Decisions and revisions

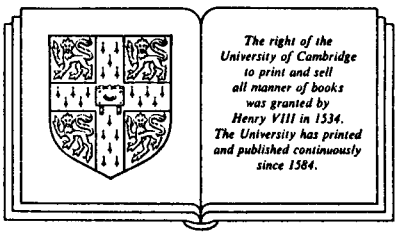
Philosophical essays on knowledge and value

Decisions and revisions

Philosophical essays on knowledge and value

ISAAC LEVI

PROFESSOR OF PHILOSOPHY,
COLUMBIA UNIVERSITY



CAMBRIDGE UNIVERSITY PRESS

CAMBRIDGE

LONDON NEW YORK NEW ROCHELLE
MELBOURNE SYDNEY

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 2RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521254571

First published 1984
This digitally printed first paperback version 2006

A catalogue record for this publication is available from the British Library

Library of Congress Catalogue Card Number: 83–26355

ISBN-13 978-0-521-25457-1 hardback
ISBN-10 0-521-25457-4 hardback

ISBN-13 978-0-521-02762-5 paperback
ISBN-10 0-521-02762-4 paperback

TO ERNEST NAGEL

Contents

Preface

Scientists must make value judgements. The conduct of their diverse activities requires this of them. The kernel of truth in the contrary claim that scientific inquiries are value-free or value-neutral is more accurately expressed by asserting that scientific interests are autonomous. The values which ought to be promoted in scientific inquiries exhibit features distinctive of scientific as opposed to political, economic, moral or aesthetic deliberations. This way of speaking admits that scientific inquiry is value-laden, that every specific scientific inquiry addresses some problem or family of problems and is, as a consequence, goal-oriented, and that the solution which ought to be adopted from among the strategies available is one which best promotes the goals of inquiry.

To concede the value-laden character of scientific inquiry is not to grant the reducibility of the goals and values of scientific inquiry to moral, political, prudential and aesthetic goals and values. The reconstructed version of value-neutrality I favor denies this reductionist view and insists that scientific inquiries seek or ought to seek to promote values and goals distinctive of the scientific enterprise.

The autonomy of scientific values is compatible with the plurality of human values. Scientists are human beings with personal goals, plans and moral commitments distinct from scientific interests which sometimes compete with the research aims that some human beings also have. Universities, research institutes and other agencies of scientific research have economic, political, educational or other interests distinct from cognitive ones.

But cognitive goals are distinct from moral, political, aesthetic or prudential objectives and may even come into conflict with them. In this sense, conflict between cognitive and other interests can generate moral struggles of the second kind noted by Dewey and Tufts (1932, p. 174) just as surely as can conflicts between competing moral values or between moral values, aesthetic values or life plans.

Insisting on the autonomy of scientific values may be thought to create an untenable dualism between theory and practice. I think not.

That one and the same person or institution may pursue diverse aims and values on different occasions or might be conflicted between different values on the same occasion is generally conceded. There is, to be sure, considerable

controversy concerning how conflict in values is to be understood and addressed. Most of those who write about such conflicts assume that they should be settled (at least ideally speaking) by the time the moment of choice is reached or that they should be settled through choice made at that moment.

In my judgement, such views are mistaken. Rational agents may and sometimes should refuse to resolve the conflicts between the rival aims and values they are committed to uphold even while they are constrained to make a decision. This is one of the main messages of 'Conflict and social agency' (1982*e*) which is reprinted as chapter 16 of this volume.

No matter how one stands on the question of when conflicts between values should be resolved, we should not deny the existence of such conflicts altogether on the grounds that they presuppose untenable dualisms. Perhaps, we would be hard pressed to draw a precise and general distinction between aesthetic and moral values. Even so, artists are sometimes in conflict between moral commitments and artistic ideals. It would be wrong to deny the existence of such conflict on the grounds that it presupposed an untenable dualism.

The difference between cognitive and other non-cognitive values should not be thought to create a chasm between theory and practice any more or less grand than any other contrast between kinds of values which might come into conflict.

There is a more important point about theory and practice which ought to be emphasized. Once we have acknowledged the value and goal-directed character of scientific inquiry and assessment of evidence, we are in a position to emphasize that criteria for assessing evidence and exercising cognitive options in science share much in common with grounds for evaluating options in other contexts of deliberation. In this respect, there is a unity of theory and practice implied by the view I am considering which could not have been emphasized so readily if we did not acknowledge at the outset a contrast between cognitive and non-cognitive values.

To sustain a view such as this, it is important to explore how a conception of the cognitive values of scientific inquiry combined with a view of rational goal attainment might produce criteria for evaluating potential solutions to problems in various contexts of inquiry. I began considering this question in the late 1950s in a relatively limited way. Subsequently, I have broadened the scope of my reflections in several directions. Over the past two and half decades my views concerning epistemic utility, cognitive decision making and the improvement of knowledge have gone through many minor and some major changes. *Gambling with Truth* (1967*a*) reports on what was my considered view *circa* 1965 when I completed the manuscript of that book. *The Enterprise of Knowledge* (1980*a*) is another snapshot of my evolving doctrine as of 1978.

However, from 1960 onwards, I published several articles on epistemic

utility and cognitive decision making which fill in many of the gaps in the developments which led in the first instance to (1967*a*) and then to (1980*a*). The first series of essays in this collection document various changes in outlook from 1960 to 1978 and may, when taken together with the two books, give a fuller picture of the various considerations which have led to the account of cognitive decision making I favor.

'Must the scientist make value judgements?' (1960*a*) was written in 1958. It points to the possibility of understanding statistical inferences in particular and scientific inferences in general as cognitive decisions justified in the context of cognitive decision problems relative to epistemic goals. 'On the seriousness of mistakes' (1962) was written in 1959. It proposes a pair of models for decision making in statistics together with epistemic utility functions suited to the models. In this paper, I introduced the notion of a degree of caution understood as reflecting a trade-off between an interest to avoid error and relieve doubt.

These papers were written before I was acquainted with Hempel (1960) and (1962), in which he makes use of epistemic utilities. In my subsequent work, I borrowed his term although I had already used the idea independently of him.

Aside from terminology, there was one other point suggested to me by Hempel's papers – namely that the trade-off between interest in avoiding error and relieving doubt might be clarified by representing the extent to which doubt is relieved by a measure of content or information. I did not think that Hempel had exploited this idea in a useful way. However, it did seem like a good idea to determine whether Popper's ideas on falsifiability and content could be of any help.

With this in mind, I hit upon the idea of trying to interpret Popper's measures of corroboration as discussed in the appendices of (Popper, 1959) as measures of expected epistemic utility which one seeks to maximize in order to determine which of rival hypotheses to accept. I soon realized that this could not be done satisfactorily. I also noticed, however, that the measure $P(x; y) - P(x)$ – which is often taken to be a measure of relevance of y to x – could be represented as an index of expected epistemic utility. Since some of Popper's measures were functions of probabilistic relevance (among other things), I proposed substituting expected epistemic utility as measured by positive relevance as a substitute for Popper's measures of corroboration. These ideas were written up by early 1961 and published as 'Corroboration and rules of acceptance' (Levi, 1963*a*). My suggested reconstruction of Popper's proposal was clearly unacceptable to Popper and his followers (Michalos, 1966) – much to the detriment, in my judgement, of Popper's own view. I include the paper here, however, not because of its bearing on Popper but because all of the subsequent versions of expected epistemic utility functions I have proposed are generalizations of the idea of probabilistic relevance.

Shortly after writing this paper, I was invited to participate in a symposium of the American Philosophical Association, Eastern Division and, for that purpose, wrote a paper on Hempel's ideas. This paper (1961*a*) relied on the results of my work on Popper's measures for its main results. Even though it appeared in print before (1963*a*), it is the derivative of that paper and I have, therefore, omitted it here.

Aside from these gropings for a sensible characterization of epistemic utility, I was concerned to develop a more sophisticated representation of the options in a cognitive decision problem than the representations I had used in my earlier papers. By 1962, I was thinking of inductive inferences as involving the accepting of a hypothesis along with the deductive consequences of that hypothesis and the available background information and evidence. In that way, one could distinguish between sentences accepted as strongest via induction and sentences accepted via induction but not as strongest. Thus, a potential answer to a question or cognitive option could be characterized as the acceptance as strongest via induction of some element of a Boolean algebra consisting of equivalence classes of sentences where equivalence is relative to background information and evidence.

To make this idea work, it was necessary to address Henry Kyburg's argument that if high probability is sufficient for acceptance, the set of accepted sentences should not, in general, be deductively closed. In 1963, I wrote the first draft of (1967*a*) and worked out to my own satisfaction an account of epistemic utility which, together with the view of cognitive options as cases of acceptance as strongest via induction, yielded a rule of inductive acceptance which rejected the requirement that high probability was either necessary or sufficient for inductive acceptance. This rule preserved deductive closure, assigned a role to indices of caution as determining trade-offs between avoidance of error and relief from doubt or agnosticism and saw expected epistemic utility as a generalized notion of relevance.

Kyburg's point about deductive closure was dramatized for philosophical intuition by the so-called 'lottery paradox'. This paradox became an object of concern among those interested in the epistemological outlook articulated by R. Chisholm (1957). In 'Deductive cogency in inductive inference' (1965*a*), I took the opportunity to respond to Kyburg's lottery paradox and the reactions expressed in (Lehrer, 1964) and (Sleigh, 1964).

The epistemic utility functions introduced in (1967*a*) were not the same as the one I considered in connection with my discussion of Popper in (1963*a*) – although they may be seen as generalizations of it. I did not, however, explain why I had abandoned the earlier proposal, largely because there seemed to be no good reason to discuss technical considerations of this kind for disagreeing with my earlier self when no one else had, so it seemed, joined in the debate. However, while (1967*a*) was in press, I read an important essay by Jaakko Hintikka and Juhani Pietarinen (1966) in which the proposal I

made in (1963*a*) was introduced independently. This persuaded me that I should write a paper in which I explained systematically why various proposals alternative to those I used in (1967*a*) ought to be rejected.

While writing 'Information and inference' (1967*c*), I realized that my characterization of relief from agnosticism in (1967*a*) had been unnecessarily restrictive. By liberalizing it in certain ways, I was able to think of many more contexts of scientific inference as occasions where the aim is to obtain new information while avoiding error. Due to the vagaries of publication, it appeared in the same year as (1967*a*), which it supersedes in certain important respects. It remains, in my judgement, the best short statement of the technical details concerning epistemic utility I have written. For this reason, I have included it here.

One of the most insightful critical discussions of my (1967*a*) appeared as an article by Kenneth Goosens (1976). Goosens raised the question as to how my proposals concerning epistemic utility related to the design of experiments and the evaluation of evidence. I wrote 'Epistemic utility and the evaluation of experiments' (1977) in response to Goosens' paper. Because the relation between epistemic utility and the design of experiments is not discussed in any other publications of mine, it seemed desirable that it be included in this section on epistemic utilities.

Ian Hacking (1967), Richard Jeffrey (1968), Risto Hilpinen (1968) and Ilkka Niiniluoto (1975) all worried about the relativity of my account of inductive acceptance and epistemic utility to a particular set of exclusive and exhaustive hypotheses (the 'ultimate partition'). Although I did not then and do not now see anything objectionable in the relativity, I acknowledged already in (1967*a*) that there was a problem concerning the conclusions an agent is justified in accepting into evidence via induction when he is committed to recognizing several such partitions as representing problems worthy of being considered.

One of the byproducts of work I began doing in the early 1970s on indeterminate probabilities is an account of indeterminate utilities which can be used to address issues of indeterminacy in epistemic utilities which arise when an investigator is conflicted between different cognitive values without being in a position to resolve the conflict. Utilizing the ideas I was then elaborating in the preparation of my (1980*a*), I prepared a paper addressing the problems posed by Niiniluoto, in particular. Although some of the material in 'Abduction and demands for information' (1979*d*) appeared in (1980*a*), some important technical details were omitted and the order of presentation was substantially modified. For this reason, I have included this paper here.

The papers just reviewed constitute a supplement and commentary on the ideas concerning epistemic utility discussed in (1967*a*) and (1980*a*). They are not substitutes for the treatments found in these books; but they should supply some motivational background for issues addressed in them. In the

case of (1967*c*), (1977*d*), and (1979*d*), they supply technical details bearing on my current position in a form which cannot be found so explicitly elsewhere.

In the cognitive decision problems discussed in (1967*a*) and (1967*c*), a cognitive option is represented as accepting a sentence as strongest via induction from background information and evidence – i.e., accepting that sentence together with all deductive consequences of that sentence and the background information and evidence. What does 'acceptance' mean in this context? In (1967*a*), I was prepared to discuss several different senses of acceptance (mere acceptance, acceptance as evidence, acceptance for test). However, I saw the problems of acceptance with which I was concerned as ancillary to the problem of revising the system of sentences accepted as evidence. For me, to accept a sentence as evidence was to be certain that it is true, to assign it probability 1 and use it as background information in subsequent investigations. I framed my acceptance rules to handle mere acceptance and not acceptance as evidence on the understanding that rules for mere acceptance concern what should be accepted into evidence were it the case that the problem under consideration and the potential answers identified for that problem the sole issue regarded as worthy of investigation at the time. Hence, I could say that a hypothesis is merely accepted even though it is not accepted as evidence.

My persistence in putting my view in this way derived from two kinds of consideration. In the late 1960s, I still lacked an account of how one should proceed when there is a conflict between the conclusions justified relative to different questions. This issue is the one considered in (1979*d*).

The other problem concerned epistemological fallibilism – the doctrine that each item in a body of knowledge is possibly false. I was wedded to that Peircean view while at the same time committed to the view that scientists do and ought to accept hypotheses into evidence and assign them probability 1. By distinguishing mere acceptance from acceptance as evidence, I was able to say that an agent could recognize the hypotheses he merely accepts as bearing probability less than 1. But this way of speaking generated the appearance of Peircean fallibilism without its substance; for I also wanted to say that scientists do and ought to accept sentences as evidence. An agent who accepts h as evidence cannot consistently assign it probability less than 1. Yet, I wanted to say that such acceptance is fallible.

Like everybody else, I was perfectly prepared to say that such acceptance is fallible in the sense that it is revisable; but I took for granted that such revisability committed one to the possibility that what is accepted is false. Admitting that possibility appears to require assigning probability less than 1 to hypotheses other than truths of logic and mathematics. Hence, at most such hypotheses could be accepted as evidence.

In the 1960s, I solved this problem (unwittingly) for myself by remaining in a fog of confusion. By 1971, I had come to realize the incoherence incurred by insisting that sentences accepted as evidence are fallibly so-accepted. To

address the difficulty, I proposed to draw a distinction between the fallibility of human knowledge and its corrigibility or revisability. The distinction is simple enough and *prima facie* can be made safe from contradiction. Yet, there seems to have been a pervasive reluctance to make it. In considering this matter, I noticed that two of the most persistent fallibilists, Peirce and Popper, could be understood as having equated fallibilism and corrigibilism due to a shared vision of the ultimate aim of scientific inquiry as being to converge on the true complete story of the world. The explanation is not a causal one. Anyone who has such an aim would find it counterproductive to revise his doctrine if he counted it as certainly true. Rejection of fallibilism brings rejection of corrigibilism in its wake. To escape this predicament, I concluded that one should consider abandoning the Peirce–Popper conception of the ultimate aim of scientific inquiry.

I began reading 'Truth, fallibility and the growth of knowledge' (1983) to diverse audiences in 1971. It was accepted for publication in 1975 but not actually published until 1983. Since it represents the first expression of what I have taken to be an important revision in my thinking and since it seems to me to have a direct bearing on many of the discussions of convergence and truth which have been carried on in the late 1970s and the 1980s, it has been included here.

'Four types of ignorance' (1977*c*) elaborates further on different senses of belief, knowledge and ignorance I have found useful to consider in my work. It contains a brief discussion of the ramifications of my analysis for John Rawls' arguments from the 'veil of ignorance'.

'Escape from boredom' (1981*e*) was provoked by the strong exception I took to Richard Rorty's efforts to group the pragmatism of John Dewey together with the views of Martin Heidegger and Ludwig Wittgenstein, and his persistent hostility to science and the notion of truth. I do not include it here, however, because of the polemic against Rorty. Whatever its origin, it does offer a perspective on the way I see the various more technical questions I explore elsewhere which I have not otherwise articulated.

'Serious possibility' (1979*c*) compares my conception of epistemic possibility with Jaakko Hintikka's conception in his (1962). 'Subjunctives, dispositions and chances' (1977*b*) restates the view of the relation between objective statistical probability I proposed in (1967*a*) in the light of the many discussions of modality and propensity interpretations of probability which had emerged in the interim. 'Direct inference' (1977*a*) focuses attention on Henry Kyburg's important work and its bearing on the interpretation of statistical probability and fundamental issues in statistical inference. It reports on a basic result concerning the relation between Bayesian inference and direct inference which has not been fully appreciated in the literature.

Since 1966, I have insisted on the importance of G. L. S. Shackle's notion of potential surprise and the conception of degrees of belief related to that idea. In this volume, I have taken my latest discussion of Shackle's ideas,

'Potential surprise: its role in inference and decision making' (1980*c*) and modified it by the addition of material taken from a review of L. J. Cohen's *The Probable and the Provable* (Levi, 1979*e*).

Although I fancy myself a critic of Bayesian decision theory, some criticisms of Bayesianism have seemed to me less serious than others. A well-known line of criticism has developed out of consideration of puzzles such as 'Newcomb's problem'. In recent years, there has been a growth of a family of related 'causal decision theories' which modify Bayesian decision theory to accommodate features of causal structure alleged to be relevant to rational choice but which Bayesian theory ignores. I continue to believe that my 'Newcomb's many problems' (1975) addresses some decisive objections to such theories and, indeed, did so before others began formulating such theories in an elaborate manner.

Aside from side references, this volume contains no essays focused on what I take to be the main defect in Bayesian doctrine – namely, its insistence that rational agents be committed (at least ideally) to numerically-definite probability judgements and assignments of value or utility. My (1980*a*) contains a rather extensive proposal for modifying Bayesian doctrine so as to accommodate these complaints and to develop an account of decision making under uncertainty suitable to this proposal. My current work is focused on extending these ideas to contexts where value conflict appears regardless of whether it derives from uncertainty or not.

'Conflict and social agency' (1982*e*) represents my first published statement of an attempt to extend my ideas concerning individual decision making under uncertainty to social and group decision making. In addition, it contains a treatment of the well-known Allais' paradox.

It is well-known that there is a tension between efforts to promote social welfare and liberal protections of certain rights. In (Sen, 1970), A. K. Sen presented a striking argument which converted this tension into a demonstration of the impossibility of a Paretian liberal. 'Liberty and welfare' (1982*b*), the final paper in this collection concedes the existence of the tension but insists that there are conditions under which the tension can be broken.

Ernest Nagel has often chided me for misleadingly distinguishing between fallibilism (which I reject) and corrigibilism (which I endorse). He contends that most readers will find it difficult to disassociate the term 'fallibilism' from the view I call 'corrigibilism'.

Even so, Nagel himself contrasts two kinds of fallibilism: that of Peirce and that of Popper. Peircean fallibilism, so Nagel claims, insists that we do have knowledge and that there are no incognizables. On the other hand, settled knowledge might become unsettled and be subject to revision. Popper, on the other hand, denies that science ever achieves knowledge of any matter of fact. Every assertion of science is merely a guess or conjecture. Newton was mistaken in insisting on a distinction between propositions which are settled certainties and hypotheses which are conjectured and in demanding that we

treat conjectures differently from certainties. For Popper, hypothesizing is all we ever do.

Although I am on record as disagreeing with Nagel concerning the relation between Peirce and Popper, Nagel's contrast between two kinds of fallibilism corresponds very closely to my distinction between fallibilism (Popperian fallibilism) and corrigibilism (Peircean fallibilism as Nagel understands Peirce). For me, the agreement is far more profound than the disagreement.

At the close of the 'Quest for uncertainty', Nagel comments on a distinction between science and myth drawn by Santayana as follows:

> Although I think with Santayana that myth and science differ, it is well to remember that much that passes as science is indeed myth; and the strength of fallibilism lies in the recognition that continuing effort is required not to confound the two. However, the versions of fallibilism I have been considering (the fallibilisms of Popper and Feyerabend) cannot in the end admit that myth is not the same as science and can often be told apart; and I have tried to show that a fallibilism which leads to such a conclusion is neither coherent nor supported by an adequate analysis of the logic of science.

The projects I have undertaken have sought in their own way to understand how the boundaries between myth and science ought to be revised through inquiry. I owe Ernest Nagel a debt of gratitude for having helped guide me in the direction of this problematic and take this opportunity to show my gratitude by dedicating this volume to him.

New York
June 1983