to know two things. First, how likely is it that any lawyer will have the character profile in question? This may well be thought of as, How "representative" of "lawyers-in-general" is this particular character profile? And, to estimate this figure, one can hardly do more than to a) recall all the lawyers one has encountered, and estimate what proportion of them displayed the character, and b) consider whether the character in question for some reason would be expected to occur among a certain proportion of (unobserved) lawyers. For instance, one might conclude that "an interest in politics" would be a characteristic shown by a large proportion of lawyers, while a fascination with mathematical puzzles would be characteristic of an equally large number of engineers.

The *key* consideration, however, would be the one that might easily be overlooked--one must ask, How likely is it that anyone who is *not* a lawyer (or, alternatively, an engineer) would display that characteristic?

Suppose, for instance, one decides that **90%** of all engineers would be interested in mathematical puzzles, while only **10%** of the general population would be so interested. Then, the probability that X, who is interested in puzzles, is an engineer would be about 75%.

Notice, too, that given any more realistic estimate of the likelihood of lawyers for the general population, the "usefulness" of the character trait as an indicator of occupation would plummet; for instance, assuming a prior probability that X is a lawyer of .1% (undoubtedly a high estimate, even given our litigious society), the conditional probability that X is a lawyer, given the interest in politics, would be .18%, in other words, not very likely.

<u>Summary</u>

In this chapter I have first, discussed the concept of probabilistic reasoning, indicating the range of possible interpretations of the central concept, viz., probability. No consensus presently exists among theorists of probability as to the "best" interpretation, or even as to whether there is a single interpretation. It may be that the concept is fundamentally ambiguous, so that, depending on the context of the problem at hand, different senses of probability are most appropriate. The decision about the required interpretation has definite consequences for the sort of calculation one undertakes, and hence will significantly affect the judgment of probability at which one ultimately arrives.

Second, I have indicated the sorts of "everyday reasoning" activities that involve one in probabilistic reasoning. Judgments of probability are required in making predictions, in assessing evidence, in making causal attributions, in making generalizations about group characteristics, and in drawing conclusions about individuals on the basis of beliefs about group membership. These judgments may be made without any process of conscious appraisal, without the deliberate application of a selected set of reasoning norms, i.e., on an "intuitive" basis.

Third, I have set out and discussed some of the large body of psychological research which supports claims about the existence of persistent, systematic errors and biases in individuals' intuitive beliefs about probability. The gist of such claims is that, at least under test conditions, individuals frequently violate the norms of probabilistic reasoning, which are taken to require a Bayesian analysis of the data at hand.

There is, however, considerable controversy among philosophers concerned with theoretical accounts of probability, of induction, and of rationality as to whether or not the Bayesian norms accepted by the psychologists are indeed normative. Clearly, if Bayesian reasoning is not really required for "rational" thinking, the research demonstrating persistent non-Bayesian patterns of thought would not be indicating a problem, and there would be no reason to include instruction in Bayesian reasoning in the critical thinking course. In the next chapter, then, I will consider the controversy over the status of Bayesian reasoning.

1. L. Jonathan Cohen, *An Introduction to the Philosophy of Induction and Probability* (Oxford: Clarendon Press, 1989), p. 43-47.

2. Max Black, *Margins of Precision: Essays in Logic and Language* (Ithaca: Cornell University Press, 1970), p. 122-124.

3. Cohen, *The Philosophy of Induction and Probability*, p. 40-65.

4. Daniel Kahneman, Paul Slovic and Amos Tversky, eds., *Judgment Under Uncertainty: Heuristics and Biases* (Cambridge: Cambridge University Press, 1982).

Richard Nisbett and Lee Ross, *Human Inference: Strategies and Shortcomings of Human Judgment* (Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1980).

5. Daniel Kahneman and Amos Tversky, "Intuitive Prediction: Biases and Corrective Procedures", in *Judgment Under Uncertainty*, p. 414.

6. Cohen, *The Philosophy of Induction and Probability*, p. 116.

7. Ibid., p. 117.

8. Daniel Kahneman and Amos Tversky, "On the Psychology of Prediction", in *Judgment under Uncertainty*, p. 66.

9. Nisbett and Ross, *Human Inference: Strategies and Shortcomings of Human Judgment*, p. 231.

10. Ibid., p. 237.

11. John H. Holland, Keith J. Holyoak, Richard E. Nisbett, and Paul R. Thagard, *Induction: Processes of Inference, Learning and Discovery* (Cambridge, Massachusetts: The MIT Press, 1989), p. 232.

12. Ibid., p. 20.

13. G.A. Quattrone, and E.E. Jones, "The perception of variability within in-groups and out-groups: Implications for the law of large numbers." *Journal of Personality and Social Psychology*, 1980, 38, p. 141-152.

14. Nisbett and Ross, Human Inference: Strategies and Shortcomings of Human Judgment, p. 239.

15. Maya Bar-Hillel, "The Base-rate Fallacy in Probability Judgments", Acta Psychologica, 44 (1980), p. 212.

16. Amos Tversky and Daniel Kahneman "Availability: A Heuristic for Judging Frequency and Probability," in Judgment Under Uncertainty, pp. 166-168.

17. L. Jonathan Cohen, "Can human irrationality be experimentally demonstrated?" The Behavioral and Brain Sciences 4 (1981), p. 323.

18. Nisbett and Ross, Human Inference: Strategies and Shortcomings of Social Judgment, p. 242.

19. Ibid., p. 254.

20. Kahneman, Slovic and Tversky, eds., Judgment Under Uncertainty, p. 53-57.

21. Richard E. Nisbett, Eugene Borgida, Rick Crandall, and Harvey Reed, "Popular Induction: Information Is Not Necessarily Informative", in Judgment under Uncertainty, p. 111.

Daniel Kahneman and Amos Tversky, "On the Psychology of Prediction", Psychological Review, 1972, p. 430-454.

22. Gerd Gigerenzer et al, The Empire of Chance: How Probability Changed Science and Everyday Life (Cambridge: Cambridge University Press, 1989), p. 149.

23. Cameron R. Peterson and L. R. Beach, "Man as an Intuitive Statistician", Psychological Bulletin 68 (1967) 29-46.

Cameron R. Peterson and W. M. Ducharme, "A primacy effect in subjective probability revision." Journal of Experimental Psychology, 73 (1967), 61-65.

Lawrence D. Phillips and Ward Edwards, "Conservatism in a simple probability inference task." Journal of Experimental Psychology 72 (1966), 346-354.

Ward Edwards, "Conservatism in human information processing" In B. Klienmuntz (ed.) Formal Representation of Human Judgment, (New York: Wiley, 1967).

24. Z. Ginosar and Y. Trope, "The effects of base rates and individuating information on judgments about another person." Journal of Experimental Social Psychology, 16 (1980), 228-242.

25. Ibid.

26. Kahneman, Slovic and Tversky, eds., Judgment Under Uncertainty, p. 56.

27. Ibid., p. 5.

CHAPTER IV

NORMATIVE ISSUES IN PROBABILISTIC REASONING

1. Introduction

A claim that has become commonplace in the psychological literature, as indicated above, is that people tend to do poorly in various ways when faced with probabilistic information. Yet the _descriptive_ accounts in the psychological literature of the cognitive strategies individuals employ cannot by themselves yield this conclusion. Before concluding that people do well or poorly in their reasoning, one must also have a normative account of probabilistic reasoning, i.e., one must determine how people _should_ reason in problems involving uncertainty. But, the normative questions are no less controversial than the empirical. In this chapter I will consider some of the conflicting positions.

Consider, for instance, the error in probabilistic reasoning discussed above--the claimed failure to consider relevant distributional data. The claim is that, in making probabilistic judgments people tend to ignore the base-rate data and rely exclusively on singular information, and do so even while recognizing that the particular singular information they possess is irrelevant. The assumption underlying Kahneman's and Tversky's psychological research into probabilistic reasoning is that the Bayesian model is correct, and sets the norm for such reasoning. In the next section, for the convenience of the reader, I will set out and then discuss this theorem.

2. Bayes' Theorem

Bayes' theorem was first set out by Thomas Bayes in 1763.[1] The theorem, which is directly derived from the basic axioms of the probability calculus, allows one to calculate the conditional probability of A given B, that is, Pr(A/B), from: 1) the conditional probability of B given A, Pr(B/A); 2) the probability of A, Pr(A); and 3) either the probability of B, Pr(B), or (what is equivalent) the probability of (B given A) plus the probability of (B given -A), [Pr(B/A + Pr(B/-A)]. So, given Pr(B/A) and some additional information, the inverse probability, Pr(A/B), can be calculated.

The theorem can thus be used to provide a means to "update" one's estimate of the conditional probability of a hypothesis after acquiring new information or evidence, (e), that is relevant to one's hypothesis, (H). When used in this way, the theorem requires that there be two hypotheses under consideration, and that these be mutually exclusive and exhaustive, i.e., H and not -H.[2]

It might be useful here to briefly sketch Bayes' theorem, and to set out a model case of the use of the theorem. Bayes' Theorem can be used to assess the probability that a hypothesis, H, is true, given some evidence, e. In such a case, the value to be determined is the <u>conditional</u> probability <u>of the hypothesis</u> given the new evidence, i.e., the Pr(H/e). There are three distinct, independent pieces of information that are required if one is to use the theorem. The first[3] is the initial probability of the hypothesis, Pr(H), called the "prior probability" of the hypothesis. This value is "assigned" by the reasoner (hence, a subjectivist interpretation of probability is required). The Pr(-H), which is also required (when using a straightforward form of the theorem), is simply a function of the Pr(H), i.e., Pr(-H) = (1 - Pr(H)).

The second distinct piece of information needed is the probability of the <u>evidence</u> given the <u>truth</u> of the hypothesis, Pr(e/H), i.e., the probability that the evidence at hand <u>would</u> be observed if the hypothesis were in fact true.

The third distinct piece of information is the probability of the _evidence_ given the _falsity_ of the hypothesis, Pr(e/-H). The theorem states that

$$Pr(H/e) = \frac{Pr(H) * Pr(e/H)}{[Pr(H) * Pr(e/H)] + [Pr(-H) * Pr(e/-H)]}$$

A Simple Model Case:

> Suppose that you are confronted with two unlabelled urns, each of which contains 100 small objects. In the first ("M"), there are 100 marbles; in the second ("D"), there are 70 marbles and 30 diamonds. You will be allowed to choose and to keep one of the two urns. The probability that you will choose Urn D, the prior probability of D, is 50%. Before making the choice you ask for some further information, and are allowed to take _one_ sample from one of the urns, i.e., to obtain a single piece of evidence ("e").
>
> Suppose that, with great luck, the object you pull out is a diamond. What is the probability that you have in your hands Urn D?

Note, first, that the correct answer is intuitively obvious. Clearly, given this evidence, the probability that this is Urn D is 100%. And Bayes' theorem generates the intuitively right answer.

$$Pr(D/e) = \frac{Pr(D) * Pr(e/D)}{[Pr(D) * Pr(e/D)] + [Pr(-D) * Pr(e/-D)]} =$$

$$= \frac{.5 \quad * \quad .3}{[.5 * .3] + [.5 * 0.0]} = \frac{.15}{.15 + 0.0} = 100\%$$

The evidence in this case is <u>conclusive</u>, both intuitively and according to Bayes' theorem, and this is so because the [Pr(-D) * Pr(e/-D)] is zero. And, <u>this</u> is so because the Pr(e/-D) is zero.

Suppose, though, that you've been less lucky in your sampling, and have come up with a marble. What does this evidence tell you about the probability that you're holding Urn D? The temptation, I take it, would be to say that, since the evidence was possible given either urn, the evidence is inconclusive and hence of no use. But, while it is quite true that the marble is inconclusive, it is not true that it is worthless. It is possible to discover the significance of that evidence by applying Bayes' theorem.

$$Pr(D/e) = \frac{.5 \quad * \quad .7}{[.5 * .7] + [.5 * 1.0]} = \frac{.35}{.35 + .5} = 41\%$$ [4]

Given this analysis, one would be well advised to opt for the other urn, which would have a 59% probability of being the "diamond" urn.

Note that, though superficially complex, the theorem is quite simple to apply (when the needed information is readily available). There are a number of observations that can be made that would permit an "intuitive" use of the theorem, i.e., an informal and very simply calculated use, to generate rough estimates of conditional probabilities.

First, the conditional probability, Pr(H/e), will always be equal to 100% whenever the second term of the divisor is 0; and, second, the _larger_ the second term of the divisor is relative to the first (which is the same as the dividend), the _smaller_ will be the Pr(H/e). For example,

let 'x' be the first term, [Pr(H) * Pr(e/H)], and

let 'f' be the second term, [Pr(-H) * Pr(e/-H)]. Then the **conditional probability of the hypothesis given the evidence**, Pr(H/e), is:

$$\frac{x}{x + f} \quad \text{or, equivalently,} \quad \frac{1}{1 + R},$$

where R is the ratio of x to f.

So, if: f = x.....Pr(H/e) = 50% ;
f = 2x.....Pr(H/e) = 33% ;
f = 3x.....Pr(H/e) = 25% ;
f = 4x.....Pr(H/e) = 20% ;
f = 9x.....Pr(H/e) = 10% .

And, if: f = 1/2x...Pr(H/e) = 66% ;
f = 1/3x...Pr(H/e) = 75% ;
f = 1/4x...Pr(H/e) = 80% ;
f = 1/9x...Pr(H/e) = 90% .

Notice also that if the two prior probabilities are equal to one another, i.e., Pr(H) = Pr(-H), then that term drops out of the equation, and can be ignored in one's rough figuring.

Black notes, however, that this move, simply setting the initial probabilities of the hypotheses at .5, involves the use of the widely discredited Principle of Indifference, and hence is problematic.[5]

Suppose that the prior probability of both the hypothesis and the negation of the hypothesis are for some reason judged to be equal, i.e., suppose that there is initially a 50% chance that the hypothesis is true, and a 50% chance that it's false. Then, whenever the **Pr(e/H)** equals the **Pr(e/-H)**, the probability of the hypothesis given that evidence, Pr(H/e), will be 50%, and this will be true no matter what the Pr(e/H) is.

Notice, too, that if the sum **(x + f)** = 1, then the **Pr(H/e)** equals the **Pr(e/H)**, and this, again, is always true, no matter what x and f are.

If the prior probabilities of the hypotheses are not equal, the necessary calculations become slightly more complex, since the Pr(e/H) and the Pr(e/-H) must be weighted by the Pr(H) and the Pr(-H), respectively.

Since f and x are probabilities, their values will always be between 0 and 1, which further simplifies things.

These relationships, once understood, would seem to be fairly easy to use in making estimates, particularly when one is willing to accept an answer that is a "rough estimate". And, to accept rough estimates would seem reasonable, since the probabilities employed in the formula are generally themselves rough estimates.

So, what are we to make of the claim that, in general, people fail to make use of this sort Bayesian assessment when estimating the probabilities of their hypotheses? In the next section I will consider possible explanations of the empirical observations.

3. The Claimed Failure to Use Bayesian Reasoning

Several possibilities immediately arise. First, the empirical claim could simply be wrong, i.e., perhaps people in general do make use of this sort of information. This could be the case if the psychological evidence advanced to support the claim were in some way flawed. For instance, it could be that insufficient data has been obtained to support this conclusion. Ginosar and Trope, reinvestigating the "neglect of base-rates" studies of Kahneman and Tversky, found that, when presented with "...similar categories, inconsistent information, and unrelated information... subjects incorporated base rates into their judgments in an orderly fashion."[6] They conclude that "in sum, the present study suggests that the base-rate fallacy is not as prevalent as earlier work would lead one to expect."[7]

Second, it could be that some alternative hypotheses that would "fit" the data have been overlooked; for instance, it might be that when using this information, additional common-sense (or what are believed to be common-sense) premises are added by the subjects to the information given in the problem scenario. Such assumptions might lead the subjects to very different, but nevertheless reasonable estimates of probability.

Third, it could be that the empirical claim is correct; perhaps in some circumstances people <u>do</u> fail to use a Bayesian approach. We would then still have to ask the more important question, viz., Why? Again, several possibilities immediately emerge. First, it may simply be that this phenomenon is one that is eminently an educational problem. Perhaps the failure to intuitively reason probabilistically is simply the result of a lack of prior attention by the student to the significance of this sort of information. If so, simply directing the student's attention would lead to improvement in probabilistic reasoning.

Or, perhaps an "intuitive" recognition of the importance of such reasoning does occur, but errors arise because of a failure to "intuit" the finer points of probabilistic reasoning, the subtleties and problems that may arise. Similarly, while the basic rules of language usage are universally intuited by all normal children exposed to human speech, it seems unarguable that, quite often, considerable "room for improvement" persists, even among adult native speakers of a language. Hence, explicit instruction is given in the rules of usage, and is combined with practice, coaching, etc., and this often leads to the desired improvement. If the problems in probabilistic reasoning are analogous to problems in the usage of language, again, such problems would seem to be ameliorable via instruction.

Or, it may be, contrary to the claims of many in psychology, that statistics does not "speak with a single voice". Gigerenzer and Murray demonstrate that different theories of statistics when applied to the Cab problem of Kahneman and

Tversky, give different, and non-Bayesian, answers as the "correct" estimate of the probabilities. The Neyman and Pearson theory, for instance, deliberately introduces an openly subjective element, requiring that the reasoner decide upon a "criterion" setting, which determines the ratio of types of errors, i.e., the ratio of "misses" to "false alarms". And, when that criterion is selected so as to minimize false alarms, the correct answer according to the Neyman Pearson theory is very close to that which Kahneman and Tversky observed in their studies, which they designated as demonstrating a reasoning error.[8]

4. Should Bayesian Reasoning Be Used?

This brings us to the normative issue, an issue which remains controversial. The question is, Should Bayesian reasoning be adopted as a norm for probabilistic reasoning? In the psychological literature, the assumed answer is that Bayesian reasoning is indeed normative, in all contexts, and that any divergence from such reasoning should be counted as an error, or even as evidence of irrationality. But, this may be too facile a conclusion.

For instance, in the Lawyer/Engineer problem described above, one might accept the empirical findings, but deny that the observed "failure" to make use of base-rate information constitutes a mistake in reasoning. The question to be asked is, Should one make use of that group-based information whenever it is available, or should one, on the other hand, concentrate exclusively on the information one has about the particular individual? Kahneman and Tversky hold that base-rate data are always pertinent. But others have argued that, in at least some circumstances, it is the use of the base-rate data in arriving at a judgment that actually counts as "irrationality".

L.J. Cohen charges that "some investigators of irrationality ...proceed as if all questions about appropriate norms have already been settled...as if existing textbooks of logic or statistics had some kind of canonical authority."[9] He maintains that, given the controversies that still exist with respect to the norms of probabilistic reasoning, the purported "errors" in probabilistic reasoning may not constitute errors at all.[10] Cohen in fact defends the reasoning of the everyday reasoner, arguing that in Kahneman and Tversky's "Cab problem", the intuitive, non-Bayesian probability estimates of the subjects are preferable to the Bayesian answers which Kahneman and Tversky consider correct. At this point I will set out the "Cab problem", and discuss the normative issues that have arisen in relation to it.

The Cab Problem

This version of the cab problem is set out by Kahneman and Tversky in Judgment under Uncertainty (1982):

> A cab was involved in a hit and run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:
>
> (a) 85% of the cabs in the city are Green and 15% Blue.
>
> (b) a witness identified the cab as Blue. The court tested the reliability of the witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colors 80% of the time, and failed 20% of the time.
>
> What is the probability that the cab involved in the accident was Blue [as reported] rather than Green?[11]

According to K&T, this problem "permits the calculation of the correct posterior probability under some reasonable assumptions"[12].

Their research, however, shows that the "base rate" data, the 85% Green and 15% Blue figures, which are taken to represent the prior probability of involvement in the accident, seem not to be used by most subjects. "The median and modal answer [given by their subjects] is typically .80, a value which coincides with the credibility of the witness, and is apparently unaffected by the relative frequency of Blue and Green cabs."[13] This phenomenon has been repeatedly observed, and has been found to be stable under a number of variations of the problem. According to Bar-Hillel, "The genuineness, the robustness, and the generality of the base-rate fallacy are matters of established fact."[14]

The Bayesian calculation that is taken to be required is:
Pr(**Blue cab**/"blue report") =

$$\frac{Pr(B) \quad * \quad Pr("b"/B)}{[Pr(B) * Pr("b"/B)] \quad + \quad [Pr(-B) * Pr("b"/-B)]} =$$

$$\frac{.15 \quad * \quad .80}{[.15 * .80] + [.85 * .20]} = \frac{.12}{.12 + .17} = .41$$

and hence the probability that the errant cab was Green would be .59.

So, according to Kahneman and Tversky, "...in spite of the witness's report...the hit-and-run cab is more likely to be Green than Blue, because the base-rate is more extreme than the witness is credible."[15]

The subjects seem to have been making two mistakes. First, they seem to be assuming that the two different conditional probabilities are interchangeable, i.e., that Pr(a/b) = Pr(b/a). That is, they seem to believe that the probability of the witness reporting "Blue" on seeing a Blue cab, which is 80%, is quite simply equal to the probability that the cab is Blue, given that the witness says that it is, i.e., that the Pr("B"/B) = Pr(B/"B"). This of course is an invalid inference, since there is no necessary equivalence between inverse conditional probabilities.

This sort of mistake would seem to be most easily accounted for as simply an ignorance on the part of the subjects of the basic workings of the probability calculus.

The second "mistake" is the more problematic. The subjects seem not to be inclined to make any adjustment for the base-rate data, but treat it as if it were entirely irrelevant to the problem at hand. This move cannot be attributed to a lack of knowledge as to how to properly take into account this sort of information. It is not that the subjects are "using the data incorrectly"; rather, they are simply ignoring its existence entirely.

However, in another variation of the problem, subjects are given what are considered to be "causally relevant" base-rate data--subjects are not given the relative frequency of cabs in the city, but instead are given the relative frequency of cab accidents in the city, i.e., the information that 15% of the cabs involved in accidents in the city are Blue, while 85% of the cabs in accidents are Green. In this case, subjects' answers were highly variable, but generally subjects did seem to make some allowance for the "prior probability" given by the base-rate data. Kahneman and Tversky conclude that "causal base-rates", those which "...suggest...the existence of a causal factor that explains why any particular instance is more likely..."[16], <u>are</u> employed in intuitive assessments of probability, although "incidental base-rates", which do not suggest any such causal link, are not. Kahneman and Tversky maintain, however, that <u>both</u> causal and incidental base-rate data <u>ought</u> to be used, that both are equally relevant.[17] This normative claim, though, is one that is hotly disputed.

<u>Cohen's Critique</u>

The normative issue thus raised has been considered by Cohen.[18] Cohen rejects the reasoning given by Kahneman and Tversky. He argues that the .41 probability figured by Kahneman and Tversky is "...the value of the conditional probability that a cab-color identification by the witness is ...[correct], on the condition that it is an identification as...[blue].[19] (I should note here that in Cohen's presentation of the problem, he refers to a variation that has the witness identifying the cab as green, which confuses the issue by changing all the numbers involved. I have tried, for the sake of clarity, to make Cohen's account match the original version by making the appropriate substitutions.)

Cohen argues that the issue before the court does not, and ought not to, "concern the long run of cab-color identification problems", but rather "...just the probability that the cab actually involved in the accident was blue, on the condition that the witness said it was [blue]."[20] Cohen concludes that "if the jurors know that only 20% of the witness's statements about cab colors are false, they rightly estimate the probability at issue as [80%]...the fact that cab colors vary according to an 85/15 ratio is strictly irrelevant..."[21]

The jurors, says Cohen, should be interested only in the "causal propensity" of the witness to correctly identify cab-colors, and this is dependent only on "causal properties, such as the physiology of vision, [which] cannot be altered by facts...that have no causal efficacy in the individual events...the mere relative frequency of blue and green cabs...does not generate any causal propensity for the particular cab in the accident."[22]

Cohen thus vindicates the ordinary "common-sense" judgments of the man in the street (and not surprisingly, his conclusion has a great deal of intuitive plausibility). To do so, Cohen adopts the propensity account of probability. But having adopted this account, it would seem that the propensity of the witness to identify cab colors accurately would be far less significant than the propensity of the particular driver in question to a) be involved in an accident, and b) having been involved, to leave the scene. The propensity of the witness would in no way alter the propensity of the driver.

The propensity interpretation of probability is elaborated by Popper;[23] it seems particularly plausible in this problem scenario, since such probabilities can be attributed to particular individuals. The difficulty is that given this account there is no clear way to determine just how various operative factors contribute to the final propensity. According to Cohen, "The main weakness of a propensity analysis is that it does not intrinsically carry with it any distinctive type of guidance in regard to the actual evaluation of probabilities...since the talk of propensities has no distinctive numerical implications, it provides no inherent basis for the assignment of actual probability-values."[24]

An Alternative View: Mill

Interestingly, the same sort of a problem is addressed much earlier by John Stuart Mill in A System of Logic.[25] For Mill, to engage in the sort of detailed calculations of prior and posterior probabilities which, according to Kahneman and Tversky are required for correct probabilistic reasoning, would involve "...the misapplication of the calculus of probabilities which have made it the real opprobrium of mathematics."[26] In his comments, Mill is referring specifically to the then current attempts to apply probabilistic reasoning to the "...credibility of witnesses, and to the correctness of the verdicts of juries."[27]

According to Mill, there are several problems with such attempts. First, Mill argues that it makes no sense to speak of, or to try to determine, "a general average of the veracity ...of mankind..."[28]; second, even if such a prior probability, i.e., distributional data for the class as a whole, could be reasonably assessed, that class information would be virtually worthless in determining the position of any single individual within the overall pattern of distribution of values. Averages, writes Mill, "...are of extremely small value as grounds of expectation in any one individual instance, unless the case be one of those in which the great majority of individual instances do not differ much from the average."[29] Mill sets this out as a fallacy of probabilistic reasoning--"...the fallacy of reasoning from a wide average to cases necessarily differing greatly from any average."[30]

Like Cohen, Mill maintains that a rational juror would concentrate on the characteristics of the individual witness, "...the degree of consistency of his statements, his conduct under cross-examination,...the relation of the case itself to his interests, his partialities, and his mental capacities"[31], and would disregard the distributional data as irrelevant to the determination of the particular instance.

Mill's issue, one will note, is not identical to the "Cab problem". But it seems that the reasoning Mill uses can be applied to the contemporary puzzle. For instance, it seems that Mill would have the rational juror in the cab case concentrate on the characteristics of the individual cab drivers involved, the accounts of their actions given by the drivers, their prior accident records, any relevant maintenance records, the proficiency of the drivers, etc.. Mill, to be consistent, would have to argue that "distributional data" such as the number of taxi-cabs of various colors in the city, would be irrelevant to the determination of the guilt of these two possible culprits.

Similarly, Mill would note that being involved in a hit-and-run accident is a characteristic which is not one for which "the great majority of individual instances do not differ much from the average".[32] Thus, even if we were informed that 85% of all the accidents in the city were caused by Blue Cabs, this still would not be useful to us in determining whether or not it is likely that this particular accident was caused by this particular Blue Cab.

As noted above, this variation on the original problem was in fact tried by Kahneman and Tversky.[33] Cohen argues that this change in the data (to indicate the distribution of accidents with respect to the two companies) is quite significant, and hence, once again, the naive subjects demonstrate the ability to reason rationally by attempting to include that new data.

There is another remarkable similarity between Cohen's argument and Mill's explication of probability. This is the emphasis on the significance of causal properties, which

Cohen refers to as "causal propensities". Cohen distinguishes between "probability functions that measure relative frequencies and probability functions that measure causal propensities...the propensity-type probabilities may be estimated from frequencies..."[34] Mill maintains that one's primary interest or goal in any calculation of "frequency-type" probabilities is to discover the causal relationships that give rise to the frequency data. The very purpose of probabilistic reasoning, according Mill, is to allow us to separate the effects of chance from the effects of causation, and so allow us to advance our knowledge of the causes that ultimately control the occurrences of events.

It is possible, though, that Mill might take another tack. As noted above, Mill does provide for the calculation of the probability that a particular effect was caused by one or another possible cause, and he sets out a formula similar to Bayes' theorem in his explanation of how this is to be done. Mill writes, in what seems to be an example similar in many ways to the Cab Problem, "Let the causes [A and B] be alike in...[this] ...respect: either A or B, when it exists, being supposed equally likely...to produce M; but let A be in itself twice as likely as B to exist, that is, twice as frequent a phenomenon. Then it is twice as likely to have existed in this case, and to have been the cause which produced M."[35] But notice here that the two possible causes are by hypothesis equally likely to produce M. In the cab case, we are not given any information as to the number of accidents per Green Cab and the number of accidents per Blue. Only if these numbers were equal could we make use of the prior probability in this way.

Mill goes on to consider cases where the two possible causes differ both in their frequency and in their propensity to cause the effect. In such a case, "the probability that M was produced by either cause is as the antecedent probability of the cause (its frequency), multiplied by the probability that if it [the cause] existed it would produce M."[36] But, again, in the Cab case, although we have access to the first number, we do not have access to the second. So it would seem that, according to Mill's argument, the only conclusion in the Cab problem would be that we cannot calculate in any meaningful way the probability that the accident was caused by a Blue or Green Cab.

There is one final point of similarity between Mill's treatment and Cohen's. Cohen admits that, although when given the eye-witness testimony one should rely exclusively on that evidence, one might not have any such evidence. Cohen claims that in such a case one should revert to the base-rate data, and one then should make the judgment on the basis of that evidence. "Of course, if no testimony is mentioned and subjects know nothing except the relative frequency of the differently colored cabs, then no causal propensity is at issue and the only basis for estimating the required probability is indeed the relative frequency."[37] This, though, casts considerable doubt on the consistency of Cohen's account. Why would it be permissible to use the base-rate data in one case, and not in the other? Cohen offers no justification. It would seem that one sort of information cannot lose its pertinence simply because additional information has been gathered. Cohen's reasoning here seems reminiscent of Mill's willingness to accept the Principle of Indifference as the basis of the assignment of probabilities, provided we have "nothing better to go on". It would seem to me that both Mill and Cohen too readily fall back on what each admits is an inferior sort of data, data which is held to be entirely useless whenever something better is available. Both would seem to be more consistent if they were to conclude that no judgment of probability can be made without the correct sort of data, whatever that may be.

Mill closes his discussion of this issue with a rather scathing comment, which has some clear educational implications. He writes that these are the sorts of errors are "frequently committed by men who, having made themselves familiar with the difficult formulae which algebra affords for the estimation of chances...like better to employ those formulae...than to look out for means of being better informed."[38] (An example of the "hammer" phenomenon--give a small boy a hammer, and he will very quickly discover that everything needs pounding...)

A Variant on the Cab Problem

Let us consider an alternative version of the same problem. Suppose that the accident occurred in a city in which the general population is 85% Black, and in which the cab-drivers are also 85% Black. Suppose, for good measure, that out of all past cab accidents, 85% were determined to have been caused by Black cab-drivers. Suppose, though, that in this case an eye-witness has identified the driver as Caucasian, and that this witness is able to make correct identifications in 80% of the test-cases (80% of the Black persons are identified correctly; 80% of the Caucasians are identified correctly); and, suppose further that only two cabs were in the vicinity, and two suspects were brought in--one Caucasian, the other, Black.

You are a juror, and you must decide whether it is rational to rely on the Kahnemann and Tversky calculations, or to reject Bayesian reasoning as normative in this context. If you decide that Bayesian reasoning is the rational choice, according to Kahneman and Tversky's analysis, it would seem that you must conclude, despite the contrary testimony of the very accurate eye-witness, that the Black cab-driver is more likely to have been involved in the accident. If, on the other hand, you reject the Kahnemann and Tversky account, you must decide that it is rational to ignore the base-rate data and to rely on the pertinent individual data. It would seem that in this scenario, to continue to believe, despite the contrary evidence, that <u>this</u> particular Black person is guilty of this particular crime, merely on the grounds that there are a large number of Black persons in the neighborhood, would be an example of the most blatant and egregious racial prejudice, and not at all a shining example of rational thought.

But, the problem with this conclusion is that, since Bayes' theorem is directly derivable from the basic probability calculus, it would seem that we cannot simply reject Bayesian reasoning without relinquishing the probability calculus itself. How can this conundrum be resolved?

Perhaps the Bayesian reasoning is being misapplied in these examples. Cohen, for instance, points out that on a frequency interpretation of probability, one cannot make the move from a general inference about the probable characteristics of an individual randomly drawn from a population, which would be sanctioned by Bayes' theorem, to a specific inference about the characteristics of a particular individual, say, George. This is because George may be significantly different from the other members of the supposed reference class, so that the probability that George is "a bassoon-historian" may not be at all the same as the probability that any randomly chosen individual is a bassoon-historian. And, if we assume that George is unique, the only appropriate reference class is the class containing only George himself. But then, "we should only obtain a reliable probability for George's being a bassoon-historian if and only if we are 100 percent certain that he is one or that he is not one."[39]

Moreover, the usual interpretation of probability adopted by the proponent of "Bayesianism" is the personalist interpretation of Savage and De Finetti. The statement that "the probability that x is F is .85%" is understood to express one's degree of belief that x is F; and this value is to be assigned, not discovered. Hence there is no reason why the "prior probability" of the hypothesis in the cab problem must be taken to be equal to the relative frequency of cabs in the city, or to any other empirically determined value. Similarly, neither the "probability of the evidence given the truth of the hypothesis" nor the "probability of the evidence given the falsity of the hypothesis" need be set equal to the experimentally determined "accuracy of the witness".

Moreover, the "accuracy of the witness" is a problematic attribute. Note that what we are given are the conditional probabilities:

1) the probability of "blue" reports given Blue cabs = (80%); and
2) the probability of "green" reports given Blue cabs = (20%).

But what would be of more interest, it seems, would be the experimentally determined inverse conditional probability, i.e.,

the probability of Blue cabs given "blue" reports, under the experimental conditions.

And this information can easily be acquired by using Bayes' theorem, along with the findings of the accuracy tests.

I.e.,

$$Pr(Blue/"b") = \frac{Pr(Blue) * Pr("b"/Blue)}{[Pr((Bl) * Pr("b"/Bl] + [Pr(G) * Pr("b"/G)]} .$$

Assuming equal numbers of Blue and Green cabs were used, this is:

$$\frac{50\% * 80\%}{(50\% * 80\%) + (50\% * 20\%)} = \frac{40\%}{40\% + 10\%} = \frac{4}{5} = 80\%$$

In other words, using Kahneman and Tversky's numbers, the <u>diagnostic</u> information, under experimental conditions, i.e., the probability that a Blue cab has been seen when a "Blue" report is given by this witness, is <u>the same as</u> the "witness accuracy", the probability that a "Blue" report is given when a Blue cab is seen. This unusual result is purely artifactual; it only occurs in this case because the probability of "true positives" (80%) has been arbitrarily chosen by Kahnemann and Tversky to be complementary to the probability of the "false positives" (20%). (And, because it can reasonably be assumed that, in the test, equal numbers of Blue and Green cabs were used.)

Bar-Hillel has conducted empirical investigations similar to Kahneman and Tversky's. She considers, and rejects, the possibility that her subjects are "mistaking" the retrospective probability actually given, Pr("b"/B) = 80%, for the diagnostic probability actually required, Pr(B/"b"). But surprisingly, Bar-Hillel goes on to acknowledge that "...if you believe you are told that...when the witness says 'the cab was Green' (or Blue...), he stands an 80% chance of being correct, then you are quite right in giving 80% as your answer, irrespective of the base-rate conditions."[40] But she asserts that "a very bizarre perceptual mechanism would have to be assumed to produce...[the diagnostic information]...given that we take percepts to be caused by external events and not *vice versa*."[41]

But what seems strange here is that Bar-Hillel fails to note that the desired diagnostic information, though obviously not given by perception, is readily generated simply by applying Bayes' theorem and a plausible assumption about typical testing conditions. It is this inference about the relation between Pr("b"/B) and Pr(B/"b") that constitutes paradigmatic Bayesian reasoning. But Bar-Hillel rejects this possible interpretation out of hand, and maintains that the subjects err by failing to use Bayesian reasoning at all.

One problem with the "cab problem" in its many variants seems to be that several very different sorts of information are simply "plugged in" to Bayes' formula, which is then used to crank out a "correct answer".

First, what one is looking for is a judgment of subjective probability. No one, I take it, can doubt that this particular cab either was, or was not, involved in the accident, and the statement, "The probability is x% that the cab in the accident was this blue cab" can only be taken to express one's degree of confidence in that conclusion. But, the probability taken to be the "prior probability" of that hypothesis is an objective, "frequency" sort of probability, simply the number of blue cabs divided by the number of all cabs in the city.

Third, the "probability of the evidence given the truth of the hypothesis", another "frequency", is one that is unrelated to the last --i.e., we are given as evidence the frequency of "blue" identifications given a set of blue cabs in a test situation.

When Bar-Hillel gave subjects a variant of the cab problem in which the individuating information given related <u>direct-</u>

ly to the populations of blue and green cabs, rather than to the accuracy of a witness, (a radio intercom was definitely heard in the offending cab; it is known that 80% of the Green cabs and 20% of the Blue cabs have intercoms) the subjects did in fact attempt to integrate the base-rate and the individuating information, though the wide range of the probability estimates suggests they had no clear idea how to manage that integration.[42]

The "Signal Detection" consideration

Birnbaum argues that Kahneman and Tversky's normative claim is incorrect, because for human beings the detection of the relevant signals (in this case, cab color) is not independent of the base-rate.[43] Birnbaum concludes that when signal detection theory is considered, it turns out that the commonly given answer, that the probability that the errant cab was blue was 80%, in fact turns out to be consistent with Bayes' theory.[44] He also points out, though, that given the various theories of signal detection extant and the various theories of judgment, and, moreover, the range of possible subjective input values (e.g., the witness's criterion for saying "Blue" or "Green"), any answer could be made consistent with Bayes' theorem.

This account cannot be taken, though, as a vindication of the ordinary, man-in-the-street judgments of probability, since the signal detection theory that one must assume the subjects to be intuitively applying is exceedingly complex.[45] (And Birnbaum notes this in the conclusion.)

5. Further Examples of Probabilistic Reasoning.

In this section I will take a detailed look at several additional examples of scenarios that require probabilistic reasoning, leaving aside the empirical psychological questions about how people do reason to consider the normative issue, viz., How should one reason in these scenarios?

The School Enrollment Scenario

Kahneman & Tversky cite a study done by Ajzen (1977), which purports to show that "causal" base rate information is incorporated into intuitive probabilistic reasoning, even though "incidental" base rate data is not.[46] Kahneman and Tversky maintain that both causal and incidental base-rates should be employed.

I shall consider this example in detail. The question Ajzen asked his subjects was, What is the probability that student A (for whom a personality sketch has been provided) will choose to take a history rather than an economics course? (Or vice-versa?)

As base-rate data, Ajzen provided his subjects with the proportion of students enrolled in each of the two courses, viz., 70% and 30%. These figures were intended to indicate to the subjects the relative popularity of the two courses, and this data was supposed to establish the relative attractiveness of the two courses. The course-attractiveness data was expected to be used by the correct reasoner as the "causal base rate data". I.e., it was assumed that the "attractiveness" would at least in part cause the differences in enrollment numbers, and so would constitute a prior probability that should be incorporated by the reasoner in a Bayesian fashion.

Ajzen found that when attempting to estimate the probability that A would choose one course over another the subjects did make use of the data given them, intuitively incorporating it (in some way) with the personal data about student A.

But, should they have done this? Is this "good" reasoning, or yet another example of a faulty intuition with respect to probabilistic reasoning?

Let's consider a slightly different example, one that is, I take it structurally the same as the one given by Ajzen. This example, I think, will suggest that Ajzen's so-called causal base rate data is irrelevant to the problem, and so ought to have been ignored by his subjects.

The Pass/Fail Scenario

Suppose that you are the administrator of a small college at which, for many years, all students have been required to take one Phys. Ed. course per year. Each student has the option to take this course either as a "Pass/Fail" or as a "Graded" course (PF or G). Records have been kept, and it is known that over the years 70% of the students have chosen the "PF" option, while only 30% have chosen the "G" option. Moreover, suppose that this ratio has been remarkably consistent over time, so that, at any given time, 70% of the students are in the PF group, and 30% are in the G group.

You conclude that the "Pass/Fail" option is the more popular, and indeed, the more "attractive" option.

Suppose you then ask yourself, What is the probability that any randomly chosen student from the incoming class will decide on the "Graded" option? Your answer will undoubtedly be, 30%,. You would rely entirely on the base-rate data, and it seems uncontroversial that you would be correct to do so.

But, suppose instead that you ask a question about a _particular_ student, Student A, whose dossier you have just been studying. Let us say that this particular incoming student has in general had very low grades, but has always excelled in athletics. Let us also say that a well-known and unbendable policy in this college is to dismiss any student whose overall grade-point average falls below a C.

What is the probability, you might ask yourself, that Student A (who knows all the relevant facts and who is definitely planning to stay in college) will choose the "Graded" Phys. Ed. option over the Pass/Fail option?

You would undoubtedly conclude that, in this case, _this_ student is virtually _certain_ to take the "Graded" option. I.e., you would _ignore_ the so-called "base-rate" data, and rely on what seems to be the rather conclusive data about the _individual_ student. Moreover, as I think seems obvious in this example, you would be _correct_ to do so.

You might then go on to critique your thinking process, and ask, But _shouldn't_ I have made some effort to "incorporate" the (70%:30%) base-rate data into my reasoning? The answer here seems to be, No. But, if not, <u>why not?</u>

To answer this, let's look in detail at the formal calculation that you would be attempting to approximate intuitively.

The "<u>probability of Pass-Fail (PF) given student s</u>", that is, Pr(PF/s), would be what you wish to estimate. If Bayes' theorem were to be used, one would need to determine:

$$\frac{Pr\ (PF)\ *\ Pr\ (s/PF)}{[Pr(PF)\ *\ Pr(s/PF)]\ +\ [Pr(-PF)\ *\ Pr(s/-PF)]}\ .$$

The prior probabilities for each option seem simple; Pr (PF) = .70, and Pr (-PF) = .30 . This would be the probability in the long run, given a large number of students, that each option would be chosen.

But, what can be said about the two conditional probabilities, Pr (s/PF) and Pr (s/-PF) ? How are we to understand these terms? Should this be read as "the probability, given that PF has been chosen by Student s, that Student s has chosen the PF option? (And for the second term, the probability, given that "graded" has been chosen, that Student s will be found among those being graded?)

But if so, the respective probabilities are obvious, viz., 100% in each case. And if so, we would find that the conditional probability, Pr(PF/s), would necessarily be the same as the prior probability, i.e.,

$$\frac{.70 * 100\%}{[.70 * 100\%] + [.30 * 100\%]} = \frac{.70}{1} = .70$$

This seems to be because the "condition" PF in Pr (s/PF) is entirely determined by s. PF can be chosen by Student S if and only if Student S does indeed choose PF. But, this would seem clearly to be a less than useful statement of the problem.

Does it make sense to think of the student as being, somehow, "evidence" that stands a certain chance of being observed if the hypothesis is true, that is in some sense "likely" to be found by taking a random sample of the population? Perhaps the difficulties arise because the hypothesis itself has not been clearly stated. Recall that in the model case, the urn problem, the prior probability of Urn D was taken to be 50%, since it was one of two possible choices. The prior probability had nothing to do with the "content" of the urns. Suppose, in this case, we consider the Pr(PF) to be .5, and the Pr(-PF) to be .5. Then these figures would drop out, and we would be left to consider the relation between "true positive" and the "false positive" figures, i.e.,

the Pr(e/H) and the Pr(e/-H).

In our case, let's say the Pr(s/PF) = 1% and the Pr(s/-PF) = 99%. (The ratio is 99:1.) Then the Pr(PF/s) = 1%.

Of course, this was obvious, since the sum of the two numbers in the divisor, 99% and 1%, is 1.00, and hence the Pr(s/PF) is in this case equal to the Pr(PF/s).

But, note that it is not necessary that these two numbers equal 1.00, since there is in fact <u>no</u> necessary relationship between these two conditional probabilities.

So why, in this example, does it seem like they <u>ought</u> to, and <u>have to</u>, add up to one? Maybe <u>this</u> is the problem--obviously, either the student is in <u>one</u> class, or in the other, and the sum of the two probabilities <u>would</u> equal one--But, in Bayes' theorem, (a) the probability of finding the <u>evidence</u> given the <u>falsity of the Hypothesis</u> is **independent** of (b) the probability of finding the evidence given the <u>truth</u> of the Hypothesis--witness the fact that in the model case, there could be <u>any</u> number of marbles and/or diamonds placed in the two urns.

In this example, it would seem that it is inappropriate to use Bayes' theorem here, since the only significant factors affecting the choice of the student are the characteristics of the individual, Student S.

There is an alternate interpretation of the problem: We could be looking for the probability that PF will be chosen given a subset of students to which Student S belongs, that is,

Pr (PF / {Students similar to St.X}).

This would solve some problems, but unfortunately, it generates others. We would then need to determine, for Bayes' formula, the Pr({Students similar to S} / PF), and the Pr ({Students similar to S} / -PF).

Note, first, that this information is not given anywhere in the problem, and can not be derived in any way from the given information. So subjects could not be faulted for failing to "use" this information, although they could be faulted for failing to recognize that necessary information is unavailable, and that the "problem" as stated cannot be solve.

It seems clear at least that there is considerable room for clarification of the norms of "correct" probabilistic reasoning, and that simple possession of the "correct formulas" will not guarantee correct reasoning.

6. Is Bayesian Reasoning Normative?

Gigerenzer and Murray (1987) observe that while the "new psychology of thinking" "...depends on the assumption that Bayes' theorem is normative...it is only by neglecting content, context, and information that this normative assumption is made tenable."[47]

They argue that neither a "concept isomorphism" nor a "structural isomorphism" exists between Bayes' theorem and the everyday problems faced by the critical thinker. Concept isomorphism requires that there be an unequivocal matching of the formal concepts used in the theorem, i.e., prior probability and conditional probability of evidence/hypothesis, with the features of the problem. This isomorphism is lost, for instance, when two or more possible assignments of prior probability are plausible.

In the cab problem, for example, Gigerenzer and Murray argue that one cannot be sure whether the prior probability of Blue cab involvement should be taken to be 15%, because the relative frequency of Blue cabs is 15%, or whether, using the indifference principle, it should be taken to be 50%, because no cause is known that would make one color cab more accident-prone than the other. Hence concept isomorphism is lost, and Bayesian reasoning cannot properly be used.

Structural isomorphism exists only when all of the structural requirements of the theorem are matched in the problem. I.e., the hypotheses in question must be mutually exclusive and exhaustive, the pieces of evidence must be obtained by random sampling, and those pieces of evidence must be mutually independent.[48]

Gigerenzer and Murray argue persuasively that, in the absence of unequivocal isomorphism, one cannot calculate a single, unequivocally correct Bayesian "answer", since that answer will vary depending upon the assumptions made by the subjects. And so, they argue, one cannot simply compare subjects' answers to the "correct" one, and tell whether or not Bayesian reasoning was used. Hence, studies purporting to reveal "irrationality" [read "non-Bayesian reasoning"] show no such thing. This point seems correct, at least with respect to some of the more complex problems studied. It seems difficult to say with any confidence precisely what reasoning process has been undertaken simply by observing the answer that results. A fortiori, it is difficult to decide what reasoning process have been used by the individual subjects of a study, from observing the "average" answer generated.

But it seems at least as implausible to suggest that the brain somehow intuitively sorts out the issues that, when made explicit, appear so difficult to resolve.

Gigerenzer and Murray also suggest that, in light of the absence of a single, unequivocal Bayesian answer, perhaps Bayesian reasoning is itself suspect. But, Bayes' theorem, and hence Bayesian reasoning, is directly derived, in an unproblematic way, from the axioms of the probability calculus. The difficulty thus cannot be taken to lie with the theorem itself, unless one is prepared to reject the probability calculus as well as the theorem. The problem seems to be that, in complicated real-life problems, one cannot easily tell which pieces of information one ought to use in the theorem, or even whether one has available the necessary information. But this is only to say that Bayes' theorem cannot be applied uncritically to a problem situation, nor can it be counted upon to generate "certain" answers about the probability in question. Instead, critical judgment is required, first to judge which information to make use of, and how to do so, and second, to consider and evaluate the real possibility that one perhaps could have made better judgments, and, that, if so, one's judgments of probability might be wrong. But this is only to say that these judgments are judgments, i.e., that no absolute certainty is possible.

Summary

In this chapter I have set out and discussed the controversy with respect to the norms of probabilistic reasoning. The point chiefly at issue is whether Bayesian reasoning is or is not normative, whether, in certain contexts and given certain interpretations of the concept of probability, other norms should be used in deciding whether an individual's reasoning is to count as "rational".

In the psychological literature purporting to show the existence of probabilistic reasoning errors, the assumption is that subjects always ought to be using Bayesian reasoning, i.e., ought to be incorporating base-rate data with individualized data. Some philosophers of inductive reasoning and the theory of probability, most notably, L. Jonathan Cohen, have strongly disagreed, arguing that in some cases, and in particular, in those cases most widely cited by psychologists, it is the use of a Bayesian analysis that should be counted as irrational.

I have examined Cohen's views on this issue, and have also set out a similar set of considerations offered by John Stuart Mill. I have also set out a variant of Kahneman and Tversky's Cab problem which seems to support the Cohen/Mill contention that Bayesian reasoning is, in at least this case, inappropriate. I have also, however, argued that it is not the use of Bayesian reasoning per se, but the misapplication of Bayesian reasoning, that is at issue. That is, the use of Bayes' theorem will give "irrational" results when irrelevant, inappropriate data about prior probabilities and conditional probabilities are "plugged into" the formula.

In the next chapter I will examine the evidence in the psychological literature respecting the child's development of the concept of probability, intuitions about probability that influence the child's reasoning, and the current approaches to the teaching of probabilistic reasoning.

1. Thomas Bayes, "Essay towards Solving a Problem in the Doctrine of Chances", Philosophical Transactions of the Royal Society of London, 53 (1763), 371-418.

2. L. Jonathan Cohen, An Introduction to the Philosophy of Probability and Induction (Oxford: Clarendon Press, 1989), pp. 67-70.

3. Even if one is reduced to guesswork, it is always possible to assign some initial probability to the hypothesis.

4. Alternatively, you could look for the Probability of Urn M given this evidence. If so,

$$Pr(M/e)= \frac{.5 * 1.0}{[.5* 1.0] + [.5 * .7]} = \frac{.50}{.50 + .35} = 59\ \% .$$

5. Max Black, Margins of Precision: Essays in Logic and Language (Ithaca: Cornell University Press, 1970), p. 124.

6. Zvi Ginosar and Yaacov Trope, "The effects of base rates and individuating information on judgments about another person", Journal of Experimental Social Psychology 16 (1980), p. 240.

7. Ibid.

8. Gerd Gigerenzer and David Murray, Cognition as Intuitive Statistics (Hillsdale, N.J.: Lawrence Erlbaum Associates, 1987), p. 167-174.

9. L. Jonathan Cohen, "Can human irrationality be experimentally demonstrated?", The Behavioral and Brain Sciences 4 (1981), p. 328.

10. Ibid.

11. Daniel Kahnemann and Amos Tversky, eds., Judgment under Uncertainty: Heuristics and Biases (Cambridge: Cambridge University Press, 1982) pp. 156, 157.

12. Ibid., p. 156.

13. Ibid., p. 157.

14. M. Bar-Hillel, "The Base-rate Fallacy in Probability Judgments", Acta Psychological 44 (1980), p. 215.

15. Amos Tversky and Daniel Kahneman, "Evidential Impact of Base-rates", in _Judgment under Uncertainty: Heuristics and Biases_ (Cambridge: Cambridge University Press, 1982), p. 157.

16. Ibid., p. 155.

17. Ibid., p. 158.

18. Cohen, "Can Human Irrationality Be Experimentally Demonstrated?", _The Behavioral and Brain Sciences_ (1981).

19. Ibid., p. 328.

20. Ibid.

21. Ibid.

22. Ibid., p. 329.

23. Karl R. Popper, _Realism and the Aim of Science_ (London: Hutchinson, 1983).

24. Cohen, _The Philosophy of Induction_, p. 56.

25. John Stuart Mill, _A System of Logic: Ratiocinative and Inductive_ (London: Longmans, Green, and Co., 1893).

26. Ibid., p. 353.

27. Ibid.

28. Ibid., p. 354.

29. Ibid.

30. Ibid.

31. Ibid.

32. Ibid.

33. Kahneman and Tversky, _Judgment Under Uncertainty_, p. 157.

34. Cohen, "Can Human Irrationality Be Experimentally Demonstrated", p. 329.

35. Mill, _A System of Logic_, p. 357.

36. Ibid., p. 358.

37. Cohen, "Can Human Irrationality Be Experimentally Demonstrated?", p. 329.

38. Mill, A System of Logic, p. 354.

39. Cohen, The Philosophy of Probability, pp. 48-49.

40. Bar-Hillel, "The Base-rate Fallacy in Probability Judgments", Acta Psychologica, 44, (1980), p. 221.

41. Ibid., p. 221.

42. Ibid., p. 228.

43. Michael H. Birnbaum, "Base Rates in Bayesian Inference: Signal Detection Analysis of the Cab Problem", American Journal of Psychology, Spring 1983, Vol. 96, No. 1. pp. 85-94.

44. Ibid., p. 87.

45. Birnbaum gives the following equations:

$P(\text{"B"}/B) = N(b-t)$ and $P(\text{"B"}/G) = N(g-t)$, and

discriminability = $[b-g] = N^{-1}[P(\text{"B"}/B] - N^{-1}[P(\text{"B"}/G]$,

where:

"g" is the mean value of Green cabs on the discriminal continuum;

"b" is the mean value of Blue cabs on the discriminal continuum;

"t" is the criterion for responding "Blue" or "Green", and

"N^{-1}" is the inverse cumulative standard normal density function.

It strains the imagination to believe that subjects are intuitively incorporating these considerations into their judgments, and hence are reasoning correctly.

46. Kahneman and Tversky, Judgment under Uncertainty, p. 155.

47. Gigerenzer and Murray, Cognition as Intuitive Statistics, p. 162.

48. Ibid., p. 163.

CHAPTER IV

THE TEACHING OF PROBABILISTIC REASONING

1. Introduction

In this chapter I will consider the issues related to the teaching of probabilistic reasoning, including the empirical studies on the learning of such reasoning, and current curricular approaches to teaching probabilistic reasoning. Much of the work on the "learning of probability" has focused on the ontogeny of probabilistic reasoning in elementary school age children. In the next section, this work will be briefly reviewed.

2. Probabilistic Intuitions

Early work on the development of an intuitive understanding of probability was done by Piaget and Inhelder in 1951.[1] They found evidence of an understanding of the notion of "chance events" in children as young as ten years old. Specifically, they found that when faced with the task of predicting the outcome given various "randomizing devices" (e.g., a spinner, or a slanted pegboard down which a marble is rolled), children of this age typically recognized the impossibility of the task. By the age of 12, children were able to recognize the existence of "non-uniform probability distributions", i.e., to recognize that given the structure of some devices, certain outcomes were more likely to occur than others. These children also recognized that the probability distribution provides a stable basis for prediction of "long-run" results, even though specific individual events remain unpredictable.

Yost, Seigel and Andrews have conducted studies that indicate an ability to recognize the chance nature of some events, and to behave accordingly, in very young children, four years of age.[2] [3] Similar observations were made by Davies.[4]

In contrast, other researchers have denied that there exist <u>any</u> probabilistic intuitions, and that it is this lack of intuition which leads to problems in probabilistic reasoning. Engels, for instance, proposes the teaching of probability in the general mathematics curriculum, from kindergarten through the 12th grade, but asserts that the reason probability has had less impact on mathematics than geometry is "because we have a natural geometric intuition but no probabilistic intuition."[5]

The early development of probabilistic reasoning in children has been studied more recently by Fischbein, who reviewed "...the data concerning the *conceptual organization* relevant to the ontogenesis of thinking, of the notion of chance and the estimation of odds, and of the notion of probability."[6] Fischbein asks two particularly pertinent questions: "...whether conceptual understanding of probability could benefit from practical training...[and] whether everyday practical behavior, to which the estimation of odds is intrinsic, could benefit from instruction in the theory of probability."[7]

Intuitions with respect to probability emerge early, Fischbein finds. But these intuitions, he writes, are "relatively inconsistent and ambiguous."[8] By 'intuition' Fischbein means a particular cognitive operation, which is accompanied by a "feeling of conviction, of certitude", and which results from one's experience of "stable patterns" in the world.[9] These intuitions are taken to constitute an intrinsic part of the reasoning process.[10] "Intuition is the means whereby cognition meets the requirements of speed, fluency, and coherence of effective action....They are components of intelligence in action."[11]

Fischbein hypothesizes that there exists even in quite young children a "natural intuitive substrate for notions of chance and probability...[resulting from]...the day-to-day experience...[of]...stochastic processes."[12] "The germ of intuitive reasoning about probability lies in natural 'experiments' with stochastic results, which involve predictions and random draws or other equivalent actions.[13] This intuitive substrate may facilitate or impede the child's acquisition of a more appropriate set of "secondary intuitions" through instruction in the theory of probability.[14]

Fischbein concludes that, in developing a curriculum to teach the fundamentals of probability in the context of a mathematics curriculum, the pre-existing intuitions of the child with respect to probability must be considered. "The systematic construction of a new conceptual system within the process of education cannot afford to ignore the intuitive endowment of the child."[15]

Moreover, this instruction must begin at an early age if it is to be effective in re-vamping the early, perhaps faulty intuitions. Fischbein claims that "once the basic cognitive schemas of intelligence [including one's intuitions] have stabilized (after 16-17 years of age), modifications to the intuitive substrate seem to be difficult, if not impossible."[16]

According to Fischbein, educational programs to attempt to modify pre-existing intuitions must "include motor, representational and conceptual elements..."[17]; i.e., the instruction must involve the student in the same sort of "natural experimentation" which led to the establishment of the original inadequate intuitions.

Piaget and Inhelder, in contrast to Fischbein, suggest that the child does not, and cannot, develop any "conceptual schema" with respect to chance before approximately the age of seven. They maintain that the notion of chance rests on the notion of the "irreversibility" of certain processes; hence the development of an understanding of the notion of chance cannot occur, they maintain, until the child reaches the level of "formal operations".[18]

Fischbein, however, argues that Piaget and Inhelder fail to distinguish between a) having "primary intuitions" with respect to chance (as opposed to necessity), for which the child has no explanation, and, b) having an understanding of the concept of chance.[19]

3. Fallacies of Probabilistic Reasoning

Gambler's Fallacy: Some psychological studies of the gambler's fallacy are instructive. Jarvik noted in 1951 what he termed the "negative recency" effect, the tendency of subjects to predict an outcome which had not recently occurred, rather than to predict the outcome that had occurred recently and with greater frequency, even while recognizing that the outcomes were independent of one another.[20] This study has been criticized, though, on the grounds that the sequences presented to the subjects consisted of "runs" that were in fact generally quite short. Given this fact, the subjects' use of the so-called "gambler's fallacy" in the context of Jarvik's study was not fallacious, in as much as the strategy did in fact yield predictions that were "more often right than wrong...".[21]

In other studies in which relatively long "runs" were used, the gambler's fallacy was not observed; instead, the subjects' predictions illustrated a "positive recency effect", i.e., the probability that "red" would be predicted as the next outcome increased as the length of the "red" run increased.[22]

Fischbein describes the origin of the related fallacious belief, which he calls the "sampling fallacy". This fallacy is the belief that the probability that "the next item sampled" will be of a certain kind either increases, or decreases, depending on the nature of the previously sampled items. He notes that when one draws samples from a homogeneous population, it is in fact true that, most likely, the proportion of elements in the sample will equal the proportion of elements in the parent population, and it is less likely that they will depart from that value. This is a correct intuition. But, this is not to say that the next item in the sample is either more or less likely to be of a certain kind.[23]

For instance, suppose one considers the proportion of boys and girls in families of six children. The most likely proportion is in fact "three girls and three boys", and most would find this unsurprising. But despite this, the probability of each particular sequence of boys and girls is identical, e.g., Pr(GBGBGB) = Pr(BBBBBBB) = Pr(GGGGGG) = (.5*.5*.5*.5*.5*.5). I.e, no matter what sequence has been observed, the probability that the next child will be a girl or boy remains the same, i.e., .5. But this sometimes seems intuitively surprising.

Cohen notes that the gambler who estimates that there is a high probability of "Heads" coming up after a long series of

"Tails" may not be committing an error in reasoning, but rather an error in communication. It may be that the gambler has in mind not the probability of finding Heads on the next toss, but the probability of not tossing at least one Heads in a series of n tosses. And that probability does decrease as the number of tosses in the series increases. Hence, in this sense, it does become more and more likely that the series will come to include a "Heads".[24]

Fischbein hypothesizes that the correct belief about the expected proportion is correctly induced by children from their experience, but then is inappropriately generalized, and the belief is formed that the particular sequence [GBGBGB] or [GGGBBB] is more likely than the particular sequence [GGGGGG] or [BBBBBB] This would seem to be the origin of the "representativeness" heuristic described by Kahnemann and Tversky.[25]

Another source of fallacious reasoning in probabilistic matters, according to Weir, is the persistent belief among adult subjects that some rational pattern exists in the observed events, that the complex patterns observed are indeed fully determined, according to some extraordinarily complex rule. This leads to the belief that, with enough persistence and ingenuity, one can discover and use that rule, or set of rules, to give the events one is interested in a 100% predictability. Weir suggests that the instructions given to children in experiments designed to test the understanding of probability often give the impression that some such pattern exists. Given this belief, the persistent variability of responses that are observed, as the children search for the particular rules "governing" what are in fact random events, do not indicate erroneous reasoning. Rather, the continued search for a rule constitutes "the rational way to play the game", as they understand it.[26 27 28]

Fischbein terms this a tendency toward the "rationalization" of events. For example, if the economy is in decline, people tend to believe there must be some cause, a cause which can be discovered and removed. Such a "causal explanation for everything" approach to problems of everyday life can become pernicious when individuals begin to search not just for the cause, but for a causal agent, asking not, What could have caused this deplorable event?, but, Who could have caused this? And, for what reason? While it may be possible to trace the proximate causes of an event, the role of chance in determining the particular concatenation of circumstances leading up to the event, and those ensuing after the event, ought not to be overlooked.

4. Improving Probabilistic Reasoning

Fischbein asserts that strong social factors contribute to the failure of the child to improve upon and gradually refine his or her early, primary, intuitions of chance. Instead of becoming increasingly sophisticated probabilistic reasoners, the children he studied seemed increasingly determined to carry out the search for "univocal" answers, "...invoking causal dependencies where none exist in reality. This preference...[Fischbein argues]...is not generated by the operational structure of thought, but by the influence of the social environment, in particular that of the school..."[29]

In one study,[30] the predictions of pre-school age children about the course of a marble through channels on an inclined board were far more accurate than those of 12-13 year-olds. Although in their explanations the pre-schoolers failed to draw any distinction between "equiprobability" and "subjective whim", explaining, for instance, "...'the marble goes where it wants to...we can't know where it will go'",[31] they nonetheless did recognize that the outcome was unpredictable.

The older children, in contrast, consistently opted for elaborate causal hypotheses, sometimes devising new "mechanical-geometric principles" in their efforts to successfully predict the results ("'It will come more often down...[the right-hand channel]..., because it has a longer and narrower trajectory'".[32] Only when the complexity of the channel board was increased, so that some of the possible outcomes actually were more probable than others, did the accuracy of the older children's predictions come to equal, and then to surpass, those of the younger children.

Fischbein explains these results as a negative effect of schooling. "The teaching process--particularly as it is determined by schools--orients the child toward a deterministic interpretation of phenomena, in the sense of looking for and explaining in terms of clear-cut, certain, and univocal relations."[33]

In school, one's answers are expected to be certain, one's explanations causal, and neither an epistemic uncertainty nor a hypothesis of "unpredictability" is considered acceptable. Consider, for instance, one response to a typical math word-problem--

> Q: Mr. Jones lives 50 miles away from you. You both leave home at 5:00 and drive toward each other. Mr. Jones travels at 35 mph, and you drive at 40 mph. At what time will you pass Mr. Jones on the road?

A: Given the traffic around here at five o'clock, who knows?[34]

Anyone who has been through school knows that this is not going to be the right answer.

One might object that what any student knows is that in this context, in "word problems", one is supposed to assume a perfect constancy, and hence a perfect predictability, but it is this unspoken expectation that Fischbein argues can over time lead children to devalue, and ultimately to reject, probabilistic considerations. Given this effect of schooling, it would seem unsurprising that the rudimentary intuitions and unarticulated beliefs developed in early childhood persist unremarked and unchallenged into adulthood.

Fischbein fails, however, to give any evidence in support of the speculation that schooling is the primary agent producing this effect. But, the artificial certainty of the answers required by typical math word-problems would illustrate just such an omission of the effects of chance.

Fischbein concludes that as a result of this omission "the intuition of chance remains outside of intellectual development, and does not benefit sufficiently from the development of operational schemas of thought, which instead are harnessed solely to the service of deductive reasoning."[35]

Fischbein, Pampu and Minzat (1970) also investigated the effect of instruction on the development of intuitions of probability. They found that, even with prior instruction, preschool children could not accurately solve any but the simplest problems in judging probabilities. The 12-13 year old children, in contrast, were equally proficient in the problem-solving with and without instruction. They did find that instruction markedly improved the performance of the 9-10 year-olds.

Studies of the ability of children to recognize the probability of particular events, and to make appropriate predictions about those events, have indicated that improvement occurs following instruction in the basic concepts of probability.[36] [37] [38]

In Keller's 1971 study,[39] two different sorts of instruction were given to different groups of children, a) a specific "programmed" instruction in probability and the making of successful prediction, including practice with dice, marbles and other materials, and b) a comparatively general introduction to the concepts of probability. Keller found very little difference in the performance of the two groups when making predictions of events (although the "programmed instruction" group did outperform the "generalized instruction" group on pencil and paper post-tests on probability.) The probability task involved here was very much simplified, and quite unlike "real-life" problems that might involve chance or uncertainty. The children were simply to predict which of two possible responses, pressing the left or right button, would be rewarded with a marble; one button was always correct, but was not always "reinforced" by the reward of the marble. Keller found, unsurprisingly, that the "schedule of reinforcement" was the most significant factor leading to successful prediction. When the reinforcement was 100%, children quickly discovered the best strategy; when the reinforcement was 66%, or 33%, children seemed to search for a pattern, and took a longer time in discovering that only one button was ever rewarded.

Similar findings were noted by Falk, Falk and Levin 1980),[40] who found that, in playing a "Lottery Game", children of 6 to 7 years demonstrated the ability to recognize which of two devices gave a higher probability of a successful outcome. They conclude "...our findings indicate unequivocally the appearance of a potential for discriminating between probabilities around the age of school entry."[41]

Falk et al. suggest that games of chance in which the children must make decisions about probabilities ought to be introduced early. Such games would not only teach children how to make good probabilistic decisions, but would also acquaint children with the principle that "...even 'good decisions' (i.e., those with high a *priori* probability of success) are only probabilistically reinforced... It is only in the long run that our good decisions will pay."[42] Such games, they argue, would also "...restore the balance in favor of indeterminism...", countering the effect of schooling noted by Fischbein.

Nisbett and Ross, after examining the nature and origin of errors in social judgment conclude that such errors stem in large part from "cognitive failings", i.e., from a lack of understanding of appropriate norms of inference from one's evidence, rather than from any motivational or "moral" deficiencies. "...'moral' causes of inferential error--that is, causes involving wishes, values, or motives--are never sufficient; they require the collusion of intellectual shortcomings in the acquisition or evaluation of evidence."[43]

There is, however, evidence which suggests that such shortcomings in probabilistic reasoning can be ameliorated by training in statistics or in certain sciences.[44] Holland, Holyoak, Nisbett and Thagard, in Induction, argue that induction can be best understood as the process of developing gradually improved sets of rules that govern action.

These rules are, in general, "condition-action rules", i.e., rules having the form "If x is the case, then do y".[45] A similar sort of rule may be used to describe one's environment, to make predictions about its future state[46]; such a rule would take the form "If x is the case, then y will occur". Both of these sorts of rules are diachronic, i.e., they specify expected changes over time. Alternatively, rules may be used to express "...relations holding atemporally between alternate state descriptions.."[47]; these "rules" would be concerned with the categorization of objects or events, and would take the form "If x is an A, then x is an F".

Holland et al. stress that the "rules" they refer to are not to be construed as imperatives. Many different sets of rules may be activated simultaneously, each set providing, perhaps, conflicting "advice" as to the actions that ought to be taken, or the categorizations that ought to be made.[48]

These three types of rules are used to create "mental models" of the problem situations the "cognitive system"[49] finds itself facing, and these models organize the system's knowledge. The models can be used "to generate predictions about the outcomes of potential solution attempts--that is, possibilities can be tested mentally before...[the system]...risk[s]...an overt attempt."[50]

These rules are "empirical rules"; they serve to model elements of the environment. Holland et al. theorize that cognitive systems also require "inferential rules", rules which direct the formation of particular empirical rules. An example of such an inferential rule would be a rule that states: "...If a prediction based on a strong rule fails, then create a more specialized rule that includes a novel property associated with the failure in its condition and the...unexpected outcome as its [prediction]."[51]

Such inferential rules take into account the unusualness of a situation or unexpected result, as well as a number of statistically based heuristics. One such heuristic is "The Law of Large Numbers", "...If S is a sample providing an estimate of the distribution of property P over some population, then create a rule stating that the entire population has that distribution, with the strength of the rule varying with the size of S."[52] Inferential rules allow the system to incorporate its prior knowledge of the consistency or variability of the population with respect to some feature in making predictions. Holland et al., following Piaget and Inhelder, Kahneman, Tversky among others, conclude that "...the adult has a fairly sophisticated statistical rule system for analyzing manifestly probabilistic events."[53] This system includes "...a wide variety of abstract, relatively domain-independent inferential rules that comprise pragmatic reasoning schemas."[54]

The organization of the stored knowledge of a cognitive system in the form of condition action rules provides a mechanism for what Holland et al. term "...the most fundamental learning mechanism: prediction-based evaluation of the knowledge store...A rule that leads to a successful prediction should be strengthened...one that leads to error should be modified or discarded."[55]

The notion of thinking as the development and the refinement of "rules" on which further prediction are based is similar to that presented by Dewey; he states, for instance, that the conclusion drawn in any particular problem situation "...not only settles that particular case, but...helps to fix a rule or method for deciding similar matters in the future..."[56]

Holland et al. stress several additional educational implications of their model of thinking. First, they note that persons possess a large number of statistical rules for making probabilistic assessments. They also note, however that in many every-day situations, people err in their judgments through failure to apply those rules. This phenomenon is attributed to a variety of factors, for instance: a) a failure to encode certain types of information in a form that can be utilized in statistical reasoning, e.g., the variability of a population may not be expressly noted, or may be incorrectly assessed, leading to "stereotyping" of a population; or b) a failure to correctly assess the role of chance in a particular situation. [57]

Holland et al. argue, however, that fairly simple instruction in the workings of statistical laws, which make explicit the sorts of rules that the system has induced through experience, combined with practice through examples in the proper encoding of essentially statistical information, can make significant improvement in the ability of individuals to correctly assess probabilistic situations. Moreover, the teaching of statistical relations and laws seems to have a greater effect on the subsequent reasoning of the student than does the teaching of syntactic logic. This effect, according to Holland et al., is because statistical instruction "...amounts to swimming downstream, educationally", and they note that "...rule systems that are foreign to the rules governing everyday pragmatic reasoning cannot readily be made to influence such reasoning."[58]

Holland et al. also note an educational implication of their model of concepts as a "sets of rule clusters", which give "...probabilistic assumptions about what features go with what other features...".[59] This is that "...within limits, complex sets of interrelated features will be learned more readily than isolated feature co-occurrences...[for]...once a feature is found to be predictive of any other, this can be used as the 'entering wedge' into the set of inter-relationships."[60]

Nisbett, Krantz, Jepson and Kunda, 1983,[61] suggest that attention has long been centered too squarely on the evidence of subjects' failures to use statistically sound heuristics in probabilistic reasoning. They present evidence that at least in certain clearly defined problems people do indeed display an untutored, perhaps intuitive, appreciation of appropriate norms of statistical reasoning. Moreover, Nisbett et al. set out evidence that "...training in statistics has a marked impact on reasoning. Training increases both the likelihood that people will take a statistical approach to a given problem and the quality of the statistical solutions."[62]

Nisbett et al. accept the contentions of Fischbein, that initial understandings of probability arise out of experience with randomizing devices of various sorts. They argue that under certain circumstances it becomes relatively easy for an individual to see an analogy between a real-life problem situation involving chance and uncertainty, and the operation of simple randomizing devices. And under those circumstances people do reason quite well, and make judgments on the basis of sound probabilistic intuitions. The problem Nisbett et al. set is two-part: first, to determine more exactly "...what characterizes events where an analogy to randomizing devices can be seen.."[63], and second, to discover "what factors encourage statistical reasoning, and what factors discourage it..."[64]

Nisbett et al. found that when, under experimental conditions, subjects were "prompted" to reason probabilistically, they tended to do so successfully. The prompting was accomplished by the inclusion of cues in the problem scenario stressing the element of chance in the problem, or the variability of the population being sampled.[65]

In these studies it was also found that subjects who had a higher degree of experience in a given domain were more likely to recognize the existence of probabilistic factors when presented with a problem from that domain. I.e., when asked to explain why a player who had performed brilliantly during "try-outs" turned out, over time, to be only slightly better than average, subjects who had themselves been active in sports generally recognized the phenomenon of regression to the mean, explaining that on the day of the try-outs, the "brilliant" player had simply had a very good day, and that neither all nor even most days could be expected to be equally good. Students lacking experience in the domain of sports, in contrast, tended to prefer strictly causal explanations, suggesting, for example, that the brilliant player had begun to coast, or had other interests, or, perhaps, deliberately played more poorly than he might have, so as not to make the other players jealous [!].

Nisbett et al. conclude that training in statistics would be an effective means for improving probabilistic reasoning in every-day problems, since "...people's intuitive reasoning skills include strategies that may be called statistical heuristics. Formal training in statistics ...should represent less a grafting on of procedures than a refinement of preexisting ones."[66]

Some confirmation of this hypothesis has been found in studies by Fong, Krantz and Nisbett[67] on the effect of prior training in statistics on the tendency to reason probabilistically in explaining events, rather than to devise causal explanations. Subjects who had prior training in statistics tended to offer more probabilistic explanations of phenomenon than did those without prior training, and the answers of those with prior training were more often judged superior in quality.

Fong et al. provided subjects lacking any prior training with brief training sessions in which subjects were either given formal explanations and demonstrations of statistical theory, or, in the "modeling" version, given a set of "good answers", i.e., probabilistic answers, to a series of problems.[68] Improvement was also observed following both an abstract instruction about the appropriate rules, and a more concrete instruction in which the reasoning norms were demonstrated by examples. In addition, the particular type of problem used in the training seemed to unimportant. I.e., subjects who had received "objective" problems in the training session, and subjects who had received "subjective" problems, did equally well whether tested on objective or subjective problems. This suggests that the inferential rules learned were not domain-specific.

Holland et al. maintain that this lack of domain-specificity exists because the subjects already posses, and use "pragmatic reasoning schemas", i.e., sets of inferential rules for generating and evaluating rules of action.[69]

However, there is evidence to suggest that even individuals with considerable expertise in statistics are likely to make basic errors in probabilistic reasoning, when called upon to make judgments outside their area of expertise.[70] For example, when presented with the "flight instructor" problem, even graduate students with a background in statistics failed to recognize that the observations can be explained by the phenomenon of regression to the mean. Apparently, when faced with unfamiliar problems even individuals knowledgeable about probability fall back on the earlier, largely intuitive and often inadequate inferential strategies.

This suggests that if probabilistic reasoning with respect to every-day problems is to be improved, there is a need to give attention to such reasoning in a general educational context.

These studies of the effects of training in statistical inference[71] raise several interesting points. First, it seems that students are not only able to learn and correctly apply statistical rules, improving upon earlier intuitive beliefs, but are also able to "transfer" those rules, to apply them to dissimilar problems. That is, statistical inference does not appear to be domain-specific.[72]

Second, an improvement in probabilistic reasoning was observed following both instruction that focused on the presentation of abstract rules, and instruction that focused on concrete examples. Not surprisingly, instruction that used both strategies was found to be most effective.

5. Programs To Teach Probabilistic Reasoning

The fundamentals of probabilistic reasoning, e.g., the probability calculus, if taught at all, are generally taught in the context of courses on mathematics, or statistics, or in college level logic courses, or in advanced courses in decision theory. Thus, currently, most students would seem to be left without any introduction to the norms of probabilistic reasoning. And existing treatments of probability may leave even students introduced to the concepts with little understanding of the import of probability in informal, every-day contexts. Some texts present probabilistic reasoning as "puzzles", or "brain-teasers"[73]; such problems, while perhaps fun and fascinating in themselves, fail to impart the significance of the subject to every-day reasoning. One such "fun and educational" lesson plan for teaching probability is presented by Marilyn Burns. Although apparently intended to introduce third to fifth grade students to the concepts of chance and uncertainty, this game merely involves "investigating the probabilities of the sums that come up when two dice are tossed".[74] While the children observed did successfully discover that some outcomes are more likely than others, it seems doubtful that this realization would, in any major way, help them in solving the probabilistic problems of everyday life.

When the concepts and norms of probabilistic reasoning are given attention, it is primarily within the context of the study of mathematics. Indeed, Freudenthal argues that one chief advantage to the inclusion of probability in the mathematics curriculum is precisely that it serves to draw a <u>critical</u> response from the student, and engenders independent thinking.[75]

Alfred Renyi, however, advocates the teaching of probability in the elementary or secondary mathematics curriculum, but, not simply as an interesting challenging and useful branch of mathematics. Rather, he suggests that the study of probability can serve to improve the character of the students. He writes, "The study of probability teaches the student that clear logical thinking is of use also in situations when one is confronted with uncertainty (what is in fact the case in almost every practical situation). [Moreover,] it strengthens their courage if they understand that some failure may be due simply to chance and is no reason to give up some effort."[76]

A Curriculum to Improve Thinking Under Uncertainty

There is one notable exception to the general absence of explicit attention to probabilistic reasoning, in the work of Beyth-Marom and Dekel, of the Hebrew University of Jerusalem. Working with Kahneman, Beyth-Marom and Dekel have developed a curriculum, including a text (1985) and teacher's manual, An Elementary Approach to Thinking Under Uncertainty.[77] This curriculum is presently being used in Israel, with fourteen year-old students[78]; the effectiveness of the curriculum is currently being studied by Beyth-Marom and Dekel.

Unfortunately for elementary students and teachers in the U.S., the original text is in Hebrew. The text has been translated into English, but has at the same time also been revised and adapted "to include material more suited to American adults."[79] This English language version is designed to be "accessible to any adult with a minimal knowledge of arithmetic..."[80] I am unaware of any evaluative studies with respect to the effectiveness of the English language, adult-oriented version of this text.

Building on the psychological research into the existence and nature of systematic errors and biases in probabilistic judgments, Beth-Marom and Dekel set out two goals for the curriculum: "...showing students when and how their judgments are wrong and...presenting corrective procedures to improve their inductive reasoning."[81] The procedural means to this end is "...to make students' implicit thought processes explicit by getting them to talk about their own beliefs."[82]

Beyth-Marom and Dekel note that there are several major problems in improving probabilistic reasoning via such a course. First, the course must bring the student to recognize his or her existing intuitions with respect to probability, and to realize that what may seem a compelling and intuitively obvious inference may nevertheless be unwarranted. Second, the course must "convince students that probability theory is relevant to life events, instead of being just the 'science of coins and playing cards'."[83] And third, the course must, if it is to be effective in a general "critical thinking" context, avoid the rather complex mathematical treatment often characteristic of courses in statistics, probability theory and decision theory.[84]

<u>Summary</u>

In this chapter I have examined the research on the early development of probabilisitic reasoning in children, considered the origins of intuitive beliefs about probability, and hence the origins of the biases and fallacies of probabilistic reasoning referred to in the psychological literature.

According to Fischbein, children begin at an early age to recognize chance phenomena, but social pressures in schooling, particularly the insistence on certainty and the weight given to causal explanation interfere and cut short the development of sophisticated understandings of the workings of chance and probability.

Holland et al. theorize that individuals begin to develop sets of inferential rules, licensing certain sorts of inference and disallowing others, at an early age, and these sets include rules of probabilistic reasoning. The existence of a set of crude, preliminary rules, however faulty, that have been induced from experience is considered to be conducive to the later improvement of probabilistic reasoning through direct instruction in more appropriate rules. Holland et al. report that these inferential rules, unlike empirical rules, do not appear to be domain-specific.

Though conflicting evidence also exists, Nisbett et al. present evidence that training in statistics does lead to an improvement in probabilistic reasoning, even in everyday problems. The individual's reasoning is improved by instruction that helps the student to encode variability more effectively, and that helps the student to see analogies between everyday experiences and the aleatory systems in which the preliminary notions of probability were originally developed.

Finally, I have noted that most attention to "probabilistic reasoning" seems to be in the context of the mathematics curriculum. This is problematic in two ways. First, children not taking advanced mathematics courses would receive no introduction to the subject. And second, the in-depth mathematical treatment may not aid the students to "see the analogies" between mathematics and everyday decision making, nor would it aid them in more effectively encoding variability in their experience.

1. J. Piaget and B. Inhelder, The Origin of the Idea of Chance in Children (New York: Norton, 1975 (originally published, 1951).

2. P. Yost, A.E. Siegel and J.N. Andrews, "Non Verbal Probability Judgement by Young Children", in Child Development, 33 (1962), p. 778.

3. Susan Goldberg, "Probability Judgement by Preschool Children", in Child Development, 36, (1966), p. 159.

4. Carolyn M. Davies, "The development of the probability concept in children." Child Development 36 (1965), p. 779-788.

5. Arthur Engel, 'Teaching Probability in Intermediate Grades", in L. Rade (ed), The Teaching of Probability and Statistics
(Stockholm: Almqvist & Wiksell, 1970), p. 88.

6. E. Fischbein, The Intuitive Sources of Probabilistic Thinking in Children (Boston: D.Reidel Publishing Company, 1975), p. 66.

7. Ibid., p. 17.

8. Ibid., p. 4.

9. Ibid., p. 7-9.

10. Ibid., p. 7-9.

11. Ibid., p. 14.

12. Ibid., p. 18

13. Ibid., p. 17.

14. Ibid., p. 18.

15. Ibid., p. 3.

16. Ibid., p. 12.

17. Ibid., p. 12.

18. Piaget and Inhelder, The Origin of the Idea of Chance in Children.

19. Fischbein, The Intuitive Sources of Probabilistic Thinking in Children, p. 70.

20. Murray E. Jarvik, "Probability learning and a negative recency effect in the serial anticipation of alternative symbols', in Journal of Experimental Psychology, 41 (1951), pp. 291-297.

21. H. Lindman and W. Edwards, "Unlearning the gambler's Fallacy", in Journal of Experimental Psychology, 62 (1961), 630-631.

22. Ward Edwards, "Probability Learning in 1,000 Trials." Journal of Experimental Psychology 62 (1961), pp. 381-390.

23. Fischbein, The Intuitive Sources of Probabilistic Thinking in Children, p. 59.

24. L. Jonathan Cohen, "Can human irrationality be experimentally demonstrated?" The Behavioral and Brain Science 4 (1981) p. 327.

25. Daniel Kahneman and Amos Tversky, "Subjective Probability: A Judgement of Representativeness", in Judgment under Uncertainty: Heuristics and Biases (Cambridge: Cambridge University Press, 1982), pp. 32-47.

26. Morton W. Weir, "Effects of Age and Instruction on Children's Probability Learning", in Child Development 33 (1962), p. 730.

27. Morton W. Weir, "Developmental Changes in Problem Solving Strategies", in Psychological Review 71 (1964), pp. 473-490.

28. Morton W. Weir, "Children's Behavior in Probabilistic Tasks", in The Young Child: Reviews of Research, (National Association for the Education of Young Children, Washington, 1967), pp. 136-154.

29. Fischbein, The Intuitive Sources of Probabilistic Thinking in Children, p. 71.

30. E. Fischbein, Ileana Pampu and I. Minzat, "The Child's Intuition of Probability", in Fischbein, The Intuitive Sources of Probabilistic Thinking in Children, p. 156-174.

31. Ibid., p. 166.

32. Ibid.

33. Ibid., p. 169.

34. Bill Watterson, The Essential Calvin and Hobbes (Kansas City: Andrews and McMeel, 1988), p. 239.

35. Fischbein, The Intuitive Sources of Probabilistic Thinking in Children, p. 73.

36. Ojemann, R.H., Maxey, E.J., and Snider, B.C., 1965 'Effects of Guided Learning Experiences in Developing Probability Concepts at the Fifth Grade Level' in Perceptual and Motor Skills, v. 21, 415-427.

37. Ojemann, R.H., Maxey, E.J., and Snider, B.C., "The Effect of a Program of Guided Learning Experiences in Developing Probability Concepts at the Third Grade Level", in Journal of Experimental Education, 33, 321-330.

38. Ojemann, R.H., Maxey, E.J., and Snider, B.C., "Further Study of Guided Learning Experience in Developing Probability Concepts in Grade Five", in Perceptual and Motor Skills, 23 (1966), 97-98.

39. Harold R. Keller, "Children's Acquisition and Reversal Behavior in a Probability Learning Situation as a Function of Programed Instruction, Internal-External Control and Schedules of Reinforcement", Journal of Experimental Child Psychology, 11 (1971), 281-295.

40. Ruma Falk, Raphael Falk, and Iris Levin, "A Potential for Learning Probability in Young Children" Educational Studies in Mathematics 11 (1980), pp. 181-204.

41. Ibid., p. 199.

42. Ibid., p. 201.

43. Richard Nisbett and Lee Ross, Human Inference: Strategies and Shortcomings of Social Judgment (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1980), p. 228.

44. John H. Holland, Keith J. Holyoak, Richard E. Nisbett and Paul R. Thagard, Induction: Processes of Inference, Learning, and Discovery (Cambridge, Massachusetts: The MIT Press, 1989), p. 257.

45. Ibid., p. 14.

46. Ibid., p. 15.

47. Ibid., p. 16.

48. Ibid., p. 16.

49. Holland et al.'s term, intended to include possible non-living thinking beings.

50. Ibid., p. 14.

51. Ibid., p. 43.

52. Ibid., p. 43.

53. Ibid., p. 44.

54. Ibid., p. 45.

55. Ibid., p. 16.

56. Ibid., p. 126.

57. Ibid., p. 230-236.

58. Ibid., p. 285.

59. Ibid., p. 17.

60. Ibid., p. 17.

61. Richard E. Nisbett, David H. Krantz, Christopher Jepson and Ziva Kunda, "The Use of Statistical Heuristics in Everyday Inductive Reasoning", Psychological Review 90:4 (1983), pp. 339-363.

62. Nisbett et al., p. 339.

63. Nisbett, et al., p. 344.

64. Nisbett, et al., p. 340.

65. Nisbett, et al., p. 348-354.

66. Nisbett, et al., p. 357.

67. Fong, Krantz and Nisbett, "Improving inductive reasoning through statistical training". Unpublished manuscript, University of Michigan, 1983, discussed in Nisbett, Krantz, Jepson and Kunda, p. 358-359.

68. Ibid.

69. Holland et al., Induction: Processes of Inference, Learning and Discovery, p. 262.

70. Ibid., p. 265.

Amos Tversky and Daniel Kahneman, "Belief in the Law of Small Numbers" Psychological Bulletin 76 (1971), pp. 105-110.

71. Holland et al., Induction: Processes of Inference, Learning and Discovery, p. 257.

72. G. T. Fong, D.H. Krantz and R.E. Nisbett, "The Effects of Statistical Training on Thinking about Everyday Problems", Cognitive Psychology 18 (1986).

73. For instance, Frederick Mosteller, Fifty Challenging Problems in Probability, with Solutions (New York: Dover Publications, Inc., 1965).

74. Marilyn Burns, "Dice Advice", Instructor (January, 1988), p. 26-29.

75. Hans Freudenthal, "The Aims of Teaching Probability", in L. Rade (ed.), The Teaching of Probability and Statistics, (Stockholm: Almqvist and Wiksell, 1970), pp. 151-168.

76. Alfred Renyi, "Remarks on the Teaching of Probability", in The Teaching of Probability and Statistics, ed. by L. Rade, p. 274.

77. Ruth Beyth-Marom and Shlomith Dekel, An Elementary Approach to Thinking Under Uncertainty (Hillsdale, N.J.: Lawrence Erlbaum Associates, Publishers, 1985).

78. Ibid., p. ix.

79. Ibid..

80. Ibid.

81. Ruth Beyth-Marom and Shlomith Dekel, "A Curriculum to improve Thinking under Uncertainty", Instructional Science, 12 (1983), p. 69.

82. Ibid., p. 70.

83. Ibid., p. 71.

84. Ibid., p. 71

CHAPTER V

SUMMARY AND CONCLUSION

1. Introduction

Given the continuing controversy with respect to both the empirical claims about how people _do_ reason when making judgments of probability, the normative claims about how people _should_ reason, and the variety of conceptual interpretations of both rationality and critical thinking, what sort of conclusions can be drawn about the teaching of probabilistic reasoning?

The first point that should be noted is that, although the formulas pertinent to probabilistic reasoning, i.e., the basics of the probability calculus and Bayes' theorem, are relatively simple to apprehend and to manipulate, major problems in probabilistic reasoning arise when individuals try to apply those simple formulas in appropriate ways to complex real-life problems. The principal educational problem, then, has to do _not_ with simply teaching students the formulas, the "math" of probability, but with teaching the meaning and significance of probability and chance, and instilling a critical attitude toward the numbers that crop up when one attempts to deal with probabilistic information.

2. Summary and Conclusions

In the first chapter, I have set out and briefly discussed the notion of probabilistic reasoning, and suggested that there is an important pedagogical connection between teaching for critical thinking and teaching probabilistic reasoning.

In chapter II, I have discussed a range of current conceptions of the term 'critical thinking', focusing primarily on the work of two theorists, Siegel and McPeck, whose views differ radically.

I have argued that teaching the norms of probabilistic reasoning would contribute to the students' development of the "critical attitude" required by McPeck, viz., the attitude of judicious skepticism which leads one to "suspend belief in the available evidence", since those norms would provide students with a set of appropriate questions to ask about that evidence, and a means to estimate the value of the evidence. Moreover, since the norms of probabilistic reasoning are themselves still rather controversial in some respects, students would be led to "critically" examine the very criteria of judgment that they accept, thus meeting another of McPeck's requirements for "critical thinking". Further, a study of probabilistic reasoning would lead the student to question the validity of his or her own "intuitively obvious" beliefs, since the study of norms which are at times strongly counter-intuitive would demonstrate that "intuition" can easily lead one to make erroneous judgments.

Second, I have that, if 'critical thinking' is interpreted as "skilled" thinking, and if one takes the demonstration of "skilled performance" to involve a knowledgeable application of a set of appropriate norms, then, again, the study of probabilistic reasoning would be a useful addition to the critical thinking course. I have also argued that other interpretations of the "thinking skills" metaphor provide a less than adequate basis for the interpretation of critical thinking.

Third, if critical thinking is taken to require logical thinking, and if 'logical thinking' is understood to refer to both deductive and inductive reasoning, then attention to probabilistic reasoning, which is basic to the evaluation of inductive arguments, again, would be required. I have also argued, contra McPeck, that logical norms, whether of deduction or induction, are not inherently domain-specific. The psychological question about transfer of learning across different domains can only be resolved by empirical evidence, but, there does exist evidence, in the work of Holland et al., indicating that knowledge of inferential rules of probabilistic reasoning does transfer across disciplines.

Fourth, I have argued that, to perform McPeck's "evaluation of knowledge claims", students would require an acquaintance with the norms of probabilistic reasoning, since such evaluation requires estimating the probability that one's premises are in fact true, and an evaluation of the significance of the evidence in the premises, i.e., an estimate of the conditional probability of the truth of the conclusion, given the truth of the premises.

Fifth, I have discussed and critiqued the interpretations of "rationality", and the proposed relation between rationality and critical thinking, offered by Siegel and McPeck. I have argued that, at the least, "rationality" would imply acceptable reasoning in problems that require induction, i.e., in "everyday" problems. And, as previously noted, to reason well inductively requires the understanding and application of norms of probabilistic reasoning.

Hence, given any of the common interpretations of the term 'critical thinking', probabilistic reasoning would be either an appropriate or a required element in a course designed to teach critical thinking.

In chapter III, I have turned from the conceptual question to the more practical question, viz., Is there any need for instruction in the norms of probabilistic reasoning? This discussion falls into two parts. First, I have set out and discussed the range of interpretations of probability that might be invoked in estimating probability, and indicated that the theoretical questions are not insignificant, and do affect the results of one's estimates. I have then discussed the common reasoning activities that involve assessments of probability in some way, e.g., predicting, generalizing, making causal attributions. These activities would seem to be a fundamental part of what is termed 'everyday reasoning'. And, it is just such everyday reasoning that most critical thinking theorists agree must be improved.

Second, I have set out and discussed the evidence in the psychological literature which suggests that people commonly do err in this sort of everyday reasoning, and that the errors are directly linked to errors in probabilistic reasoning. I have also noted that that research is not uncontroversial, and that a variety of different explanations of the origin, and hence the meaning, of the "errors" are possible. Nevertheless, the bulk of the research to date does indicate that this sort of reasoning does present a problem for many. It would seem that regularly falling into Cohen's "cognitive illusions" would be as disadvantageous, practically, as would the regular reliance on Kahneman and Tversky's "faulty intuitive heuristics". This research suggests, then, that the teaching of probabilistic reasoning would meet an existing need.

In chapter IV, I have discussed the controversy surrounding the identification of appropriate norms of probabilistic reasoning. Much of the psychological research introduced in the previous chapter is grounded on the assumption that Bayesian reasoning is normative, and given this, any divergence from Bayesian analyses would count as error. But this assumption has been challenged by Cohen, who argues that in some contexts it is Bayesian reasoning that should be counted as irrational. I have also set out the position taken by Mill on this issue, and given an example of a variant on the classic Kahneman and Tversky "Cab problem" that seems to support the Cohen critique. But, I have concluded that it is not the application, but rather the misapplication, of Bayesian reasoning to this scenario that creates the difficulties. It seems that "available data" is plugged into the formula by Kahneman and Tversky, without sufficient justification. This illustrates the points raised above, that learning the norms of "rational" probabilistic reasoning would involve students not only in learning a set of appropriate procedures, but also in critically evaluating those procedures and their application.

In chapter V, I have discussed issues related to the teaching and learning of notions of probability. I have begun with a review of psychological literature on the early origin of intuitions about probability, and then discussed the possible benefits to be derived from the direct teaching on probabilistic matters, e.g., the effect on everyday reasoning of statistics courses. Finally, I have discussed the one curriculum that has been devised to teach probabilistic reasoning on the elementary level.

In short, I conclude: a) that probabilistic reasoning is an appropriate, and/or required component of the critical thinking course, depending on the conception of critical thinking adopted; b) that errors in probabilistic reasoning do commonly occur, and these errors are significant, since they affect significant "everyday" reasoning activities; c) that philosophical controversies exist with respect both to the interpretation or interpretations of probability, and to the particular norms of probabilistic reasoning that are most appropriate, but that the existence of such controversy makes the subject of more, not less, value in the critical thinking course; and, finally d) that students' reasoning in probabilistic matters is amenable to improvement via instruction in a classroom setting.

Hence I conclude that probabilistic reasoning should be incorporated in the critical thinking curriculum.

3. Probabilistic Reasoning in A Critical Thinking Course: Goals

There would seem to be three distinct goals that ought to be pursued in such a program. The first goal would be to provide the student with an **understanding** of the concept of probability. The student should acquire propositional knowledge as to what probability *is*, and should learn to recognize the element of probability in common problem situations.

The second goal would be for the student to acquire some knowledge as to the **norms** of probabilistic reasoning. The student should have some idea as to what "good" probabilistic reasoning is, and should acquire through practice some skill in probabilistic reasoning, i.e., a "knowing how", or performative knowledge.

The third goal of such a program would be to give the student an appreciation of the concept of probability. I.e., the student should not only acquire propositional knowledge and performative knowledge, but should also become aware of the significance of the use, and misuse, of probabilistic reasoning in the "real world" (which might be called a "knowing 'so what?'", i.e., an understanding of the significance and usefulness of that which has been learned). The student would then (one would hope) come to develop a particular attitude, viz., a willingness to engage in critical probabilistic reasoning. In the long term (again, one would hope), the student would acquire a disposition to reason critically.

The objective of this instruction would be, first, to make students aware of the probabilistic element common in everyday judgments; b) to make students aware of the often counter-intuitive rules which are central to good probabilistic reasoning; and c) to encourage the student to make practical use of such reasoning. This instruction would:

a) bring the student to a conscious awareness of seemingly plausible patterns of inference that he or she may habitually use; b) illustrate the errors that can result from the use of faulty inferential strategies, and c) offer the student more useful inferential strategies, and provide practice in using them.

Students should become aware of the common occurrence of problems that require such reasoning, and the usefulness in every-day thinking of the an understanding of the notions of chance and probability. Second, students should recognize that there are norms by which probabilistic reasoning, one's own as well as that of others, can be judged. Such norms apply equally well both to explicit calculations of probabilities and to informal, rough estimations. Third, the students should know what those norms are, and understand them well enough to be able to reason both explicitly and informally in accordance with then norms.

Fourth, the students should realize that there are subtle snares and pitfalls in this sort of reasoning (just as there are in deductive and informal arguments). There are, for instance, out and out fallacies--invalid inferences that nevertheless seem quite plausible at first sight. The student should be able to avoid such fallacious inferences, and spot them when they occur in other's arguments, whether formal or informal. (And it seems that it is in informal arguments that such errors would be hardest to spot and or avoid.) Such fallacies of probabilistic reasoning include, for instance, the assumption that $Pr(e/H) = Pr(H/e)$, and the assumption that $Pr(e/H) + Pr(e/-H) = 1.0$.

There are other problems, in addition to the making of clearly fallacious inferences, that students should come to recognize and avoid. For instance, problems may arise in choosing an appropriate reference class, in accepting a "handy" but inappropriate number as a "prior probability". One may also err in failing to recognize that the information one needs to make an inference is not in fact available, and cannot or shouldn't be merely estimated. One should know when to conclude that a problem cannot be solved with the information given.

Fifth, the student should develop a set of attitudes and dispositions that will lead them to actually engage in "good" probabilistic reasoning, to scrutinize and critique the arguments others may offer, and to extend this favor to their own work.

4. Some Possible Objections

The most obvious objection to the proposal to incorporate instruction in probabilistic reasoning in critical thinking programs would, I think, be the assertion that people do not in fact make errors in probabilistic reasoning, or that if they do, such errors are themselves a matter of chance, have no systemic pattern, and are not amenable to improvement. These, of course, are empirical questions, and can only be answered by reference to empirical research. Given the body of research to date, it would seem that in at least some experimental situations, the errors described do occur.

A second possible objection might be that, although errors do occur in probabilistic reasoning, those errors are in some sense unimportant. Although it is beyond the scope of this paper to detail the full array of problems calling for probabilistic reasoning, I think that from even the few examples given it is evident that probabilistic reasoning has a role to play in deliberations about quite common, every-day problems; and these problems, although "ordinary", nevertheless have significant implications. It should perhaps be noted again that I am not suggesting that there is a need for children to become sophisticated statisticians. Rather, my contention is that an improvement in the basic level of understanding of probability and the effects of chance would aid in day-to-day decision-making.

A third objection might be that "judgment" cannot be reduced to the manipulation of numbers, or that an attempt to "make judgments statistically" would discourage the very "critical" attitude that is desired. There are several responses to this objection. First, the applicability of probabilistic reasoning is limited to problems involving either chance or insufficient evidence, i.e., uncertainty. There may be significant problems for which such reasoning would be inappropriate, but, when chance or uncertainty exist, one has no alternative but to make assessments of degrees of probability--the only question is how those judgments will be made, whether one's estimates of the "numbers" will be justifiable or not.

Second, it must be acknowledged that, even for problems involving the element of chance, it may not be possible to come up with equally acceptable numbers to represent the probabilities involved. It is easy, for instance, to establish an acceptable figure for the probability of rolling double sixes; it is quite another matter to establish the "probability" of a nuclear meltdown, or the probability of success in one job rather than another.

In the first case, one can appeal to any of several straightforward "objective" theories of probability, e.g., the frequency interpretation, the classical interpretation; in the second case, one must rely on some "subjective" theory of probability, i.e., probability as degree of rational belief.

Predictions of this second sort involve judgment, and educated estimation[1], and this element may well be irreducible in any given problem. Nevertheless, the sorts of errors and biases that plague simpler probabilistic problems remain. Teaching about probabilistic reasoning will not eliminate the need to make qualitative estimation of probabilities, but it seems that it would at least improve those judgments, and any reasoning based on such probabilities.

And, finally, one might object that present programs in informal logic adequately address the problem, and there is thus no need to focus specific attention on probabilistic reasoning. For instance, one might point out that the "Gambler's Fallacy", discussion of which is often included in informal logic courses, involves a probabilistic reasoning error, or, similarly, that when one teaches students to avoid the logical fallacy of "Hasty Generalization" or "Sweeping Generalization" one is, in effect, teaching probabilistic reasoning. In a limited sense, this is true; however, merely teaching students to recognize that probabilistic errors can occur and ought to be avoided is not enough to improve probabilistic reasoning in practice. A program in probabilistic reasoning would need to provide students with some general criteria by which to evaluate their generalizations, predictions, causal attributions, etc., and with practice in making such judgments.

5. Recommendations for Further Investigation

Further investigation is required of both the empirical and the philosophical issues relevant to probabilistic reasoning. Empirically, much more could be discovered, first, about how people do in fact reason about probabilistic matters, prior to instruction, and second, about how successful instruction is, or can be, in improving probabilistic inference.

In addition, much remains to be said about the philosophical issues, i.e., about the interpretation(s) of the concept of probability, about the conceptualization of rationality, and about the evaluation of the strength of inductive reasoning and the justification of inductive inference. And finally, more could be done to sort out and evaluate the merits of the various interpretations of the concept of critical thinking.

BIBLIOGRAPHY

Bar-Hillel, Maya. "The Base-rate Fallacy in Probability Judgments." Acta Psychologica, 44 (1980), 212.

Bartlett, Sir Frederic. Thinking: An Experimental and Social Study. New York: Basic Books, Inc., Publishers, 1958.

Bayes, Thomas. "Essay towards Solving a Problem in the Doctrine of Chances." Philosophical Transactions of the Royal Society of London 53 (1763), 371-418.

Beyth-Marom, Ruth, and Dekel, Shlomith. An Elementary Approach to Thinking Under Uncertainty. Hillsdale, N.J.: Lawrence Erlbaum Associates, Publishers (1985).

Beyer, Barry K. Developing a Thinking Skills Program. Boston: Allyn and Bacon, 1988.

Beyer, Barry K. Teaching Thinking Skills: A Handbook for Secondary School Teachers. Boston: Allyn and Bacon, 1991.

Birnbaum, Michael H. "Base Rates in Bayesian Inference: Signal Detection Analysis of the Cab Problem." American Journal of Psychology 96:1 (Spring, 1983), pp. 85-94.

Marilyn Burns, "Dice Advice", Instructor (January, 1988), p. 26-29.

________. The Book of Think (Or, How to Solve a Problem Twice Your Size. Boston: Little, Brown and Company, 1976.

Charlton, Ronald E. "The Direct Teaching of Analysis." Thinking Skills Instruction: Concepts and Techniques. Edited by Marcia Heiman and Joshua Slomianko. Washington, D.C.: National Education Association, 1987.

Costa, Arthur L., ed. Developing Minds: A Resource Book for Teaching Thinking. Association for Supervision and Curriculum Development, 1985.

Cohen, L. Jonathan. "Can Human Irrationality Be Experimentally Demonstrated?" The Behavioral and Brain Sciences 4 (1981), 323.

________. An Introduction to the Philosophy of Induction and Probability. Oxford: Clarendon Press, 1989.

Black, Max. Margins of Precision: Essays in Logic and Language. Ithaca, N.Y.: Cornell University Press, 1970.

Davies, Carolyn M. "The Development of the Probability Concept in Children." Child Development 36 (1965) 779-788.

De Bono, Edward. Teaching Thinking. London: Temple Smith, 1976.

Edwards, Ronald R.; and Cook, Wanda D. Problem Solving through Critical Thinking, Grades 5-8. New Rochelle, N.Y.: Cuisenaire Company of America, Inc., 1990.

Edwards, Ward. "Probability Learning in 1,000 Trials." Journal of Experimental Psychology 62 (1961), pp. 381-390.

________. "Conservatism in human information processing." Formal Representation of Human Judgment. Edited by B. Klienmuntz. New York: Wiley, 1967.

Engel, Arthur. "Teaching Probability in Intermediate Grades." The Teaching of Probability and Statistics. Edited by L. Rade. Stockholm: Almqvist & Wiksell, 1970.

Falk, Ruma; Falk, Raphael, and Levin, Iris. "A Potential for Learning Probability in Young Children." Educational Studies in Mathematics 11 (1980), pp. 181-204.

Fischbein, E. The Intuitive Sources of Probabilistic Thinking in Children. Boston: D. Reidel Publishing Company, 1975.

Fischbein, E., Pampu, Ileana, and Minzat, I. "The Child's Intuition of Probability." The Intuitive Sources of Probabilistic Thinking in Children. E. Fischbein.

Boston: D. Reidel Publishing Company, 1975, pp. 156-174.

Fong, G.T., Krantz, D.H., and Nisbett, R.E.. "The Effects of Statistical Training on Thinking about Everyday Problems", Cognitive Psychology 18 (1986).

Freudenthal, Hans. "The Aims of Teaching Probability." The Teaching of Probability and Statistics. Edited by L. Rade. Stockholm: Almqvist and Wiksell (1970), pp. 151-168.

Gigerenzer, Gerd; and Murray, David, J. Cognition as Intuitive Statistics. Hillsdale, N.J.: Lawrence Erlbaum Associates, Publishers, 1987.

________, Swijtink, Zeno; Porter, Theodore; Daston, Lorraine; Beatty, John; Kruger, Lorenz. The Empire of Chance: How Probability Changed Science and Everyday Life. Cambridge: Cambridge University Press, 1989.

Ginosaur Zvi, and Trope, Yaacov. "The effects of base rates and individuating information on judgments about another person." Journal of Experimental Social Psychology 16 (1980), 228-242.

Goldberg, Susan. "Probability Judgement by Preschool Children." Child Development 36, (1966), 157.

Heiman, Marcia; and Slomianko, Joshua. eds. Thinking Skills Instruction: Concepts and Techniques. Washington, D.C.: NEA Professional Library, 1987.

Holland, John H.; Holyoak, Keith J.; Nisbett, Richard E.; and Thagard, Paul R. Induction: Processes of Inference, Learning, and Discovery. Cambridge, Massachusetts: The MIT Press, 1989.

Hume, David. A Treatise on Human Nature, Book 1, Part 3. Edited by L.A. Selby-Bigge. Oxford: Clarendon Press, 1888.

Hyde, Arthur A.; and Bizar, Marilyn Thinking in Context: Teaching Cognitive Processes across the Elementary School Curriculum. New York: Longman Inc., 1989.

Jarvik, Murray E. "Probability learning and a negative recency effect in the serial anticipation of alternative symbols." Journal of Experimental Psychology 41 (1951), 291-297.

Kahneman, Daniel; Slovic, Paul; and Tversky, Amos, eds.. Judgment Under Uncertainty: Heuristics and Biases. Cambridge: Cambridge University Press, 1982.

________, and Tversky, Amos. "Intuitive Prediction: Biases and Corrective Procedures." Judgment Under Uncertainty: Heuristics and Biases. Cambridge: Cambridge University Press, 1982.

________. "On the Psychology of Prediction." Judgment under Uncertainty: Heuristics and Biases. Cambridge: Cambridge University Press, 1982.

Karmos, Joseph S; and Karmos, Ann H. "Strategies for active Involvement in Problem-solving." Thinking Skills Instruction: Concepts and Techniques. Edited by Marcia Heiman and Joshua Slomianko. Washington, D.C.: National Education Association, 1987.

Keller, Harold R. "Children's Acquisition and Reversal Behavior in a Probability Learning Situation as a Function of Programed Instruction, Internal-External Control and Schedules of Reinforcement." Journal of Experimental Child Psychology 11 (1971), 281-295.

Levinson, Horace C. Chance, Luck and Statistics. New York: Dover Publications, 1963.

Lindman, Harold, and Edwards, Ward. "Unlearning the Gambler's Fallacy." Journal of Experimental Psychology, 62 (1961), 630-631.

Lipman, Matthew. Thinking in Education. Cambridge: Cambridge University Press, 1991.

McPeck, John E. Critical Thinking and Education. Oxford: Martin Robertson & Co. Ltd., 1981.

________. "Critical Thinking Without Logic: Restoring Dignity to Information." Proceedings of the Thirty-seventh Annual Meeting of the Philosophy of Education Society. 1981.

________. Teaching Critical Thinking: Dialogue and Dialectic. New York: Routledge, Chapman and Hall, 1990.

Mill, John Stuart. A System of Logic: Ratiocinative and Inductive. London: Longmans, Green, and Co., 1893.

Mosteller, Frederick. Fifty Challenging Problems in Probability, with Solutions. New York: Dover Publications, Inc., 1965.

Neimark, Edith D. Adventures in Thinking. San Diego: Harcourt Brace Jovanovich, Publishers, 1987.

Nickerson, Raymond S.; Perkins, David N.; and Smith, Edward E. The Teaching of Thinking. Hillsdale, N.J.: Lawrence Erlbaum Associates, Publishers, 1985.

Nisbett, Richard E.; Borgida, Eugene; Crandall, Rick; and Reed, Harvey. "Popular Induction: Information Is Not Necessarily Informative", in Judgment under Uncertainty: Heuristics and Biases. Cambridge: Cambridge University Press, 1982.

________; Krantz, David H.; Jepson, Christopher, and Kunda, Ziva. "The Use of Statistical Heuristics in Everyday Inductive Reasoning." Psychological Review 90:4 (1983), pp. 339-363.

Nisbett, Richard E.; and Ross, Lee. Human Inference: Strategies and Shortcomings of Social Judgment. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1980.

Nisbett, Richard E.; Krants, D.H.; Jepson, D.; and Kunda, Z. "The use of statistical heuristics in everyday inductive reasoning". Psychological Review, 84 (1983), 339-363.

Norris, Steven P. "Thinking about critical thinking: philosophers can't go it alone." McPeck, John E. Teaching Critical Thinking: Dialogue and Dialectic. New York: Routledge, Chapman and Hall, 1990.

Ojemann, R.H., Maxey, E.J., and Snider, B.C.. "Effects of Guided Learning Experiences in Developing Probability Concepts at the Fifth Grade Level." Perceptual and Motor Skills 21 (1965), 415-427.

________. "Further Study of Guided Learning Experience in Developing Probability Concepts in Grade Five." Perceptual and Motor Skills 23 (1966), 97-98.

________. "The Effect of a Program of Guided Learning Experiences in Developing Probability Concepts at the Third Grade Level." Journal of Experimental Education 33, 321-330.

Peterson, Cameron R., and Beach, L. R. "Man as an Intuitive Statistician", Psychological Bulletin 68 (1967) 29-46.

Peterson, Cameron R., and Ducharme, W. M. "A Primacy Effect in Subjective Probability Revision." Journal of Experimental Psychology, 73 (1967), 61-65.

Phillips, Lawrence D., and Edwards, Ward. "Conservatism in a simple probability inference task." Journal of Experimental Psychology 72 (1966), 346-354.

Piaget, Jean, and Inhelder, Barbel. The Origin of the Idea of Chance in Children. New York: Norton, 1975. (originally published, 1951).

Popper, Karl R. Realism and the Aim of Science. London: Hutchinson, 1959.

Pratte, Richard. Philosophy of Education: Two Traditions. Springfield, Illinois: Charles C Thomas, Publisher, 1992.

Quattrone, G.; and Jones, E. "The perception of variability within in-groups and out-groups: Implications for the law of large numbers". Journal of Personality and Social Psychology, 38, 141-152.

Renyi, Alfred. "Remarks on the Teaching of Probability." The Teaching of Probability and Statistics. Edited by L. Rade. Stockholm: Almqvuist and Wiksel (1970), p. 274.

Ruggiero, Vincent Ryan. Beyond Feelings: A Guide to Critical Thinking. Palo Alto, CA.: Mayfield Publishing Company, 1984.

Ruggiero, Vincent Ryan. The Art of Thinking: A Guide to Critical and Creative Thought. New York: Harper & Row, Publishers, 1984.

Segal, Judith W.; Chipman, Susan F.; and Glaser, Robert, eds. Thinking and Learning Skills, vols. 1 and 2. Hillsdale, N.J.: Lawrence Erlbaum Associates, Publishers, 1985.

Seymour, Dale and Beardslee, Ed. Critical Thinking Activities in Patterns, Imagery and Logic. Palo Alto, CA: Dale Seymour Publications, 1990.

Siegel, Harvey. Educating Reason: Rationality, Critical Thinking and Education. New York: Routledge, 1988.

________. "McPeck, informal logic, and the nature of critical thinking." McPeck, John E. Teaching Critical Thinking: Dialogue and Dialectic. New York: Routledge, Chapman and Hall, 1990.

Slovic, Paul; Fischoff, Baruch; and Lichtenstein, Sarah. "Facts versus fears: Understanding perceived risk." Judgment under Uncertainty: Heuristics and Biases. Edited by Daniel Kahneman, Paul Slovic and Amos Tversky. Cambridge: Cambridge University Press, 1982.

Tversky, Amos, and Kahneman, Daniel. "Availability: A Heuristic for Judging Frequency and Probability." Judgment Under Uncertainty: Heuristics and Biases. Edited by Daniel Kahneman, Paul Slovic and Amos Tversky. Cambridge: Cambridge University Press, 1982.

________. "Belief in the Law of Small Numbers." Psychological Bulletin 76 (1971), pp. 105-110.

________. "Evidential Impact of Base-rates." Judgment under Uncertainty: Heuristics and Biases. Cambridge: Cambridge University Press, 1982.

Weir, Morton W. "Children's Behavior in Probabilistic Tasks." The Young Child: Reviews of Research. National Association for the Education of Young Children. Washington, 1967, pp. 136-154.

Weir, Morton W. "Developmental Changes in Problem Solving Strategies." Psychological Review 71 (1964), pp. 473-490.

________. "Effects of Age and Instruction on Children's Probability Learning." Child Development 33 (1962), p. 729-735.

Yost, Patricia, Siegel, Alberta Engvall, and Andrews, Julia McMichael. "Non Verbal Probability Judgement by Young Children". Child Development 33 (1962), p. 778.

Young, Robert E., ed., Fostering Critical Thinking: New Directions for Teaching and Learning. San Francisco: Jossey-Bass Inc., Publishers, 1980.