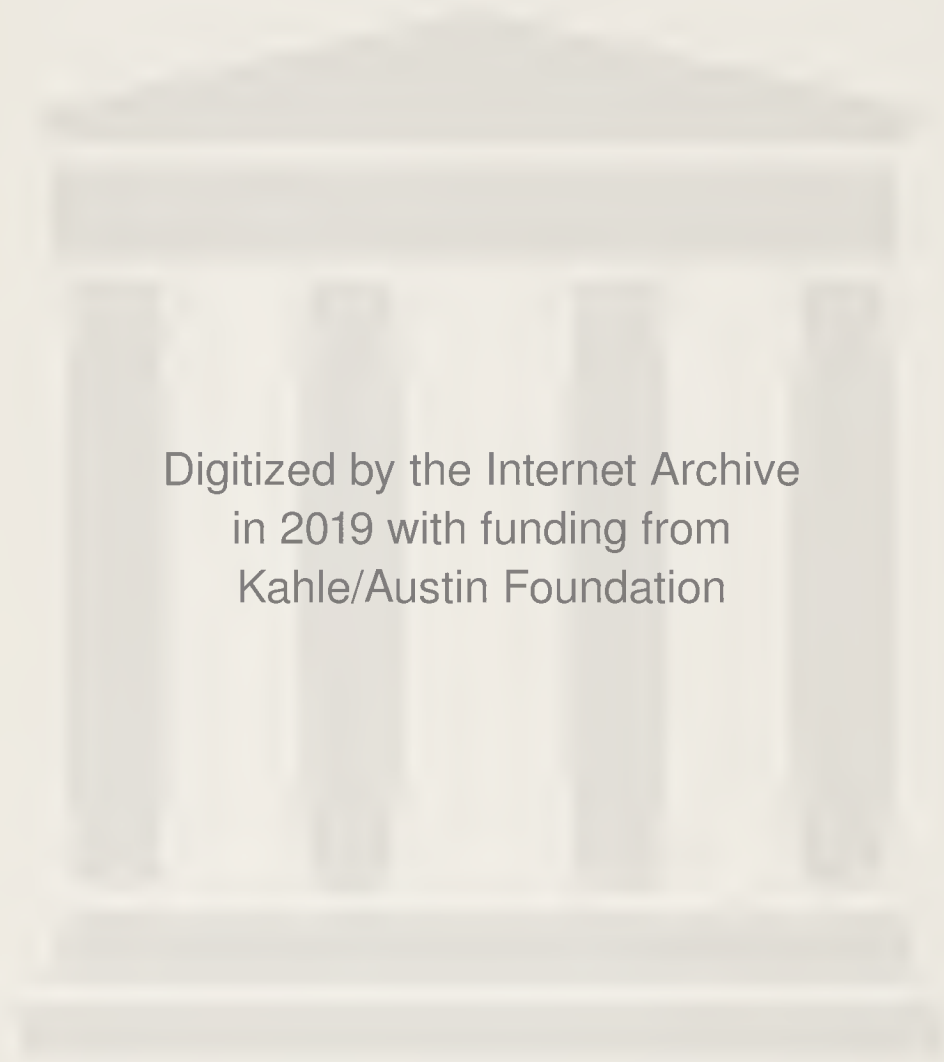
THE PREVALENCE OF HUMBUG AND OTHER ESSAYS

MAX BLACK


Digitized by the Internet Archive
in 2019 with funding from
Kahle/Austin Foundation

https://archive.org/details/prevalenceofhumb0000blac

The Prevalence of Humbug

Also by Max Black:

The Nature of Mathematics
Critical Thinking
Language and Philosophy
Problems of Analysis
Models and Metaphors
A Companion to Wittgenstein's "Tractatus"
The Labyrinth of Language
Margins of Precision
Caveats and Critiques

MAX BLACK

The Prevalence of Humbug

AND OTHER ESSAYS

Thomas J. Bata Libr...
TRENT UNIVERS...
PETERBOROUGH ONTAR...

Cornell University Press

ITHACA AND LONDON

B29 .B543 1983

Copyright © 1983 by Cornell University Press

All rights reserved. Except for brief quotations in a review, this book, or parts thereof, must not be reproduced in any form without permission in writing from the publisher. For information address Cornell University Press, 124 Roberts Place, Ithaca, New York 14850.

First published 1983 by Cornell University Press.
Published in the United Kingdom by Cornell University Press Ltd.,
Ely House, 37 Dover Street, London W1X 4HQ.

International Standard Book Number 0-8014-1514-4
Library of Congress Catalog Card Number 82-22211
Printed in the United States of America
Librarians: Library of Congress cataloging information appears on the last page of the book.

The paper in this book is acid-free and meets the guidelines for permanence and durability of the Committee on Production Guidelines for Book Longevity of the Council on Library Resources.

To Michal

uxori carissimae
cum gratia et amore

There is no safety in numbers—or in anything else.
—James Thurber

Contents

	Preface	9
	Introduction	13
	Part I / Rational and Reasonable	
1	*"Why Should I Be Rational?"*	19
2	*Reasonableness*	41
	Part II / Aspects of Science	
3	*Scientific Objectivity*	59
4	*Is Scientific Neutrality a Myth?*	77
	Part III / Beyond Rationality	
5	*Humaneness*	97
6	*The Prevalence of Humbug*	115
	Part IV / Some Puzzles	
7	*The Rationality of Voting*	147
8	*Newcomb's Problem Demystified*	160
	Notes	175
	Index	183

Preface

Nearly all of these pieces originated as talks for general audiences, often delivered from notes. In polishing transcripts and drafts I have tried to preserve a conversational style. I hope readers will enjoy reading the essays as much as I did composing them. I chose subjects that happened to interest me, in the hope of finding out what I then thought and what more needed to be done.

Collections inevitably provoke questions about coherence. I hope that readers will discern in these assembled pieces a complex unity resulting from a number of interlacing and contrasting preoccupations. Among them are concern for reasonableness, restrained by a conviction that rationality is not enough; commitment to common sense of the kind that does not shy away from science and philosophy; appreciation and distrust of abstract models; as much interest in the unformulated stratagems and implicit understandings of speech as in the normative codes of grammar and logic. While the essays might be called philosophical exercises, they presuppose no exposure to academic philosophy.

I have been greatly helped by comments from friends, students, and colleagues. The editors and publishers who have kindly permitted me to reprint some earlier versions are named at the beginning of the notes to the relevant essays.

A special word of gratitude is due to my friend Edward Connery Lathem, whose generous stylistic suggestions have been invaluable.

MAX BLACK

Ithaca, New York

The Prevalence of Humbug

Introduction

A poet, uncomfortable in the role of literary theorist, once remarked: "The reader will wonder, perhaps, what to expect of me. I should fly some recognition signal."[1]

When I tried to do so, I came up with nothing better than: "Lapsed mathematician, addicted reasoner, and devotee of metaphor and chess." Even that unsatisfactory self-identification may help to explain the variety of approaches in the present collection of essays: the reader may notice an intermittent tension between interest in finding good reasons and sound arguments, on the one hand, and on the other a standing conviction that a single fruitful idea is worth a hundred syllogisms.

The need for a recognition signal is perhaps especially felt in an occupation as problematic as philosophy. I therefore hope that the following informal remarks may at least warn an intending reader about what not to expect.

Philosophers are commonly expected to be sages, producing edifyingly profound thoughts about such matters as the Meaning of Life or the Secret of the Universe. I have little taste for generalities about capitalized abstractions, suspecting that life has no edifying Meaning and that the open Secret about the Universe is the absence of such a secret. So I leave pursuit of Wisdom to the gurus and reach-me-down mystics, who are now, as always, in plentiful supply.

Yet I believe that a philosopher, in the more modest role of a

rational critic in pursuit of clear ideas and sound reasons, can help thoughtful nonprofessionals to cope with issues of serious private and public concern.

Curious readers who sample philosophical work are apt to be dismayed by a seeming preoccupation with farfetched questions. The resulting frustration is well expressed in Mencken's amusing squib: "A metaphysician is one who, when you remark that twice two makes four, demands to know what you mean by twice, what by two, what by makes, and what by four."[2] Ironically, Mencken's prime examples of philosophical folly are vitally important in the philosophy of mathematics, where splendid and lasting advances have been made in modern times.[3]

Questions of the form "What do we mean by *X*?" and "What *should* we mean by *X*?" are fine starting points for serious reflections and will repeatedly arise in the following pages. Yet, while the precision of mathematical concepts and the existence of well-crafted symbolisms for their precise expression have greatly helped us to understand what mathematics is all about, the situation is very different in ethics, the philosophy of action, and political philosophy.

In these controversial fields, the intellectual and political games played are reminiscent of the immortal croquet game in Wonderland, in which the croquet balls were "live hedgehogs, and the mallets live flamingoes," while the players quarreled so dreadfully that one couldn't hear oneself speak—and if there were any rules, nobody attended to them.[4] The basic words and ideas of current controversy are as nimble as the paraphernalia of Wonderland croquet. Such indispensable staples of moral and political discourse as rights, duties, justice, freedom, equality, welfare, and many more (including my own standing preoccupation, rationality) are as prickly as hedgehogs and as flighty as flamingoes.

Indeed, basic but highly controversial notions, in any field of action or inquiry, are usually "essentially contested concepts,"[5] ideas of such established prestige that opposing partisans will fight for the right to use them—usually in question-begging senses that subvert their ordinary meanings. One's own preferred ideal of freedom is called "real" or "true" freedom, while

the opposition advocates "license," or "freedom, falsely so called." Even those who disdain such disreputable maneuvers must cope with the endemic ambiguity and duplicity of ordinary language, in fields where precise terminology is too often a disguise for shoddy doctrine.

Here, then, there is work aplenty for semantic hygiene or the clarification of concepts.[6] The reader may expect to find me often preoccupied with the meanings of words in common use. My concern, however, is not with words alone, or merely with their dictionary senses, but especially with concepts, hidden rules of usage, and the nonverbal realities or fictions designated. The elusive character of fundamental concepts partly accounts for the imperfection of much philosophical as well as popular argument.

All responsible thinkers have hoped to find good reasons and sound arguments for their convictions. Yet a fine scholar once said:

> Almost all philosophical arguments are invented . . . to recommend, or defend from attack, conclusions which the philosopher was from the outset bent upon believing, before he could think of any arguments at all. This is why philosophical reasonings are so bad, so artificial, and so unconvincing. To mistake them for the causes which lead to a belief in the conclusion, is generally to fall into a naive error.[7]

Reasons can be "bad" either in themselves or in the uses made of them: they may need more support than the conclusions they are supposed to establish, or, while themselves acceptable, may be used in unsound argument.

A reasoner need not be guilty of objectionable dogmatism in offering premises without further support: for an argument must start somewhere and a demand that everything should be supported by reasons would be self-defeating. But one's most basic premises ought not to be held with obdurate commitment. It is discreditable to refuse to consider possible objections to one's most cherished convictions.

I readily confess to a strong belief that common-sense convic-

tions about the existence of other sentient persons, the reality of the "external world," and so on are more likely to be right than the skeptical objections of even the greatest philosophers. Indeed, I am unapologetically sure that no amount of argument will ever persuade me that time is unreal, that physical objects are merely collections of sense data, that persons are only machines, that induction is an illusion, and so on. I am troubled, however, by the ubiquity in private and public life of *invalid* argument.

One reason for the endemic presence of logical fallacy in the work of even the most scrupulous reasoners is the tedium of what might be called slow-motion argument: the kind of proof found in Euclid's *Elements,* where each step is carefully recorded, with explicit invocation of previous results, and the requisite principles of inference. Such formal reasoning, though invaluable for exhibiting hidden assumptions, is boring: even pure mathematicians, professionally devoted to strict argument, would rather run than crawl, can't be bothered with stating the obvious, and hurry along to assimilate the gist of an extended argument as an articulated mental structure. Yet impatience with the kind of tedious detail that might ideally be left to a computer readily generates fallacy. So proof checking, though as tiresome as the proof reading of manuscripts, is, in the end, indispensable.

The present collection, however, does not aim at the precision of exhaustively elaborated argument. My aim is not to prove anything, but to promote understanding and to provoke discussion. So I do not apologize for skating on thin ice in constant awareness of hidden depths.

I shall be disappointed if these essays do not occasionally entertain as well as enlighten. Humor need not be an enemy of rationality, and "life is too serious to be taken altogether seriously" (Hu Hsai).

Part I

Rational and Reasonable

I /

"Why Should I Be Rational?"

> I can stand brute force but brute reason is quite unbearable. There is something unfair about its use. It is like hitting below the intellect.
>
> —*Oscar Wilde*

The Challenge

The philosopher ended the last lecture with a traditional valediction: "In conclusion, ladies and gentlemen, let me urge you once again: Do try in future, do please try, to be rational!" After the cheers and jeers had died away, a student rose to say, "Professor, I have been itching for a long time to ask: *Why* should I be rational?" Was this an absurd question? If you think so, you will be agreeing with nearly every philosopher who has considered the matter. I shall try to explain why I disagree.

I invite you to think of my imagined heckler as offering a challenge to the pursuit of rationality: I shall call his performance "the challenge," not yet assuming that he is asking a genuine question. If he could properly be so regarded, as I think, it remains to determine just what his question may be.

It is initially unclear what the lecturer meant by urging the pursuit of rationality. When John Locke discussed Reason, he reminded his readers that "the word Reason in the English language has many different significations," and his careful editor,

A. C. Fraser, commented: "'Reason' is among the most ambiguous of philosophical terms."

Rational and *rationality, reason* and *Reason,* remain hotly contested notions,[1] whose users disagree even about the nature of their disagreement. We shall need to decide what my imagined professor intended when he urged others to be "rational."

The Pursuit of Reasons

A friend of rationality might prefer the admonition "Try to live rationally" when recommending a necessary though insufficient condition for the Good Life. Thus rationality might reasonably be conceived as a style of life, more general than a specific policy and less comprehensive than a "way of life." Speaking now for myself, I would partly specify my own commitment to a rational life-style by means of the following self-addressed injunctions: (1) To wish for good supporting or justifying reasons for my beliefs, feelings, attitudes and actions, whenever such primary reasons are available; hence (2) to care about finding such reasons, unless (3) I have got or can find good second-order reasons for holding in particular cases that primary supporting reasons are not needed.

You may object to some features of this sketch, but I hope you will agree that a rational life-style must include some respect for and reliance on reasons. Indeed, this much would be accepted by all contemporary partisans of conflicting ideals of rationality. This leads me to postulate now that my imagined challenger would accept the following formula as a further specification of his intended question: "Why should I pay any attention to reasons?"

Isn't the use of such a formula self-defeating, by demanding a reason for relying on reasons? Isn't that as absurd as asking why anybody should ever ask a question? Would we not be justified in retorting that this formula does not and cannot express a genuine question?

My initial reluctance to credit a user of the formula with rais-

ing a genuine question is influenced by Charles Sanders Peirce's wise remark that "the mere putting a proposition into the interrogative mode does not stimulate the mind to struggle after belief. There must be *a real and living doubt,* and without this all discussion is idle," and by his admonition that philosophical investigations should start "from the very state of mind in which you actually find yourself at the time."[2]

I suggest, accordingly, that we shall be unable to credit our challenger (let us call him Skepticus) with a genuine question, expressing some real and living perplexity, until we specify his own state of mind.

A Gallery of Questioners

There will emerge a spectrum of conceivable challengers, needing appropriately differentiated responses.

At one extreme, Skepticus might be a *militant antirationalist,* who assimilates the homilies of rationality advocates to the braying of asses. His challenge is a malicious provocation, intended to induce one more exhibition of academic folly. He uses such expressions as "reason" and "rational justification" with invisible quotation marks, as a principled atheist might in alluding to angels and devils. Such a bigot is immune to rational persuasion: if talk about rationality makes him reach for a gun, avoidance and flight are the most rational responses.

Still, Skepticus might, more promisingly, be an *agnostic rationalist,* who usually "listens to reason" in practical affairs, but despairs, like Hume, of finding rational foundations for his beliefs and actions. Conformist in thought and deed but agnostic about the ultimate articles of faith, he is likely to be easy prey for anthropologizing relativists, for whom "our rationality" is merely one culture-bound form of life among many others.

Finally, Skepticus might after all be a *confident rationalist,* wholeheartedly though undogmatically following a rational style of life, and sure that there are good reasons for his commitment, even if he cannot produce them.

In what follows, I shall usually think of Skepticus in his avatar of Hume redivivus—that is, as an agnostic rationalist—since the stance of such a person presents the most serious challenge to a would-be rational defender of rationality.

The Form of an Appropriate Answer

Consider two possible answers to Skepticus, merely by way of illustration: "You should respect reasons, because if you do, you will have a happier life than if you don't"; "You should respect reason, because if you don't, you will roast in hell." These are practical syllogisms: in each of them some relatively beneficial consequence—living more happily than otherwise or avoiding the agony of hellfire—is alleged to result from the pursuit of reasons.

There can be no serious doubt that the arguments I have suggested—or, if one prefers, the fully explicit versions of those arguments—are formally valid. But they might be valid and yet unsound; or sound but ineffective for the person addressed (our Skepticus, say). They will be unsound if they beg Skepticus' question, as alleged by contemporary philosophers; and they will be ineffective if the alleged benefits or harms are not accepted as such by him.

In what follows, I shall first try to dispose of the popular view that arguments of the sort exemplified above do beg the question of the desirability of respecting reasons (i.e., attending and seeking them in the manner previously outlined), and then I shall offer a positive defense of that essential aspect of a rational lifestyle that consists of respect for reasons.

Geach's Objection

P. T. Geach has said:

> Though it is reasonable to ask for reasons, it is not always reasonable to ask for reasons. . . . [I]t is not reasonable to ask for a reason why one should ever observe the practice of asking for

> reasons; the man who claims to reject the practice must not ask others why (i.e., for what reasons) they go in for it, or else he is shown not to be a total abstainer from it.[3]

Here Geach is unjustifiably assuming that the questioner (our Skepticus) must be somebody who "rejects the practice" and who is therefore what I have called a militant antirationalist. In such a case, I would agree that offering a serious answer might be as futile as starting a game of chess with somebody who absolutely refuses to play the game. But we have seen that Skepticus may well be, like Hume, somebody who does play the game, but has yet to be shown the rational point or justification of doing so. And a more radical questioner might still deserve some kind of answer, even if he merely uses such words as "reason" sneeringly, in the sense of "what you call reasons."

If we agreed with Geach, we ought to respond to Skepticus by saying something like: "I can't answer, because you are behaving unreasonably in asking the question." Then *we* should be behaving reasonably in having a good reason for refusing to answer, but that would cut no ice with the questioner. Yet to offer a direct answer instead would, in Geach's view, be as unreasonable as to start a game of chess with somebody who "rejects the practice" of playing the game.

Geach seems to assume that the questioner is a total abstainer from rationality: I shall argue later that this need not be so. And even somebody who is utterly skeptical about the point of the practice of using reasons might maliciously intend to embarrass an advocate of rationality. If so, there is prima facie nothing unreasonable about presenting supporting reasons for the practice.

But wouldn't there after all be something objectionably circular about offering a reason for attending to reasons? I shall now briefly discuss the considerations advanced by two other philosophers for thinking that there is indeed such an absurdity.

Peter Singer's Objection

Peter Singer has recently claimed that what I have been calling the challenge must be dismissed because it "questions some-

thing—in this case rationality—normally presupposed."[4]

Well, the alleged presupposition might plausibly be expressed in the following contention:

"One should provide reasons for assertions." (*C*)

A more defensible version might be:

"Reasons should be offered for assertions, whenever possible or desirable." (*C'*)

That *C'* is in some sense "normally presupposed" in ordinary conversation and argument cannot be expected to appeal to a questioner who is prepared to reject this very presupposition. Might he not retort that if rational conversation presupposes something like *C* or *C'*, then so much the worse for *that* kind of discourse, which he himself is happy to forgo? (Cf. "Betting in a casino presupposes willingness to lose money." "Then I'd rather not bet!")

But it has not yet been shown that a rational defense of rational procedures presupposes what such a defense is intended to achieve. I have proposed that a rationalist's defense might run: "You should attend to reasons, because if you do so, such-and-such (which you don't want) will be the case," which has the form of:

C; because if *C*, then *D* (*A*)

where *D* is understood to be something unwanted by the person addressed. So our present question becomes: Does *A* presuppose *C*, and if so, how?

Well, nobody could sincerely present or accept *A* without already believing its conclusion (*C*) to be correct. But there is nothing disturbing, from a logical point of view, about this, since we often find it useful to argue for a conclusion that we already believe to be correct, for example, in order to strengthen an antecedent belief or to establish logical connections between isolated beliefs.

To use *C* as a *premise* of *A* would, of course, to be guilty of flagrant circularity, arguing from *C* to *C*. But clearly *C* is not needed as a premise of *A*, even though we show that we accept *C* when we argue in its favor. In order for *A* to be valid, it suffices that the reason presented (if *C*, then *D*) shall be a good reason for us to accept *C*.

Now consider the position of a skeptical questioner who is

offered argument *A*. Does he need to believe *C* in order to perceive *A* as valid? No: he is required only to see that the *particular* reason supplied (if *C*, then *D*) is indeed a good reason for him to hold *C*. If he cannot or will not do that much, then of course the argument will fail to convince him of the advisability of attending to reasons—and so, no doubt, would any other argument to the same effect. No argument, however sound, can have any appeal to a questioner who is too stupid to recognize a good reason when one is offered—or too obdurate to confess that he does recognize it.

Ayer's Objection

Writing on a closely related topic, the justification of reliance on inductive procedures, A. J. Ayer has said: "What is demanded is a proof that what we regard as rational procedure really is so; that our conception of what constitutes good evidence is right. But of what kind is this proof supposed to be? A purely formal proof would not be applicable, and anything else is going to beg the question." Ayer views this outcome with some complacency: "This does not mean that the use of scientific method [in context: inductive procedures] is irrational. It could be irrational only if there were a standard of rationality which it failed to meet; whereas in fact it goes to set the standard: arguments are judged to be rational or irrational by reference to it."

No doubt Ayer would take the same position with regard to our rationality question. He might well feel that with regard to a demand for an argument in support of appeal to reasons, "no proof that we are right can be forthcoming: for at this stage nothing is going to be allowed to count as such a proof." Substituting rational argument for inductive reasoning in Ayer's peroration, we get: "When it is understood that there logically could be no court of superior jurisdiction, it hardly seems troubling that [rational argument] should be left, as it were, to act as judge in its own cause. The sceptic's merit is that he forces us to see that this must be so."[5]

Stripped of its persuasive rhetoric, Ayer's view amounts to

this: scientific method is what scientists accept as justifying the conclusions they accept; appeal to good reasons is what those of us (including scientists) treat as a "standard"[6] for justifying beliefs: "logically" speaking, there can be no "proof" or even a good supporting argument to show that we are right in having such a standard. (This implies that even if experience were to show that reliance on reasons does not pay in certain domains—an eventuality not to be excluded a priori—that would in no way discredit our continuing reliance on reasons, even in those very same domains.)

Unlike Ayer, I do find this view of the matter "troubling." For I believe it provides, however unintentionally, substantial ammunition to an irrationalist. Such a one might plausibly retort that if there can be no reasons favoring rationality, there cannot be any reasons against irrationality. Seen in this light, a settled policy of rejecting appeal to reasons will seem no shakier than the contrary policy. Currently fashionable talk about the standard *we* adopt, or the rules supposedly governing the uses of rationality and its semantic cognates, or the "game that is played," or the mores and conventions of "our" society, will then have no persuasive force against anybody seriously *revolting* against these very standards and practices.

Popper's View

I shall end this part of the essay with a remarkable illustration of the tendency of even passionate devotees of rationality to succumb to what sounds like ultimate irrationality.

In an essay titled "Utopia and Violence," Karl Popper partly identifies his own version of "critical rationalism" in the following terms: "A rationalist, as I use the word, is a man who attempts to reach decisions by argument, and perhaps, in certain cases, by compromise, rather than by violence." Popper then makes the following remarks about the basis of his own "rationalism" as thus understood: "[M]y rationalism is not self-contained, but rests on an irrational faith in the attitude of reasonableness. I do not see that we can go beyond this."

Here Popper uses the word "irrational" not in its usual sense of "contrary to reason" but rather as meaning "*non*-rational," that is, unable to be supported by argument. That this is his intention is made explicit by the following passage: "My rationalism is not dogmatic. I fully admit that I cannot rationally prove it. I frankly confess that I choose rationalism because I hate violence, and I do not deceive myself into believing that this hatred has any rational grounds."[7]

Although Popper has not explicitly raised the question I have chosen as a title, he has indicated an answer to it in the passages I have quoted. A Popperian style of reply to the questioning student imagined in my opening paragraph would presumably run somewhat as follows:

"Why should you be rational? There is no 'should' about it! I can give you no arguments in favor of being rational. To be blunt about this matter, there are no *reasons* for you to be rational (for the existence of any such reason would yield an argument for rationality). You have a *choice* between being rational and practicing violence. If, like me, you hate violence, you will choose rationalism; if not, not. That's all there is to say. If you reject rationalism, you will be irrational, and that's all there is to it."[8]

It is worth noting that this position implies that there are also no reasons for or against violence. For if violence and rationalism are mutually exclusive alternatives—as Popper seems to imply—any reason for or against violence would be a reason against or for rationalism. Whether Popper's view about the epistemological status of rationalism can itself be rational in Popper's sense—that is, supportable by reasons—I shall not stop to consider. At any rate he offers no such reason that could be expected to sway anybody already committed to some alternative to rationalism—say, praying for divine guidance.

Might not a terrorist, happy to bait the enemy, reply in the following terms?

"Bravo! You hate violence, but I hate argument (a sneaking use of force by other means). You call me irrational, but I glory in that title. Like you, I hold that there are no ultimate *reasons* for my irrationality (for that would detract from the purity of my position). The difference between us is like that between a Prot-

estant and a Catholic: your faith is my heresy; my faith is your heresy. That's all there is to say."

Leaving such a terrorist to gloat over his dialectical skill, I shall try to argue that there does remain something important to say *from a rational point of view* in defense of rationality.

The Bogey of Infinite Regress

I have claimed that a suitable response to our imagined challenger can take the form of a "practical" argument.

C, because *R* *(A)*

in which something, *R*, alleged to be the case, is offered as a reason for the problematic assertion *C*. Thus a demand for a reason is met by a piece of purportedly sound reason*ing*. In order to do the job, *A* must be valid, yet will not count as *sound* unless it is agreed that belief in *R* is justified.

Now a skeptic, an agnostic, or a rationalist with high standards of argument might object that *A*, even if valid, merely shifts the burden of establishing credibility from *C* to *R*. Accordingly, he now demands a reason for accepting *R*, that is, a further argument to show that belief in *R* is justified. If such a reason, *R'*, is supplied, *A* expands into:

C, because *R*, because *R'* *(A')*

But *A'* is as vulnerable as *A* was to a demand for further justification. So the objector contends that we cannot be satisfied with *A'*, but must move on to A", and further still, without end. (If no argument will satisfy him but the next, he can never be satisfied.)[9]

Traditionally, the way out of this predicament has been to stop at some R^n that allegedly needs no further support—either because its truth is self-evident (no sooner clearly understood than intuited as true) or incorrigible (sincere belief in its truth guarantees its truth). But when *C* is some empirical proposition, or some injunction concerning an action (or other "doing") in the real world, it will be impossible to produce such a proof from self-evident first principles or incorrigible sense-datum assertions.

This line of thought, as old as Plato, remains a potent incentive to skeptical antirationalism. For anyone endorsing it may plausibly feel that the chain of reasons can be terminated only by some external compulsion, fiat, or arbitrary choice (cf. Popper's remarks discussed earlier). In acute instances of this neurotic anxiety, the sufferer may feel like some character in a story by Kafka, protesting an unmotivated condemnation, who is sent from one functionary to another, only to find at last an ayatollah who pronounces guilt and an assassin who gives further argument a quietus. (When Wilde referred to "brute reason," did he have a similar picture in mind?)

Adequate discussion of the complex issues raised by the infinite regress argument must await another occasion. I write here only for an imagined questioner who retains sufficient common sense and self-knowledge to confess that he does not genuinely doubt that he exists, or that he sees a tomato, or such other staples of epistemological debate.

My own position, stated briefly, is that sufferers from hyperbolic pseudo-doubts are illicitly operating with a mathematico-logical model of proof, inspired by Euclid's axiomatic presentation of geometry or its latter-day refinements. This perverse inclination to demand mathematically strict standards of proof for empirical assertions or for practical injunctions is, I suggest, strengthened by a seductive metaphor of reasons as supports, which I shall examine in an Appendix.

A Landing Stage and a Preview

The various views of those who have examined the rationality question ("Why should I be rational?") can be summarized as follows: (1) the question is too incoherent or self-defeating to deserve an answer; (2) the question makes sense, but can receive no rational answer; which shows that (3) rationality must be grounded in something nonrational (faith, custom, hatred of available alternatives, or whatever); (4) the impossibility of an answer shows that rationality needs no defense. I have tried to show that such views rest on misconceptions, and I have urged

that the rationality question, when specified by attention to the state of mind of a hypothetical questioner, can sometimes deserve an answer.

The time has now come to offer what may serve as at least a preliminary sketch of an answer (for a questioner whom I shall imagine to be an agnostic rationalist).

My strategy will be to describe the most primitive form of reason-using (or something like it, detectable in the "lower animals" as well as in *homo sapiens*), to consider some consequences of abandoning this practice, and finally to sketch how progressively more elaborate and sophisticated exercises of rationality can be built on the basis of such basic or proto-rationality.

Basic Reason Using by Humans

Consider the following episode: While clambering up a mountain slope, I suddenly notice a large boulder hurtling toward me. I jump to one side—and tumble into a stream. This is a paradigm case of what I shall call basic reason using (BRU). In this imagined instance, it would be generally agreed, I hope, that my perception of something's being the case (the rolling boulder), seen as imminently dangerous, prompted evasive action. (I use the colorless word "prompted" to avoid discussing whether the relation between the perception and the action was partly a causal one, as I believe.) In other BRU episodes, such as reaching for a ripe blackberry, the induced expectation is of some positive benefit. A rough definition of a case of BRU might run: prompt action induced by a perception, not necessarily correct, of something perceived as imminently beneficial or harmful.

I have deliberately chosen as my prime example an episode in which there was no time for well-articulated thought. What passed through my mind at the moment of peril might be plausibly rendered as "Ouch! Going to hit me. Jump!" or something still more disjointed and fragmentary. But afterward, of course, I have no trouble in providing an articulated version of what I thought and did, along the lines of "I saw a large boulder rolling

toward me and expected to be hit if I didn't act at once, so I jumped." This full-blown version has the form of a practical enthymeme; in supplying it, I am, as it were, replaying what passed through my mind in slow motion, with some additional editing.

Well-articulated inference, whether from perception or antecedent belief, is comparatively rare in human experience, being normally reserved for difficult or complex problems that demand orderly approaches and critical appraisal of the reliability and strength of the reasons adduced. Commonly, fragmentary expressions of perceptions or thoughts suffice to induce satisfactory conclusions, informally derived. Even the eminently rational deliberation of a chess player will probably run somewhat like this: "That to there? Will be captured. Same if go there? No. OK" (a far cry from the decorous full sentences of De Groot's "protocols").[10]

Postulating, accordingly, little more than a kind of telegraphese of fragmentary visual and verbal images, even in sophisticated exercises of human reasoning, I regard clear cases of human BRU as normally requiring no more than a capacity to articulate and reconstruct reasoning post hoc. (And even so, a reasoner who is maladroit at retrospective reconstruction is not necessarily to be censured as an inept reason user.)

To return, then, to my original paradigm, I qualify the episode as a case of *basic* reason using, because I view such quasi-automatic response to situations of immediate benefit or risk as necessary for the adequate performance of all but the crudest human reactions (jumping when pricked). Basic reason using is obviously indispensable for human survival, but basic also for the most elaborate and distinctive manifestations of human intelligence and skill.

Quasi-Rationality in Beasts

Consider next the following samples of animal behavior, each of which might plausibly be regarded as resembling a case of basic human rationality:

1. A bear raises its head while feeding and leaps out of the path of an approaching boulder.

2. A cat, observing its mistress pouring milk, moves to its saucer.

3. A dog is seen trying to cross a busy street—looks to either side, darts into the traffic, returns to the sidewalk, and so on.

4. Some deer, grazing in a sheltered hollow, respond to a distinctive warning cry from the sentinel posted at the rim by taking to flight.

Each episode resembles my earlier paradigm of human basic reason using (the mountain walker and the rolling boulder) by involving anticipation of danger or imminent benefit: animal and human alike suit behavior to the expected-but-not-yet-occurring. We can think of such anticipatory response as the interpretation of perceptual *signs,* which as it were enlarge the actor's perceptual field: thus the deer behave as if they could peer over the hollow's rim at the approaching predator. Each actor in our miniature dramas responds as if he could actually see in advance what is about to happen.

Some readers may find a whiff of antropomorphism in my references to "anticipation," "warning," and "sign." But I am not imputing to bear or deer mental acts that phenomenally resemble what passes through the mind of an adult human in a similar situation. (Can we even guess how the world looks to a dog or a horse?) Of course, bears and deer have no *concept* of a reason or a sign and may well be incapable of producing even the visual images, let alone the fragments of verbal telegraphese, that often mark informal human inference. Above all—not to shirk the obvious—not even the talented chimpanzees now so much in the news can be plausibly expected to exercise our own retrospective power of articulating nonverbal rational action ("I jumped because I saw the boulder rolling") or the reflective self-criticism ("How stupid to ignore the chance of falling into the stream!") required for learning from our mistakes.

Granting all this, I still wish to stress the following points that will, I trust, be readily conceded: First, the animal actor's *manifest behavior* may closely resemble human behavior in a like case: we jump aside as the bear does; a child's forays into a busy street do

not significantly differ from those of my imagined dog. Indeed, if we indulge the fancy of being transported into the body of another animal—say a pigeon—with human powers of perception and inference unimpaired, our behavior, for all our cognitive advantages, could not be expected to be better than that of ordinary pigeons.[11] For evolutionary adaptation can be counted on to have eliminated, on the whole, any immediately deleterious responses to immediate stimuli. One might say that the "lower animals," given their relatively modest cognitive abilities, behave *as if* they were as rational as they need to be.

Second, intelligent human behavior obviously requires the acquisition of *automatized* skills: a child musician laboriously calculating where to place his fingers must end by getting the notes right as unthinkingly as a bird pecking at a seed. It is as if developed human skill requires a return to the spontaneity of merely animal response (releasing mental energy for higher tasks), with habit doing the work of instinct.

Given the homology in manifest behavior of basic human rationality (innate or acquired) and parallel animal behavior, I propose to call the latter exercises of *quasi-rationality*.

Answering Skepticus

Basic reason using in humans (BRU) and its homologous quasi-rationality in other animals conduce to survival and well-being by promoting benefit or averting harm. But only when the reasons used—or, equivalently, the sign signified—are well used: the reason has to be "good," the sign must be a reliable signifier. Conversely, impairment of the requisite inferential capacity, whether by fatigue, organic malfunction, or disease, tends to produce frustration of desire, bodily or mental damage, or, in extreme cases, death. The price of poor reasoning, whether through misperception or faulty anticipation, can be frustration, hurt, or extinction.

Given all this, we now have the makings of a rather obvious but compelling reply to our imagined questioner, who is du-

bious about the point of rational behavior. The prime reason for not trying to discard the basic rationality that we can assume Skepticus to be already practicing, since he is alive and in sufficiently good shape to ask disturbing questions, is that to do so invites the risk of unwelcome death. Expressed in our canonical form, the argumentative reply to the original question takes the form: "You should respect basic reasons (be basically rational) because not to do so is to expose yourself to almost certain frustration, pain, and death."

Given that Skepticus is a developed human being, he has the choice of accepting or rejecting the normative conclusion of this argument: he could on principle try to emulate an inarticulate "lower animal," trusting to his own reflexes to keep him out of danger—or, if he is sufficiently doctrinaire, could even adopt a fanatical policy of total indifference (a kind of extreme ataraxia) toward the normal perceptual signs of danger or benefit. But if so, he had better find some riskless environment or trust to friends to keep him out of mischief.[12] I have already argued that such a prudential argument involves no vicious circularity: in offering it, we presuppose nothing at all, although we can reasonably expect and hope that Skepticus will understand the argument and feel its force. If not, there is nothing more to be done.

Even if Skepticus accepts the prudential argument outlined above, it may be felt that very little has been accomplished. For Skepticus might well be basically rational, in our deliberately restricted sense of responding intelligently and prudently to perceptual signs of impending harm or benefit, and on the whole behave quite irrationally: he might have the horse sense to dodge traffic and still believe that he can command the stars and live forever. But in the present project, as in other practical undertakings, the first step can be of extreme importance. Once you can get a troubled questioner to see the point of basic rationality, the ground has been laid for consideration of what is involved in having a (basic) reason for conduct, and so by degrees for an unfolding of the need for a more comprehensive conception of rationality. Such rational development of the point of rational thought and conduct might well commence with the implications of the needs and opportunities of a *social* and *literate* animal,

focusing perhaps on the kind of implicit reasoning involved in talking and listening to others. And so by gradual steps to the consideration of vexed questions about the limits of rationality, its relations to faith and revelation, and so on. But this complex task is too much to commence at the end of what is already a long essay.

I shall be content if the line of argument can give my imagined questioner good reasons for consciously affirming basic rationality, leaving his further induction into the human race, as it were, for another and more ample occasion.

The Professor's Reply

Since I opened this essay with an imagined challenger, I shall, at the cost of some repetition, end with an imagined response:

Why should I be rational? That is like asking, "Why should I breathe?" And one answer to *that* is: "It's harder than you might think to stop—and stupid to try if one isn't tired of living." (There are less painful ways of committing suicide.)

You can't see the analogy? Well, consider. You were born an animal—an extraordinary one, to be sure—and remain so. I too. So here we are, too "educated" to be good animals, too ignorant to be satisfactory humans. Now our animal capacities to survive and satisfy our needs (by breathing, eating, mating, and so on), however modified by social "acculturation," strongly depend on a primitive repertoire of inferential skills for recognizing and responding to threat of harm and promise of benefit. We were born able to treat sights, sounds, and scents as *signs* of impending events of concern; and all our learned responses rely on this primitive survival-and-satisfaction kit. (If you see a bicycle rolling your way, you step aside; when the fire blazes suddenly, you flinch; the sight of a ripe blackberry sets your hand moving.)

Such anticipatory activity, however spontaneously and unconsciously performed, is already an exercise of primitive rationality of a very basic sort. All distinctively sophisticated human performance depends on basic sign reading—as ability to

play a Chopin prelude requires ability to strike the correct keys unthinkingly.

An act of such minimal rationality can be viewed as an episode of compressed reasoning, capable of subsequent articulation, on the model of: "That bicycle will hit me if I don't move, *so* I shall jump [Does so!]." In this reconstruction, the *so* (or such a surrogate as "therefore," "hence," or "because") stamps the episode as one of inference to a practical conclusion. The most basic use of a reason is to reason from a perceptual sign to what it signifies.

When a "lower animal" reacts to a sign portending immediate harm or benefit, it behaves to all outward appearance as we do: I jump away from the bicycle and so does my dog. But I have a reason for jumping and know that I have; my dog simply jumps. I call his action *quasi-rational* because I applaud his behavior as life-preserving or life-enhancing. (He may be "dumb" but he can't be stupid; that is our liability as reasoning human animals.) He acts as if he were guided by reasons; but we can act not just quasi-rationally but fully so.

Forgive me for reminding you of the enormous practical benefits generated by the sophisticated rational superstructure that we erect on the foundation of minimal rationality. (It takes a human, reasoning animal to forecast tomorrow's weather, even if he uses a computer and still gets the wrong answer.)

Are you, as a human animal, doomed to be minimally rational (to shun danger and seek pleasure) as mindlessly as any dumb brute? No. You may, if you wish, regard even minimal rationality as a disease of civilization. As a child you had no choice, but now you are free to try to divest yourself of your acquired habits of recognizing and respecting basic reasons.

But the price will be heavy. It is too late to rely on animal instinct and unrealistic to become someone's domesticated pet. To fight against even the quasi-rationalty of speechless animals would be to choose deliberate self-stupification. To purge one's mind of reasoning and its requisite conceptual and cognitive skills in order to live in the immediate present is to court the horrible apathy of the catatonic.

Your life would then be predictably short and nasty—remember the "wild boy of Aveyron." Anybody who ignores traffic

signals or hurtling bicycles on the doctrinaire ground that attention to reasons is prejudice and superstition won't live very long without the protection of friends or the security of a consistently protective setting. Short of seclusion on some island of lotus eaters or, more plausibly, in some padded cell, the price of principled irrationalism is a messy and early death.

You wouldn't go that far? You are not *that* committed to principle? Well, it's your choice. I hope you will be more fortunate in your friends than in your principles. Let me hear from you again, if you are still around to send greetings.

Appendix: Reasons as Metaphorical Supports

A reason (something-that-is-the-case, a *circumstance*) is so naturally conceived as supporting its *probandum* (as I shall call it) that we normally have little occasion to examine the backing metaphor. Yet the word "support," considered out of context, primarily evokes *physical* support, as when a buttress props up a cathedral wall. It is worth taking a close look at the background metaphor.[13]

The metaphorical statement "A reason is a support" determines something like the following implication complex in its secondary subject: A support is ($S1$) a material body that ($S2$) prevents another body from falling, or ($S3$) helps to do so. The parallel claims about the primary subject would be: A reason is ($R1$) a circumstance (something-that-is-the-case) that ($R2$) provides sufficient evidence for a probandum's truth, or ($R3$) helps to do so (by raising the probandum's likelihood).

When so elaborated, the metaphor looks farfetched; evidence for something is not much like force exerted; staying upright has little resemblance to being true; and so forth. My purpose, however, is not to discredit this ramshackle metaphor, but only to consider how far the alleged infinity of reasons is paralleled by a corresponding series of *material* supports.

Three points deserve special emphasis.

1. Although a material support will often itself be supported or

prevented from falling, there never is, nor could there be, an infinite series of progressively remoter supports. The roof of a house is supported by its walls; they in turn are supported by the foundation; and it is supported by soil or rock; but when we proceed in this way we get no further than the whole earth, which neither needs nor receives any support. The old fantasy of the supporting elephant, itself supported by a tortoise, after which there are "tortoises all the way," is no more than a joke.

2. Whether the walls do actually support the roof, in our illustration, we determine *hic et nunc* by considering what would happen if the walls were removed, and is quite independent of whether the walls are themselves supported. (Similarly, in the important case of the alleged infinite regress of causes: whether *A* is a necessary cause of *B* is determined by a local inspection of the consequences of imagining *A* removed and is quite independent of the issue whether *A* is an intermediate or a "first" cause.)

3. In the case of bodily support, it is easy to describe—contrary to initial intuition—a system of reciprocally supporting supports. An amusing example would be that of a number of persons standing in a circle, each with his back half turned to the next: on a word of command, all lower themselves and hey presto all are sitting, each supported by the lap of one neighbor while simultaneously supporting another. (If one member of the ring were removed, all would collapse.) A still simpler example would be that of three posts leaning against one another to form the framework of a wigwam—or two leaning boards forming a hutch open at both ends. In all such cases, the postulate of a serial arrangement of the interacting bodies is violated, without detriment to the stability of the entire system.

The three features of systems of material support briefly described above might be dubbed: (1) finiteness of support chains; (2) proximate or "local" influence of supports; and (3) reciprocal influence. In view of the merely analogical resemblances between material supports and supporting *reasons,* we cannot assume without examination that all three features can characterize *rational* support (support of probanda by reasons). But in fact they do.

1. As in the case of the unsupported earth, able to support

bodies resting upon it, we can envisage unsupported reasons able to support a probandum. Let *R* be a reason for the probandum, *P*. Then whether *R* really does "support" *P*—that is, whether the truth of *R* tends to make *P* more likely to be true than if there were no such circumstance—depends only on the relation between *R* and *P*. Thus the fact—if it is a fact—that a hundred balls drawn at random from an urn (with replacements) were all black (*R*) itself supports the probandum that most of the balls in the urn were black (*P*), without reference to any *reasons* for taking that *R* to be a fact. (I might simply be convinced, and correctly, that the *R* in question was a fact, but my conviction is not itself a *reason* for so regarding *R*.)

2. The last example illustrates the analogue, in the case of rational support, of my previous claim that the reality of material support is establishable *hic et nunc*. Of course, we cannot literally *see* that the above-cited sampling fact supports the urn generalization: the link between *R* and *P* is a generally accepted principle about the probative influence of a random sample. Still, the analogue of the second feature seems to be sound: so long as the sampling fact really is a fact, it renders the generalization probandum more credible than otherwise, regardless of the support that the *R* in question might or might not receive.

3. For an analogue of the reciprocating effect of mutually supporting bodies, we need only to produce an example of three propositions each two of which entail the third (which might be called a "prologism"—a kind of opposite to the more familiar "antilogism"). The following would be such an example:

Let a = n is a multiple of 6
b = n is a multiple of 15
c = n is a multiple of 10

Now if *a*'s truth and *b*'s truth were independently established, *a&b* would be a (conclusive) reason for asserting *c*—and similarly for *b&c*, and so on. (Examples not involving strict entailing relations between the members of a prologism are left as an exercise for the reader.)

I hope the foregoing discussion, incomplete as it is, may serve to inhibit what might be called the fallacy of misplaced regression, that is, the view that something of the form *aRb* cannot be

known to be true (or cannot even be justifiably asserted) unless we also know that for some *c*, we have *cRa* and so on ad infinitum. To believe that an empirical proposition cannot be rationally supported unless an infinity of other propositions is already known to be justifiably assertible is really as absurd as believing that one cannot identify somebody's parents without knowing his whole family tree.

2/

Reasonableness

Preliminaries

Many say and some believe that it is a good thing to be reasonable. Yet it would be heroic folly to aspire to be reasonable in all situations: those who wish to be as reasonable as it is reasonable to be may well feel uncomfortable with Reason's traditional status as a supreme value.

How far reasonableness is praiseworthy, and on what grounds, and what limitations there are upon its scope, are questions that cannot fail to engage our interest. But lurking behind them is another question, still more fundamental and at least as difficult: *What is it to be reasonable?* It is to this mind-twister that I shall address myself.

Let us remind ourselves of the kind of answer that has traditionally been offered. Two examples will suffice.

The distinguished sociologist Morris Ginsberg once explained that "a rational person is one . . . who is able to bring to bear on a particular situation a knowledge of general rules which he has laid down for himself on the basis of his experience and insight."[1] We are entitled, I think, to treat "rational" in this passage as a mere stylistic variant upon "reasonable." Now Ginsberg's criterion, if taken strictly, has the depressing consequence that hardly anybody is ever in a position to be "rational." Let the reader ask himself how often he can apply to a particular

situation rules that he has "*laid down for himself*" on the basis of "experience and insight." (And is it to matter whether the rules are correct?) Independent discovery of applicable rules is so rare that to insist upon it would make rationality as scarce as handcraft. With rules as with instruments, we must usually employ those that others have made. The best we can hope for, nearly always, is to choose wisely, not to create.

Another distinguished thinker, the late Karl Mannheim, once defined "substantial rationality" as "intelligent conduct based upon one's own insight into the connections between events."[2] Here we have the same emphasis upon the agent's originality ("one's *own* insight") already criticized above, with the extra complication of an invocation of intelligence ("*intelligent* conduct"), an idea no less obscure than the notion of rationality it is supposed to clarify.

Such statements as those quoted illustrate the common fault of inflated definition. Fired by enthusiasm for an ideal of "Reason" or "Rationality," a thousand writers have offered persuasive definitions of what they think *ought to be valued*. In this way zeal breeds distortion and sober analysis is supplanted by high-minded rhetoric. Too often also, the path to these inflated answers is what C. S. Peirce called the "high priori" way: we are invited to treat as self-evident what badly needs argumentative support.

I shall undertake here the more modest task of beginning to explore *what we actually mean* by "reasonable." I start with the adjective rather than the noun, because it resists the reification that "Reason" positively invites. Our topic is the reasonable, with a small *r,* not Reason with a capital *R*.

My replacement of the original question "What is reasonableness?" by its linguistic correlate "What do we mean by 'reasonable'?" is deliberate. For an answer to the latter will necessarily be an answer to the former: I know no better way of clarifying what philosophers have called the "concept" of something than by exploring the ways in which mastery of the concept is manifested in uses of the words expressing that concept. To render explicit how "reasonable" and related words are used is *ipso facto* to anatomize reasonableness: there are not two tasks, but one. This sober linguistic approach to our problem has the advantage also of providing specific tasks, and difficult ones at that.

The full program, then, would be to render explicit some of the rules governing the uses of "reasonable" and such closely related words as "reason" and "justification." Some such rules will look somewhat superficial, in their concern with grammatical categories, or with what linguists call "constraints upon co-occurrence" (rules determining the linguistic associates of "reasonable"); other rules, specifying conditions in which "reasonable" may be properly applied, may seem more probingly related to the meanings of "reasonable." But both types of rule, "syntactic" and "semantic," in showing how we use "reasonable," help to clarify its meanings.

As to evidence, I can rely only upon the reader's knowledge of what can properly be said in English or in any language into which English can be translated. That the reader possesses the requisite knowledge, in common with the writer and other competent English speakers, is certain. That still leaves plenty of room for disagreement about details, since it is so much easier to use a word than to explain *how* one uses it. But let us begin.

What Kinds of Things Are "Reasonable"?

Since "reasonable" and "unreasonable" are adjectives, they come attached to nouns that they "qualify." But which nouns? In a phrase of the form "a reasonable (unreasonable) *N,*" which nouns, when substituted for *N,* will make sense? Some will not do so: it is nonsensical, in a primary context, to speak of "a reasonable wart," or "an unreasonable star," although we might succeed in attaching some metaphorical sense to each of these remarkable expressions.

By a "primary context" I mean a situation in which a speaker seriously, literally, and prosaically uses the expression in question, but without talking *about* it. If there were no primary contexts for a word's use, there could be no secondary uses either: jokes and lies, tropes and word play, are parasitical upon the plain speech of primary contexts. But the "secondary" resources of speech are inexhaustible: any string of words could, with sufficient ingenuity, be used in an intelligible utterance. A cat

might jokingly be called unreasonable; "a reasonable *if*" probably occurs in some poem of e e cummings; the sentence "a reasonable footprint makes no sense" makes good sense. In such cases, the context is "secondary" or abnormal: the hearer needs special information about the speaker's intentions in order to make sense of what he hears.

Among the admissible substitutions for *N* in "a reasonable *N*" are the following: "man," "woman," "child," "parent," "husband," "chairman"—but not "horse," "fish," "table," or "lake"; also "demand," "question," "objection," "offer," "compromise," "explanation," "expectation," "prediction," "claim"—but not "change in weather," "rise in cost of living," "memory," or "knowledge." "Plan," "proposal," "strategy," "design" will fit—but not "shape," "consequence," "configuration," "constellation"; "belief," "hope," "fear," "resentment," "disappointment" will do—but hardly "joy" or "pain." (Some of the nouns cited will couple more readily with "unreasonable" than with "reasonable.")

We might usefully supplement such preliminary and unsystematic samplings by seeking admissible substitutions in the expression forms "It is (would be) reasonable (unreasonable) to *X*." In this way we would get specific instances of the general categories elicited by the first test. (Cf. "a reasonable demand" with "it is reasonable to demand an explanation for an overcharge.")

These simple exercises soon reveal a great variety of nouns or nominal expressions to which "reasonable" and "unreasonable" can be sensibly attached: they include references to persons, sometimes specialized with regard to sex ("man," "woman"), age ("child"), status ("husband") or function ("chairman"); references to some but not all actions—including those of demanding, questioning, objecting, and so on; or, by a natural extension, to what the content of such actions might be—the demand *made,* the question *asked,* and so on; also various attitudes and feelings. This catalogue is of course incomplete.

Yet if "reasonable" and its opposite seem uncommonly hospitable in their linguistic alliances, they are not wholly promiscuous. A common thread unites this miscellany of instances: all

have to do, more or less directly, with human beings, and more specifically with what human beings can *do* (actions) or with what they can *have* (feelings, attitudes). I shall ignore here anthropomorphic extensions to higher animals—or, in another direction, to spirits, angels, and deities. This observation serves to explain why such things as tables, oceans, or stars are ineligible for appraisal as "reasonable" or the reverse.

Further discriminations are needed, since some things that human beings do, such as breathing or stumbling, cannot be called "reasonable"; and the same holds for some but not all feelings and attitudes. Let us begin with actions, as being more easily discussed than features of "inner life." We may usefully employ two techniques in searching for what makes actions reasonable: the examination of linguistic affiliations and repugnancies—and the construction of "paradigm cases."

Affiliations and Repugnancies

Actions amenable to appraisal as reasonable or the reverse can also be characterized by a large variety of other adjectives: "rash," "well considered," "impulsive," "far-sighted," "intelligent," "sensible," "foolish," "prudent," "wise," "dangerous," "futile," "successful," "pointless," "inconsiderate," "clumsy," "clever," "imaginative," "willful," "irresponsible," "wicked," "vicious," "irrational," and many, many more. Most, if not all, of these imply approval, like "reasonable," or disapproval, like "unreasonable": they may be called "appraisal adjectives."

I propose now to find a way of testing how these appraisal adjectives harmonize or clash with "reasonable" and "unreasonable" themselves.

Consider an expression of the form "*A* but not *B*," where *A* and *B* are adjectives or adjectival expressions. One of the obvious uses of the ". . . but not . . ." locution is to defeat an otherwise warranted presumption. For example, on learning that something is made of wood, we are entitled in the absence of further

information to presume that the object can float in water; however, the form of words, "made of wood *but not* able to float," defeats that presumption, while acknowledging its presence. It has the approximate force of "Since the object is made of wood you might expect it to float, but for all that it doesn't!"; similarly for "well designed but unsuitable" and any number of other admissible instances of the "but not" form. This explains why "made of wood but able to float" sounds absurd. Incidentally, the admissibility of an instance of "*A* but not *B*" shows that being not-*B* is logically compatible with being *A*, since otherwise "*A* but not *B*" would be a flat contradiction, like "husband but not married"; the presumption of being *B* when *A* comes only to this: knowing that something is *A*, we are entitled, other things being equal and in the absence of special information, to think that it will be *B*. Alternatively: knowing that an object is *A* is some reason to think that it is also *B*.

Let us now apply these ideas to the formula "reasonable but not *X*" where *X* may be any appraisal adjective. We obtain the following selection of admissible values of *X*: "successful," "prudent," "wise." And if we apply the same test to the positive variant "reasonable *but X*," we get, among others: "dangerous," "futile," "short-sighted." It is significant, however, that the following seem to be inadmissible values for either formula: "rash," "foolish," "futile," "stupid," "inconsiderate"—and, on the other hand, "well considered," "intelligent," "sensible," "far-sighted," and so on. We could obtain further insight on the logical connections between these adjectives and "reasonable" by considering which of them could be admissible values in "reasonable *and X*," "reasonable *and hence X*," or their negative variants. The reader may be able to satisfy himself that the connections between the meanings of the nine appraisal adjectives last mentioned and the presence or absence of reasonableness are tighter than those of mere presumption.

Further work of this sort might lead us to some such summarizing formula as the following:

> In order for an action to be reasonable it must be well considered, intelligent, sensible, far-sighted, and so on, and may be presumed to be prudent, wise, and so on;

> while its being rash, foolish, futile, stupid, or inconsiderate will normally disqualify it for praise as "reasonable."

This is all very rough as yet, and inaccurate without some rather obvious qualifications. But my aim is to illustrate a potentially useful technique, rather than to apply it in boring detail. In support of the recommended procedure of searching for "affiliations and repugnancies," we can adduce the well-known and somewhat obvious semantical principle that the meaning conveyed by a word is a function of its admissible contrasts: when a speaker calls something "reasonable," the information he conveys depends upon how many other things he might have chosen to call that action instead. (Thus, if the only contrast intended were with "unreasonable," the information conveyed would be minimal.) A detailed exploration of what I have called the "affiliations and repugnancies" of "reasonable" would help us, perhaps better than a formal definition, to understand its role in our language. If such an undertaking also aroused interest in such linguistic cognates as "prudent," "sensible," and "intelligent," so much the better. For "reasonable" is a nodal word—a place at which many semantic threads are knotted together. Like other words, it has a trailing set of friendly and hostile cognates. More generally, our utterances always imply *this but not that* and *this, so* (necessarily, presumably) *that*. And the consequent incompleteness of all explicit formulations—as shown by the possibility of unending elaboration of the tacit implications, presumptions, and suggestions—is inescapable.

Reasonable and Rational; Unreasonable and Irrational

To use "rational" and "irrational" as pretentious variants for "reasonable" and unreasonable" is to blur some important differences of meaning. The tests we have just used can help to highlight these. I hope the reader will agree that "unreasonable but irrational" makes no sense, and that there is something wrong about "unreasonable and irrational" and about "unrea-

sonable, and hence irrational." Similar dissonances arise from trying to combine "rational" with "reasonable," although it is somewhat easier to treat these positives as straightforwardly synonymous.

One has the impression that in stigmatizing something as irrational—and the censure is part of the meaning—we are going beyond dianoetic[3] appraisal and are moving, as it were, on a different plane. This impression is confirmed by the discrepancies in the ranges of co-occurrence of the two words: we talk about an "irrational impulse," but surely not—or not so freely—about an "unreasonable impulse."

The explanation seems to be that in calling an action unreasonable we are implying that the agent's reasons are unsatisfactory. So we are at least implying that he *has* reasons to be criticized: the reproach pays him the compliment of supposing him able, as we say, to "listen to reason." But to call him irrational is to imply a more severe disorder: we diagnose his action, rather than take issue with it. (Yet, interestingly enough, we still reprehend it.) An irrational act calls for therapy or restraint rather than for argument, and this may be why we are willing to call irrational what is beyond the agent's control. I can say no more here about this interesting topic.

The Method of "Paradigm Cases"

A more direct attack upon the meaning of "reasonable" and its opposite can be made by trying to describe what philosophers call "paradigm cases"[4]—examples of the use of a word so clear and unproblematic that they exhibit part of that word's meaning. A paradigm case of the use of "reasonable action" would be a real or imaginary action of which we could properly say "*That* is a reasonable action, if anything is!" Such an exemplary instance could be used as a standard—to remind ourselves of what we mean by "reasonable," to test another's understanding of the word, or to teach its meaning to a child. For to refuse to apply

"reasonable" to such a paradigm, when fully acquainted with its character, would be to fail to understand how we use the word and so to betray partial ignorance of the English language or to manifest some eccentricity of usage.

But if we now try to produce some paradigmatic cases of reasonable action, we find an unexpected hitch. It is relatively easy to describe paradigmatic—clear, unproblematic, exemplary—instances of *un*reasonable action, but surprisingly hard to find equally uncontroversial instances of the reasonable. One may suspect that "unreasonable" is the word that "wears the trousers," to use a phrase of J. L. Austin's, and that the root idea of "reasonable" is no more than "not unreasonable."[5] At any rate, it may do no harm to begin with clear cases of unreasonable actions in the hope of capturing the notion of reasonableness by indirection.

I accordingly propose the following as clear, paradigmatic cases of *un*reasonable actions:

(a) To look for somebody's telephone number in the directory, when you know that he has no telephone.[6]

(b) To buy a dog if you don't like dogs and have no use for one (say as a watchdog).

(c) To look for pound notes on the pavement.

(d) To guess the date of the siege of Vienna when a reference book with the information is at your elbow.

(e) To assert before you have looked into a bag that it will contain black marbles (assuming that you have no special knowledge about the bag).

(f) Having drawn ten white marbles from the bag, to claim that the next to be drawn will not be white.

(g) To complain that a man is ungrateful for an injury you have done him; and so on.

In each case,[7] we can explain why the verdict of unreasonableness is in order, though the fault varies from case to case: futility (a), pointlessness (b), action on infinitesimal probability of success (c), choice of defective means to the end in view (d), absence of sufficient reason (e), action in conflict with evidence (f), action based upon assumptions plainly false (g).

The Analysis of Reasonable Action

Reflection upon the failures of reasonableness in the examples listed above and upon similar examples which I omit suggests the following budget of necessary conditions for the reasonableness of actions.

1. *Only actions under actual or potential control by the agent are suitable for dianoetic appraisal.* Behavior that may be thought undesirable, such as irrepressible grimaces or clumsy and uncoordinated gestures, can be called reasonable or unreasonable only if the person in question can be held *responsible* and so able, at least in principle, to exercise choice, decision, and restraint. Such actions may be called voluntary, with a side glance at the slipperiness of that word; they may, however, be performed without conscious intention or calculation. Here we may recall that we do not speak of reasonable sensations or of unreasonable physical pains, and it would make sense to speak of a reasonable dream only if the dreamer had the extraordinary power to choose what he dreamed.

2. *Only actions directed toward some end-in-view can be reasonable or unreasonable.* Where nothing is hoped of an action, nothing intended, it is absurd to praise it as reasonable or to denigrate it as unreasonable. This is perhaps why it sounds so strange to speak of "reasonable walking" or "reasonable eating," *simpliciter,* although both activities are voluntary; and "reasonable painting" is just as odd, because there is no definite end-in-view. But it is quite natural to say "It is reasonable to eat more in order to reduce your indigestion" or "It would be reasonable to walk in order to have a better chance to catch a bus." Reference to an end-in-view at once reinstates the propriety of dianoetic appraisal.

3. *Dianoetic appraisal is relative to the agent and to his choice of end-in-view.* It may be reasonable for you to buy a bicycle if you can afford one and enjoy cycling, and unreasonable for me to do the same if I am crippled. And the very same action, say answering a question about the time by saying "Buddha is a dung heap," may be reasonable or not under different aspects. It makes a difference whether you want the interlocutor to know the time or whether you hope to prod him into satori. The Zen masters are not

necessarily unreasonable, however absurd their actions may seem.

Here is a convenient point at which to distinguish between reasonableness (or the reverse) *from the agent's standpoint*—and the dianoetic judgment of a competent observer. What may be thought reasonable, given the agent's limited knowledge, may still be judged unreasonable from a more comprehensive perspective. But we shall soon see that the agent's subjective judgment of the reasonableness of what he is doing cannot be insulated from—indeed, positively invites—the more objective verdict of a disinterested and better-informed spectator.

4. *Judgments of reasonableness are appropriate only where there is partial knowledge about the availability and efficacy of the means.* When we are quite ignorant of what to do (which we almost never are) any action is as reasonable as any other—that is to say, in one sense not reasonable at all.

When we know the sole means to achieve our end, there is a weak irony or facetiousness about calling the obligatory action "reasonable." There is nothing reasonable about dowsing a fire—though it would be highly unreasonable not to do so. Nor is there anything merely reasonable about accepting the answer of a correct addition sum. Here, "reasonable" is an implausible meiosis.

Dianoetic appraisal, to make the point in another way, is characteristically appropriate when we have to act in conditions of uncertainty (as we do most of the time). This may be why "reasonable" will not combine with such "achievement words" as "know," "remember," "see," "recognize." The natural habitat of "reasonable" is a situation in which a task is to be performed, with the outcome still in doubt.

Among the obviously relevant criteria for reasonable action in uncertainty are the likelihood of success and the cost of the available means in relation to their probable efficacy and the value of the outcome. It is unreasonable to try slicing meat with a very blunt knife, when a sharper one is at hand; and it is unreasonable in another way to pay an expert $50 to help you save $5 on your income tax.[8]

It follows that dianoetic judgments are—or perhaps ought to

be—typically *comparative:* if we can sometimes say that one action is more reasonable than another, we are very seldom able to choose a single action as uniquely reasonable.

5. *Dianoetic appraisal can always be supported by reasons.* In the jargon of philosophers, "reasonable" and "unreasonable" always refer to "supervenient" characters of a situation: the question "What is reasonable (unreasonable) about that?" must always admit of an answer. If what has been said under the previous heading is right, a defense of a judgment recommending one action rather than another will typically invoke relative probabilities and relative utilities, in some plausible combination. "It's reasonable to do *A* rather than *B,* because it's more likely *to achieve what you want*—and, besides, it costs no more."[9] Now the moment we begin to argue in some such way, we produce reasons, and reasons are no respecters of persons. Where we have taken account of the agent's own choice of ends (point 3 above) and his own scale of relative values ("utilities" and "disutilities"—rewards and hurts), we leave no room for further subjectivity. An acceptable answer about reasonableness, relative to a given end and a given set of preferences, must be as impersonal, as "objective," *if it exists,* as a question about the right conclusion to a logical argument. Thus all attributions of reasonableness, however casually or privately made, have a built-in thrust toward objective justifiability. Whatever is reasonable must also be justifiable, that is to say supportable by reasons that *any* unbiased impartial judge would have to accept. Unfortunately, this doesn't give us much practical help in deciding what is reasonable, or relatively more reasonable, in problematic cases; for the notion of "a reason" is at least as controversial as the topic of this essay.

6. *To call an action reasonable is to approve it, and to call it unreasonable is to censure it.* This point has been implicit throughout this discussion in my frequent reference to "reasonable" as a term of *appraisal.* It is worth making explicit, however, that the praise is part of the standard meaning of the word (as also in the cases of "fine," "good," "splendid," and the like) and can be overridden in speech only by the use of irony, paradox, or some other ad hoc rhetorical device.

Some philosophers say that the method of paradigm cases fails

for "evaluative expressions," at least in establishing uncontroversial cases of the required evaluation: they grant that factual or "descriptive" conditions for application are determined by linguistic convention, but hold it essential to evaluation that the agent be left free to *make* his own evaluations. J. O. Urmson, in an influential paper, says: "There is thus a close logical connexion between an evaluative expression and the accepted standards for its appropriate use; but this cannot be identity of meaning, *for no evaluation can be identical in meaning with a description*" (my italics).[10]

This is, however, to take an indefensibly narrow view of meaning and of the purpose of appeal to paradigms. To the extent that approval is built into the proper use of "good" in certain situations, such approval is part of its meaning; and the same is true of "reasonable." In finding paradigm cases of the use of "reasonable," we are therefore producing cases where by linguistic conventions a speaker counts as approving of what he labels "reasonable." If we don't like the implied set of standards, we can try to change the language, avoid using the word in question, or override the prima facie connotation of approval by using appropriate devices. But that is another story.

I once heard some students condemn a talk about student grievances as "too reasonable." Intelligible as this comment was, it depended upon the general understanding that to call something reasonable is, *so far,* to praise it. Saying of somebody that he is "too virtuous" is an instance of the same device.

Yet if to dub something as reasonable is, at least provisionally, to praise it, the praise is faint and has in most contexts something depreciatory about it. A reasonable action, as we have seen, may turn out to be unsuccessful—or even *wrong,* however justified: a reasonable move in a game of chess might be just the one to permit a crushing yet unforeseeable combination. There are often better things for an action to be than merely reasonable, even if reasonableness is the best we can aim at. "Intelligent" and "imaginative" are rather warmer, while "reasonable" has something of the tepidity of "sensible," "defensible," and "respectable." In moral contexts, and some others, "right" is a good deal stronger than "reasonable."

It begins to look as if reasonableness is a somewhat humdrum,

pedestrian virtue, involving as it does a problematic calculation of probabilities and expected values in situations of inescapable fallibility. This is why "reasonable" with a small *r* smacks too much of the prudential for some tastes.

The Reasonableness of Feelings and Emotive Attitudes

Our preliminary investigation into the reasonableness of actions could be summarized by some such formula as the following:

> In a situation in which it is uncertain which action to take, an action is reasonable if there is some sufficient reason to take it, and no better reason to choose one of the alternatives.

Here the reference to "reasons" points implicitly to the complex tangle of factors—conduciveness to the end-in-view, probable efficiency, cost (in the wide sense), and the like—which I have mentioned.

If we try to apply some such formula as this to the case of feelings, in the belief that use of the same label of approbation is no accident, we shall encounter certain difficulties. For one thing, the notion of conduciveness to an end-in-view will have to be replaced by something like appropriateness, although the root idea of justifiable reason will still survive. We call a fear unreasonable, for instance, if the object feared is not really dangerous enough to justify the degree of fear evoked. Here, perhaps, it is not difficult to see a connection with the pattern of reasonableness in actions proper.

A more serious difficulty is that it is commonly agreed that feelings and emotive attitudes are not so plainly under control as actions. It is hardly fair, one might object, to call a fear unreasonable, and so to condemn it, if there really is no way to repress that fear. Well, its manifestation at least is more or less under control: a man beset by cowardly fears may still be able to act bravely. And an appraisal of the feeling itself may appeal from the unbridled immediacy of momentary feeling to a cooler verdict in

the future—from Richard drunk with emotion to Richard soberly recollecting his feelings. The long-term influence of such appraisal should not be underrated. David Hume said that reason was the slave of the passions, but the desire to be reasonable may itself be a "passion," able to contend with other passions and sometimes to curb them.

Being Reasonable

We have proceeded far enough to see, at least in outline, how reasonable behavior looks. A man will be acting reasonably to the extent that he tries to form a clear view of the end to be achieved and its probable value to him, assembles the best information about available means, their probable efficacy, and the price of failure, and in the light of all this chooses the course of action most strongly recommended by good reasons. (A similar account, suitably modified, will apply to the choice of attitudes and ends, to the extent that choice of these is possible.) Such a posture has quasi-moral implications, since the requisite calculation demands a sustained effort to suppress bias and, in striving for impartiality and objectivity, to pay a decent respect to the opinions of others. To act reasonably is to be willing to reason and thereby to submit to impersonal judgment. But a morality entirely based upon reasonableness would be at best anemic and frigid: a man can reason and reason and still be a villain.

To many, and perhaps especially to the young, there will be something unattractive about the character of dispassionate deliberation here evoked: the reasonable man will smack too much of the conscientious judge. But consideration, respect, kindness, mercy, and love are not really competitors for justice—except for the muddled: if perfect love casteth out reasonableness, that is no reason for being unreasonable. The case for being reasonable—when the question properly arises—is that there is nothing better to do. That, of course, is a consideration that will appeal only to those who are already at least potentially reasonable.

Part II

Aspects of Science

3/

The Objectivity of Science

The Questions

My object is to raise such fundamental questions about the nature of science as: is science a search for truth? Does science succeed in progressively approximating to truth? Does it provide an accurate map of reality? Does science yield information about objects having properties and relations independent of our choices and desires? Or is science to be regarded as no more than another elaborate myth? Is science, after all, no more than a fiction—glorious or pernicious, or both at once? In short, I am raising some basic questions about the so-called objectivity of science.

I shall use "science" throughout as shorthand for *scientific activity*. I want to think of science most of the time as investigation and research, something that is being done, rather than the technological or theoretical product of the activity.

Let me begin by persuading you, if you need persuading, of the importance of such questions. For they may seem very abstract and remote from the practical urgencies of our time. Abstract and remote from practical concerns they certainly are, but they are not unimportant or trivial. The answers that we adopt, whether unthinkingly or after full consideration of the complex issues underlying them, have important bearings on the values we attach to science. They help to determine our conception of

science as a cultural as well as a practical achievement. How we think about the objectivity of science will affect our views about its place in a well-rounded education, its social role, and much else. Time spent in becoming clearer about these matters will not be wasted.

Some thirty years ago, a remarkable Polish writer, Leon Chwistek, opened his book *The Limits of Science* with the following words:

> We are living in a period of unparalleled growth of anti-rationalism. Exact thinking based upon the principle of consistency is the sacrificial goat to which all the disasters of our time have been imputed. . . . Exact thinking is blamed for drying up the sources of the sacred enthusiasm and for causing the emotional exhaustion of our epoch. Exact thinking, it is alleged, has become the source of the excessive growth of materialistic culture, as well as the shrinkage and sterilization of spiritual life. The demand for a new logic, for new laws of thought more suitable to the needs of spiritual life has become the hobbyhorse for a whole galaxy of obscure and false doctrine. . . . We have entered into a period of incredible abasement of science, a period of the noisy superiority of groups of puffed up eulogists of irrational nonsense who are leading mankind toward open crime and violence—as a rule unknowingly but often quite consciously.[1]

Chwistek would have been appalled to know that his denunciation would still apply to the conditions of our intellectual life over thirty years later. We are living in a period of great and widespread contempt for science and contempt for the rational mode of thought of which science is a supreme embodiment. Let me take, as one example of many, some of the remarks made by Theodore Roszak in his influential book *Where the Wasteland Ends*.

In speaking of the "synthesis" achieved by Newton, Roszak says: "The world picture was homogenized and qualitatively flattened. For the first time in the Western tradition, the universe became in the most objective sense of the word a 'Universe': a cosmos uniform throughout its infinite extent, devoid of ethical

nuance, of magical centers or charmed circles, of quintessential terrain." Roszak speaks of the "nature" revealed by science as being "like a value-neuter screen on which only the measurable behavior of things might be registered." A little later he talks of "the crushing inhumanity of Western science." He says that science has produced "a nature that was felt to be dead, alien and purely functional." And in a most revealing passage he says: "To cast aside all the pre-scientific and non-scientific realities by which men and women have lived for so long is to settle for a truth that is little more than an operational superficiality; worse, a license for the making of well-informed fools."[2]

Roszak's book is well worth reading as an eloquent expression of deep revulsion from the values that underly and support science. Nevertheless, it is ironic that at a time when science is achieving unparalleled theoretical and practical successes, a thinker as eloquent and sincere as Roszak should be able to regard it as "a license for the making of well-informed fools." Discounting the ulterior motives that underpin such vehement attacks on science, I wish to consider some of the reasons that have provoked serious doubts about the entire scientific enterprise.

Science as a Cognitive Enterprise

Science, as its very name indicates, has been traditionally regarded as a *search for knowledge*. Consider, for instance, the following definition in a useful philosophical dictionary:

> *Science:* an aggregate [*ensemble*] of knowledge and investigation, having a sufficient degree of unity and generality, and capable of leading persons who dedicate themselves to it to agreed conclusions, resulting neither from arbitrary conventions nor from individual tastes and interests which they share, but based upon objective relations, gradually discovered and confirmed by definite methods of verification.[3]

Notice the prominence rightly given in this definition to "knowledge."

There was a time, not very long ago, when people hoped that science would liberate us from superstition and prejudice. Consider, for instance, the enthusiasm with which a great scientist, Ernst Mach, spoke of the achievements of Galileo:

> It was toward the close of the 16th century that Galileo, with a superb indifference to the dialectic arts and sophistic subtleties of the Schoolmen of his time, turned the attention of his brilliant mind to nature. By nature his ideas were transformed and released from the fetters of inherited prejudice. At once the mighty revolution was felt, that was therewith effected in the realm of human thought—felt indeed in circles far remote and wholly unrelated to the sphere of science, felt in strata of society that hitherto had only indirectly recognized the influence of scientific thought. And how great and how far-reaching that revolution was![4]

Mach would not have written with such fervor about the admirable work of Galileo (and of another hero of his, Charles Darwin) unless he had assumed as too obvious for argument that scientists are concerned with what Mach called "the growth of natural *knowledge*."[5] For if Galileo and the others have not been engaged in searching for knowledge, what *have* they been up to? Providing, as advanced thinkers like to say nowadays, another "myth"? Providing some elaborate fiction that may have practical applications? If we seriously believe that, we had better revise our ideas about the power of science to release us from the fetters of inherited prejudice. For does not *belief* in a myth have to count as "superstition"? Let us therefore dwell for a little while on what is implicit in the cognitive claims of science—its claim to supply us with knowledge, or at least with an approximation to it.

Knowledge and Truth

Knowledge, like other fundamental concepts that we shall be examining, turns out to be extremely elusive. One difficulty is that we use the verb "to know" and its cognates in a number of

distinct, if related, senses. We talk of "knowing *how to* drive a car," where the reference is to a skill or aptitude; we speak of "knowing somebody," where the reference is to personal acquaintance; we speak of "knowing what hunger *is like,*" where we have in mind some direct experience; and so on. There are, however, important contexts, to which I shall restrict myself, in which we use the verb "to know" in another fashion, in order to make a *truth claim*. If I say, "I know that today is Monday," there is a logical implication that if what I say is justified, then it is *true* that today is Monday—today must really *be* Monday if my knowledge claim is to be validated. And so in general. In this sense of knowledge, the one most immediately relevant to our discussion, the concepts of *knowledge* and *truth* are logically connected. So if you happen to think, as so many people do nowadays, that science does not and cannot aim at *truth,* you will have to abandon any thought that science aims at *knowledge*. Hence another way of specifying what might be called the classical conception of science is to say that science is a search for truth. Of course, such truth must somehow be made accessible to us, and that is why we are concerned not merely with truth, but with *known* truth. As a Russian proverb puts it, "Milk in a cow's udder never slaked anybody's thirst." For all practical and theoretical purposes, truth inaccessible to human endeavor might just as well not exist. That is why reliable methods of getting at the truth are indispensable components of science.

In the past, the contention that science is committed to a search for truth would have been regarded as too obvious to call for discussion. I was therefore taken aback to find a recent writer, J. Ravetz, in an elaborate treatise titled *Scientific Knowledge and Social Problems,* speaking almost casually of "the obsolescence of the conception of science as the pursuit of truth." On the next page Ravetz says, "The experience of modern scientists in their work, seeing the rapid rate of obsolescence of scientific results, makes the vision of the pursuit of truth not so much wrong as irrelevant." Later on he says, "The ideal of the pursuit of truth has functioned for generations to provide science with a defence against external enemies, and a basis for the morale and commitment necessary for the production of worthwhile work. It will

persist for some time to come, but its base in experience is eroding very rapidly."[6]

Ravetz seems rather complacent about what he calls "the epitaph of academic science and with it the ideal of truth."[7] I see no difference between "academic science" and just science, *sans phrase*. The science whose death Ravetz welcomes is science itself. With a friend like Ravetz, scientists do not need an enemy like Roszak. About truth, I would make a claim parallel to my earlier one about knowledge: once the ideal of truth is abandoned, science loses its importance and interest and degenerates into a mere fiction.

Truth and Probability

Our need to operate with a concept of truth, not only in science but also in ordinary life, can be seen sharply when we consider the consequences of abandoning truth in favor of probability. In my experience, working scientists are fond of saying to their nonscientific colleagues that science has abandoned the supposed dogmatism that goes with claims to achieving the truth, and prefer to speak always of probabilities.

This reminds me of an essay that Benjamin Franklin once addressed to young men, telling them they would obtain a reputation for judiciousness and modesty if they prefaced all their remarks by such a phrase as "I think" or "I believe" or "possibly" or "probably." For all I know, this may be a useful piece of practical advice, but consistently followed it leads to absurdities. If a scientist seriously claims that every scientific assertion must be qualified by insertion of the word "probably," we may ask him what he thinks about a probability statement such as "Copper is probably a metal." Is that probability assertion itself to be taken straight and unmodified, or should we qualify *it* by a further probability operator? If we adopt the second course, we shall have to reject "Copper is probably a metal" in favor of "Probably, copper is probably a metal." And even *that* statement will need to be introduced by another "probably" qualifier:

Probably what? Well, "Probably it is probable that copper is probably a metal."

The tongue begins to twist and the mind begins to stagger. This process produces a kind of infinite stammer, which, like Zeno's runner, never gets started. If you cannot say anything at all without prefacing it by the word "probable," you never will succeed in saying anything. However, if it is all right to say "Copper is probably a metal" without further qualification, then at least *some* unqualified scientific assertions are acceptable. One does not have to be an expert in probability theory to know that it is easier to establish the truth of assertions about copper than it is to obtain reliable assertions about the probability of statements about copper.

Another way of appreciating the point is as follows. All words in a language get part of their meaning by means of *contrast* with the other words that might be substituted for them: part of the meaning of "copper" is that copper is *not* iron, not sodium and so on. To put it another way, part of the information that we get from an assertion about copper is that other assertions about iron, sodium, and so on are logically excluded. Now a universal use of "probably," recommended by those who think this a more correct way of speaking, would destroy the background contrast, and in so doing would erase the ordinary meanings of "probability."

In ordinary life, when we hear somebody seriously qualifying an utterance by using "probably," we gather that the information in question is uncertain. However, if every statement were to be introduced by the probability qualifier, that qualifier would cease to convey information and its use would be no more than a verbal tic, like some current usages of the work "like." "I mean, like, that science, like, is like the search for truth." An indiscriminate use of "probably" really has no more to recommend it than this nervous rattle in the throat.

To be fair, I must admit that something important lies behind the idea that all scientific statements should be understood as probability statements. Those who believe this may have in mind that scientific statements have the important property of being asserted provisionally. In other words, no scientific state-

ment, however well supported by experiment, observation, and theory, is in principle exempt from further critical scrutiny and possible modification or rejection. There is no supreme court of appeal for scientific assertions. That is a correct and important point. But it has very little bearing on the feasibility of adherence to the ideal of truth, as a regulative principle for scientific inquiry.

Science and Ontology

I have made little use, so far, of the word "objectivity," which occurs in the title of this essay. This has been deliberate, since I believe that an oblique approach to the complex and perplexing notion of objectivity will be helpful. So, before we consider what we do and should mean by "objectivity," let us spend a little time considering how beliefs about the *existence* of so-called external objects are involved in the pursuit of knowledge or truth, whether at the scientific or at the common-sense level.

Let us contrast two statements, one of which we should regard as making a truth claim and the other as belonging to fiction. First, "At 7 A.M. on April 21, 1983, Ronald Reagan, president of the United States, was brushing his teeth." Then, for contrast: "On the morning of April 21, 1883, Mr. Pickwick entered a coach that would take him to Tilbury." Now, one sufficient reason why questions of truth and knowledge arise with regard to the first of these assertions but not with regard to the second is that there really *is* such a person as Ronald Reagan and there really is not and never has been any such person as Mr. Pickwick. (I believe I can hear somebody muttering, "But after all, Mr. Pickwick did exist in the mind of Dickens!" I think this a very loose and misleading way of talking.)

Imputations of existence, I am claiming, are implied by truth claims and knowledge claims. It would be absurd to say that we *know* that Mr. Pickwick took a coach on such and such a day in the way that we know copper to be a metal—or that it is *true* that he did so. And it would be just as absurd to say that it is false that

he did so: questions of truth and falsity simply do not arise in connection with fictitious characters. Students who answer questions about Pickwick in a true–false examination are displaying knowledge or ignorance about a *book,* not about a person.

I now want to generalize the point by saying that *any* assertion that embodies a truth claim must be about *something,* must have a subject matter. In making any kind of truth claim or knowledge claim, we are committed to holding that certain objects, which the assertion is ultimately about, *exist.* What the ultimate subjects of discourse are can become a topic of high controversy. Philosophers of science may contend that scientific assertions are "really" about sense data, or about future predictions, or about recondite theoretical entities, or about statistical properties, or something else. What they cannot say with any plausibility, I believe, is that scientific assertions—or for that matter assertions in ordinary life—are about *nothing at all.* In this way, ontological commitment, as philosophers say, is involved in all truth claiming, all cognitive assertion. I hope to be able to say more about the ontological commitments of science later on. However, it is time to say something about objectivity.

The Meaning of "Objectivity"

I have been arguing that scientific pursuit of knowledge is connected, conceptually, on the one hand with a search for accessible truth, and on the other with an ontology, a commitment to the existence of independently existing objects. This notion of things to be discovered, investigated, and eventually spoken about, that are what they are, independently of what we do to them or say to them, is the kernel of the notion of objectivity.

One philosophical dictionary defines the term "objective" as "possessing the character of a real object existing independently of the knowing mind, in contrast to subjective."[8] The *Oxford English Dictionary* defines the term as follows: "that is or belongs to what is presented in consciousness, as opposed to the con-

sciousness itself; that is the object of perception or thought, as distinct from the perceiving or thinking subject; hence that is or has the character of being a single 'thing' external to the mind; real."

The basic idea revealed in these definitions is of something existing and having its properties independently of human examination or intervention. The historian Leopold von Ranke once said, in a famous phrase, that the historian's purpose was to discover "*wie es eigentlich gewesen war*"—how things actually were. With apologies to Ranke, one might say that the purpose of the scientist always has been and should continue to be that of discovering *wie es eigentlich ist*—how things actually *are*. Such a "realism," as philosophers would call it, seems to me an integral part of the scientific attitude. What the mathematician G. H. Hardy said about mathematical reality could easily be adapted to express scientific commitment to objectivity: "I believe that physical reality lies outside us, that our function is to discover or observe it, and that the laws which we establish and which we describe grandiloquently as our 'creations' are simply our notes of our observations."[9] I have substituted in Hardy's statement "physical reality" for "mathematical reality," and "laws" for "theorems," and "establish" for "prove." Here we have, I suggest, a fair approximation to the scientist's credo. It is the point of view expressed by Max Planck when he said: "[I]t is of paramount importance that the outside world is something independent from man, something absolute."[10]

Let us not shy away from that word "absolute," even though it is apt to raise hackles nowadays. What is intended is a straightforward logical point. None of the interwoven concepts of knowledge, truth, reality, and objectivity that we have been thinking about can sensibly be qualified by adding "for me." It makes no sense to say, for instance, "The sun exists for me, but perhaps not for you." Truth, knowledge, and reality are not relative notions—though, of course, possession or grasp of any of them is relative to the possessor. It seems significant to me that Theodore Roszak, to whom I have already referred, says, "What is true to me is what I am *persuaded* is true."[11] Anybody who seriously believes this is, for better or worse, cutting himself off

from the enterprise of the scientific pursuit of knowledge, truth, and reality—and what is worse, cutting himself off from less technical but essential concern about knowledge in prescientific contexts. As against this, I would propose as an essential ingredient of the scientific credo a marvelous remark by Bishop Butler: "Things and actions are what they are, and the consequences of them will be what they will be; why then should we desire to be deceived?"[12]

Objective Knowledge and Common Sense

I have been urging that commitment to the ideals of truth, knowledge, independent reality—in short, the complex notion of objective knowledge—is essential to the scientific enterprise. Now, some contemporary critics of science might argue that "so-called science has proved that genuine science is impossible," reject the ideal of objectivity as an illusion, and be content to see scientific investigation demoted or even abandoned. However, the price to be paid for abandoning the pursuit of objectivity is more serious than this, since the ideal of objective knowledge is at the heart of what may be called the "common-sense view of the world."

I mean such convictions as the following, which the ordinary person takes wholly for granted before falling into the clutches of a philosophical skeptic: "I exist, but so do vast numbers of other persons, each of whom can say the same. Besides myself and other persons, the world is full of things and events that are what they are, no matter what I or anybody else would prefer to be the case: water is wet, fire burns, guns kill, and so on indefinitely. In short, I am in and part of a universe composed of enormously many things and events that I am helpless to make otherwise. Some things I know about this universe, but I am ignorant of much more, and sometimes what I take to be knowledge turns out to be only an illusion." Such a credo, to which every sane person subscribes, recognizes a distinction between the subjective and the objective, points in generalized terms toward what I

have called an ontology, and admits the possibility of knowledge, in spite of the prevalence of illusion and error. It expresses what philosophers have somewhat contemptuously called an attitude of "naive realism." The naiveté is supposed to arise from the belief that things *are* really as they look—that grass really *is* green, that hydrogen sulphide really *does* stink, and so on. Now, I hold that the naive realism of the ordinary person is *correct* in essential respects and that the "external world" really is as it appears to us. But I am now concerned not with the details of everyday ontology, but with something underlying it, which I shall call an "objective orientation" toward experience. This is so deeply entwined with our concepts, our language, and our habits of adaptation to the surrounding environment that any wholesale violation of that attitude leads to madness and self-destruction. Let me explain.

Learning "Objective Orientation"

There is good reason to think that none of us was born with what I have called the objective orientation, or, if you prefer, a feeling for reality. According to Jean Piaget, there is a period in the early history of a child's development at which no distinction is drawn between "the intrusive self" and anything else. It is a period of what might be called promiscuous acceptance, in which everything experienced is equally real or unreal. It is a stage when the distinctions between being and seeming, reality and illusion, have not yet emerged. Piaget puts it this way: the child "is a prey to perpetual confusions, between objective and subjective, between the real and the ostensible; it values the entire content of consciousness on a single plane in which ostensible realities and the unconscious interventions of the self are inextricably mixed."[13]

Elsewhere Piaget says: "The outer world [at this stage of development] is only a series of moving pictures which appear and disappear, the most interesting of which can reappear when one knows very well how to manage it (for example, crying long

enough if it is a question of someone whose return is desired). But these are only moving pictures without substantiality or permanence and, above all, without localization."[14]

It is fascinating and moving to follow in Piaget's pages the details of how a child gradually develops the notion of "the permanent object." For a very young child, say five or six months old, a watch that is covered with a piece of cloth simply ceases to exist: it makes no effort to raise the cloth or seek for the object beneath it. Later on, the child becomes clever enough to locate and seize objects that have been covered—though it takes a good deal longer to realize that this can happen on each side of the body. It takes an astonishingly long time for a child of even superior intelligence to develop "the notion of a general space . . . including all objects which have become solid and permanent, with the body itself as an object among others,"[15] enmeshed in a network of spatial relations that can be mastered by the use of appropriate gestures and movements.

It would be a gross error to think of this fundamental learning process as a gradual conversion of doubt into knowledge: it is something more fundamental than that. The infant is in no position to doubt, because no certainties are yet available. (Doubt and certainty are sophisticated concepts that emerge later.) The external world that reveals itself in the "school of hard knocks" is not a hypothesis; and the child is not yet even a primitive scientist.

Rather than knowledge, one might here prefer to speak of imprinted conviction. The process is one of learning to act, notably through painful consequences of maladapted action, in such a way as to take account through one's behavior of the presence of independent material objects. It is, as Wittgenstein would have said, the acquisition of a "form of life," and one so basic that without it the child cannot become fully human. The entire process is one of biological adaptation and social initiation into the culture of the surrounding adults.

The word "knowledge" seems too weak to express the convictions that result: that is why I have preferred to use the term "orientation." Of course, much more than "respect for facts" is

involved, and the objective orientation is laced with all sorts of magical and animistic beliefs. A child, like an adult, is a poet, a maker of fables, as well as an adapting organism. All that we call subjective is at least as important to the child as the development of outward vision. Yet children and the adults that they become do develop a serviceable view of the world, a sufficiently comprehensive map of their environment, the properties of material objects, their relative spatial and temporal positions, and their causal interactions. This common-sense view of the world is, in my judgment, the indispensable matrix out of which scientific objectivity develops. We learn, as scientists or consumers of science, to extend and refine it, but we can reject it only by actual withdrawal from the world. The price of abandoning the common-sense view is the rejection of life that we observe in autistic children or catatonic adults.

To be sure, the common-sense view of the world is in many ways an incomplete and distorted view of reality. All of us see the universe in a curious perspective, with our own bodies and our own temporal locations as having privileged status. Large areas are, as it were, vaguely sketched in, and the cognitive field is full of invisible blind spots.

Speaking for myself, the interior of my body is largely terra incognita. Nor is anybody's sense of reality an all-or-nothing affair. Our awareness of the independent existence of other persons and other things flickers and wavers: the Chinese seem less real than our neighbors and the distant past hardly more than a theoretical inference from the narratives of historians. One should think perhaps of different modalities of awareness, with absorption in the external at one extreme and absorption in the inner at the other. The orientation that I have been reminding you of is merely one way of coping with experience, one way of coming to terms with what Freud called the "exigencies of life." We are concerned with one kind of *attitude* toward experience, too basic to be called a theory or even a world view. Here what corresponds to doubt at the intellectual level is not a series of propositions for which we seek evidence, but rather a tendency toward withdrawal from external realities. A human being dom-

inated by subjective impressions would have to be uninterested in observation and the reception of information from the external world. The closest approximation that we have to this condition is that of drug-induced or alcoholic trances or, better still, our condition in sleep. "In natural sleep," as Freud said, "we withdraw our interest from the whole external world."[16]

Scientific attitudes and the scientific attempt to construct a single comprehensive and objective view of the universe are necessarily grounded in this rich and diversified compost of respect for facts, rudimentary theories about the external world, and partially adapted actions and habits of response.

Imperfect as it is, it is all that we have with which to get a grip on reality. Even the pilot of a Boeing 747 must use his hands at some point. Even the most sophisticated scientist, controlling a battery of theoretical concepts remote from immediate experience, has also to rely on the primitive accommodations to the immediate environment that I have called the common-sense view of the world. His capacity to handle the middle-sized objects that are part of his working equipment depend absolutely on his arduous indoctrination in how to cope with the elements of the external world that impinge most directly upon him. Without the skills of common sense a scientist would be no better than an autistic child or a zombie.

Science and Common Sense

The great Thomas H. Huxley went much further than this. In a famous statement he said:

> Science is, I believe, nothing but trained and organized common sense, differing from the latter only as a veteran may differ from a raw recruit: and its methods differ from those of common sense only so far as the guardsman's cut and thrust differ from the manner in which a savage wields his club. The primary power is the same in each case, and perhaps the untutored savage

> has the more brawny arm of the two. The real advantage lies in the point and polish of the swordsman's weapon; in the trained eye quick to spy out the weakness of the adversary; in the ready hand prompt to follow it on the instant. But, after all, the sword exercise is only the hewing and poking of the clubman developed and perfected.[17]

Huxley's position has some plausibility in its application to the rather primitive state of biology in his own day. He could still say that "the vast results obtained by Science are won by no mystical faculties, by no mental processes, other than those which are practiced by every one of us, in the humblest and meanest affairs of life." If one thinks, as Huxley did, that the intellectual procedures of science consist of no more than "induction and deduction," then no doubt one might be prepared to say, with him, that "the man of science, in fact, simply uses with scrupulous exactness the methods which we all, habitually and at every moment, use carelessly; and the man of business must as much avail himself of the scientific method—must be as truly a man of science—as the veriest bookworm of us."[18]

Today such assimilation of scientific method to "trained and organized common sense" looks wildly exaggerated. It sounds about as plausible as saying that driving an automobile is nothing more than a trained and organized way of walking. To be sure, both modes of locomotion are ways of getting from one place to the other—but after all, the presence of an automobile, with all that it implies of background technology and developed skills of using that technology, makes a fantastic difference.

The extraordinary development since Huxley's time of mathematical models, of abstract scientific theories involving concepts remote from immediate experience, and, above all, the growth of a remarkable scientific community, which manages to combine rigorous public criticism with the continuity of a shared tradition, puts science in a light that surely would have surprised Huxley and his contemporaries. It is the fantastic elaboration of the theoretical, instrumental, and technological apparatus by which scientific investigation is now conducted that provokes the radical skepticism of the critics of science.

Doubts about Science

An adequate examination of the barrage of critical objections that are now being leveled, from a variety of quarters, against the supposed objectivity of science would require another essay at least as long as this one. Those scientists and philosophers of science who have tried most strenuously to form an adequate working conception of the nature of science *as it is* are obviously impressed by the indirectness of scientific approaches to reality. Such knowledge as we now claim to get from science seems to be increasingly filtered through a complex of human language, concepts, and theories, so that it becomes increasingly hard to separate the objective features from the subjective ones. But such heightened and justified sensitivity to the indirectness of scientific knowledge sometimes leads to absurdity.

A scientist, Hudson Hoagland, once wrote:

> All knowledge of our universe, including a knowledge of God, comes to us strained, if you will, through a series of highly involved physio-chemical events constituting sensory reception and responses mediated by the central nervous system. We experience not the properties of objects but the properties of our own nervous systems. We can thus have no direct knowledge of reality beyond the symbols that we learn to agree upon with others who have similar nervous systems.[19]

But how can Hoagland know all that he presupposes about the "highly involved physio-chemical events" to which he refers? If he knows this, then he knows that observation and experiment do teach us something about the external world, and not merely about that special part of it that we call our nervous system.

A similar argument *ad hominem* can be applied against the overenthusiastic followers of Thomas Kuhn, who draw, from their supposed *knowledge* of the sociological determinants of scientific progress, skeptical conclusions about the very possibility of scientific knowledge. It would be consistent for such radical critics of the scientific enterprise to end by criticizing their own presuppositions of knowledge and to relapse into silence. But there is

something charmingly absurd about a skeptic who, while disclaiming the possibility of knowledge by the means that we have found to be the most reliable, still affirms confidently that he *knows* that scientists cannot know anything.

It was said that the great Duke of Wellington, at the height of his renown, was accosted in Hyde Park by a stranger who raised his hat and said, "Mr. Winterbottom, I believe." To which the duke replied, "Sir, a man who can believe that could believe anything!" I am inclined to say that a man who can disbelieve the fruits of scientific investigation could and should disbelieve anything. There is no halfway house between skeptical rejection of the cognitive claims of science and a total skepticism that should reduce its advocates to silence.

Envoi

Every philosopher needs a devil's advocate to keep him reasonably honest. I imagine my own as that splendid kind of student who is not about to be hoodwinked by fine words or plausible arguments. I imagine him saying: "This is all fine and dandy. You come across as a good old-fashioned defender of the eternal verities. [A shrewd hit, that!] Science, you say must believe, now as always, in an objectively given universe and must be committed to the pursuit of truth—not probability, or common agreement, or something that "works," or plausible conjectures. Only Grade A, unadulterated *truth* is going to slake your philosophical thirst for what you call objectivity. But you are astonishingly reticent about how to *find* this precious knowledge. And you pass over in silence the disagreements of scientists themselves and their own repeated testimony to the so-called creativity of the scientific enterprise."

An adequate reply is beyond my power. If I had the time, however, I would stress two points. First, disagreement in the *practice* of science can easily be exaggerated. Second, and more important, the ideal of objectivity is, like other basic ideals, one that we ought to pursue even while knowing that we cannot wholly achieve it.

4/

Is Scientific Neutrality a Myth?

The Implications of Neutrality

Let us begin with the notion of scientific neutrality and ask what it really means. Although the term is widely used, its precise meaning is unclear and its suggestions can easily be misleading. The dictionary defines neutrality as a policy of withholding aid from any states at war with one another or, more generally, as the attitude of taking no side in some dispute or conflict.

To call scientists neutral is to imply that they are noncombatants. But in what war, in what dispute, are scientists supposed to take no sides and to remain indifferent? (Incidentally, throughout what follows, I shall be considering the obligations of scient*ists*. I shall use "science," for the most part, merely as convenient shorthand.)

The kind of conflict in which science is supposed to stand aside can be conceived, for a start, as a conflict between practical policies. Consider, for example, the question of whether or not to smoke cigarettes. The prevailing view is that science can only provide factual information relevant to this issue and that it has nothing to say about the values that are involved. Such factual information can be ammunition for either side, because it may strengthen the case for smoking or for nonsmoking. But it would be wrong, many people think, for a scientist as a *scientist* to tell an individual smoker or the society to which the smoker

belongs which choice to make. Whether we choose to take the risk of lung cancer or to ignore that risk for the sake of the benefits or pleasures we impute to smoking, is up to you and to me. According to this view, science should have, moreover, nothing to say about whether lung cancer is an evil; that crucial judgment is also up to "the public": the consumer.

Thus expert scientists are supposed to say to their fellow citizens: "Whether you take lung cancer to be good or bad is your own affair. And anything you propose to do about this is also your own affair. As scientists, we take no position—have nothing to say about such value judgments or the appropriate policy decisions. We stand neutral." From this standpoint, a similar posture of neutrality is in order—is, indeed, required—in more controversial issues; abortion, for instance, or medical experiments on uninformed and unconsenting subjects.

The conflict in which scientists are enjoined to remain neutral may also involve a clash between moral positions rather than between practical policies. Consider, for example, the question of whether a nation is justified in committing genocide. According to the prevailing view, the scientist, as scientist, has nothing to say about this matter. If a nation wishes to engage in genocide, science can of course provide the tools—the most efficient means for achieving that unchallenged end. But should a nation reject genocide as a monstrous crime, that too is its right and privilege. As between favoring genocide and opposing genocide, science is neutral: it has, one might say, no moral dimensions. The doctors who experimented on the wretched inmates of Nazi Germany's concentration camps ought not to be reproached *as* scientists.

This conception of the scientific expert as one who stands aside from the clashes of policies and moralities—as, say, Switzerland stood aside from the war between Germany and the Western powers—is misleading in one important respect. A neutral is one who chooses not to take sides, and must therefore necessarily have the power to be a partisan if he wishes. But if scientific neutrality is a reality, not a myth, scientists have no option about taking sides with respect to the agonizing personal and social problems of our time. If so, talk about the neutrality of science smacks somewhat of the ridiculous—like calling bread neutral

because it can feed sinners or saints indiscriminately. There can be no *scientific* choice or partisanship in matters of policy or morality, and hence there can, strictly speaking, be no scientific neutrality either: science just has no direct bearing on any except scientific issues. (The qualification is, of course, important.) The neutrality of science, one might say, has to be essential, incapable of removal or amelioration. But perhaps it would, according to this view, be better to speak of scientific irrelevance. Nevertheless, I shall continue to use the convenient, if somewhat misleading, label of "neutrality."

The thesis of the essential neutrality of science might usefully be broken down into two subviews. If science really can say nothing about what should be done or should not be done, it might be called *normatively neutral.* If it can say nothing about what is good or bad, right or wrong, it might be called *evaluatively neutral.*

Whether in its normative or evaluative version, the neutrality thesis depends on the answers to some troublesome questions of logic. If science really is normatively and evaluatively neutral, that is something we shall just have to accept. To be morally indignant about this would be as absurd as to complain that 2 and 2 make 4. We have to deal in the first instance with a question of logic, not one of morality or policy.

The Logical Defense of the Neutrality Thesis

Let us now consider the reasons for thinking the neutrality thesis to be correct. So far as I know, the only reason in its favor—but a powerful one—is a supposed proof of the existence of a logical gap between "is" and "ought" or between "is" and "should." More specifically, there are supposed to be conclusive reasons for recognizing that no proposition containing such words as "should," "should not," "must," "must not"—that is to say, normative words—can follow from the *factual* propositions that science seeks to certify as true or probable.

The classical form of this argument is due to David Hume, and

it continues to be a part of the current orthodoxy of nearly all professional philosophers. Hume's argument is straightforward and deceptively simple. In a famous passage from his *Treatise,*[1] which has been quoted a thousand times, he maintains—to use his own language—that all the moralists he knows make an imperceptible transition from observations about human affairs (or assertions about the existence of God), all expressed with "the usual copulation of propositions, *is* and *is not,*" to normative conclusions "connected with an *ought* or *ought not.*" He says that this transition is "of the last consequence" and needs to be explained. In his own words: "It is necessary . . . that a reason should be given for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it." The nerve of Hume's argument, therefore, is the plausible idea that the conclusion of a valid deductive argument cannot contain any material that is not already contained in the premises. If the premises are factual, and so contain no reference to ought or ought not (or for that matter to good or bad), it does seem "inconceivable," as Hume says, that normative or evaluative conclusions should follow.

Of course there is one way in which new material of a normative or evaluation sort can be present in a valid conclusion. From the premise that omelets cannot be made without breaking eggs (which can be regarded as factual), it certainly follows that *if* omelets ought to be made, *then* eggs ought to be broken. What Hume means is that no *categorical* or *unconditional* "ought" conclusions can follow from factual premises. This argument certainly seems, at least at first sight, extremely persuasive. But the matter is not so clear. How do we know that a given proposition is factual, nonnormative, in the relevant sense? The absence of the crucial word "ought" or "should" is not a reliable criterion. For example, the proposition that murder is a sin does logically imply the conclusion that one ought not to commit murder. But how are we to tell from the mere linguistic form of the first proposition that it is nonnormative? Well, one might fall back on the idea that the proposition about murder is *unverifiable.* But that raises other and controversial questions that are no part of Hume's formal argument. His famous argument needs to be

supplemented by further considerations about what propositions shall count as being in some appropriate sense factual or objective—that is to say, true or false independently of the desires or hopes of the persons who assert and believe them. Behind *this* issue there lurk difficult questions about how we should conceive of the aims and procedures of scientific investigation. Should we conceive of science as a very special, distinctive way of approximating to knowledge, or as a discipline that shares the goals of the whole spectrum of knowledge-seeking activities?

My own considered judgment is that Hume's argument, in the form in which I have presented it, is circular. I believe that certain categorical "should" and "ought" propositions do have truth values, and can be certified as true objectively (independently of desires or hopes). If so, Hume was wrong in thinking them altogether different from scientific propositions. In a properly comprehensive sense of knowledge—which is broader than that of scientific knowledge, narrowly conceived—some normative and evaluative propositions can be known to be true. But should that be so, Hume's argument collapses and better arguments must be found for the logical segregation of scientific truths from practical and moral issues.

These questions are too intricate and controversial for detailed discussion here. I mention them merely to counteract the widespread opinion that somehow the independence of "is" from "ought" has been established once and for all by logicians and philosophers. To discredit this weak argument from philosophical authority, it is worth reporting that a sizable number of contemporary philosophers have rejected Hume's argument.

But in a way it does not matter, for our present purpose, whether Hume's critics are right or wrong. I think we are entitled to assume that human beings, however diverse their moral or religious backgrounds, can agree on certain fundamental ethical principles. We must postulate this in order to make any kind of rational dispute about means or ends possible. If we were to encounter somebody who thought it all right to eat young children as a way of augmenting our food supply, argument would be pointless and therapy or something more drastic would make

more sense: anybody who genuinely held such a position would have to be treated as monstrously inhuman or insane. So even if factual propositions had to be regarded as logically segregated from normative or evaluative ones, the introduction of some generally acceptable nonfactual premises would legitimate the derivation of normative conclusions.

More important, however, I shall later argue that any approach to our problems that confines attention to the products of scientific activity, considered as isolated *propositions,* and conceives of knowledge as merely the justified *assertion* of such propositions is misleadingly and inappropriately one-sided. The weightiest, most urgent problems of neutrality are unaffected by the supposed logical gap between scientific truths and norms or values. If we think of science more concretely and realistically, as something that human beings do, the issues (as I shall hope to show) look very different.

On the whole, the thesis of the gap between "is" and "ought," although it has had enormous appeal and very widespread currency, deserves in my opinion to be called one of the great half-truths or popular fallacies of contemporary Western culture. We should especially be wary of the seductive but invalid syllogism that runs as follows: Science is a good thing; science is necessarily neutral; therefore, scientific neutrality is a good thing. This syllogism might easily be converted, by negation of the conclusion, into: Scientific neutrality is a bad thing; science is necessarily neutral; therefore, science is a bad thing. Qualms about the so-called neutrality of science jeopardize the supporting public attitudes without which science cannot thrive.

Nevertheless, the logical version of the neutrality thesis deserves to be taken seriously, if only for its practical consequences. If neutrality, in its logical sense, has to be accepted, if science really has no intrinsic tendency to further human welfare rather than to increase human misery, the case for massive and indiscriminate support of science may well appear problematic. Even if the human usefulness of science can be grounded in its *indirect* bearing on welfare and happiness, it is natural for a firm believer in the neutrality thesis to have serious qualms about the value of science. A layman might well ask: If scientific truths are only

bridges between nonrational directives and evaluations, what is so good about science? If a scientist is really, from a logical point of view, a kind of engineer, supplying factual bridges between the normative and evaluative considerations that matter most to human beings, what is so good about being a scientist? (The roots of one deplorable kind of antiscientific revolt may be found in this crude kind of attitude.) Questions about whether science as a whole or particular scientific projects are worth pursuing are themselves not scientific questions, but they cannot be brushed aside. For science, considered as an activity, is not an uncontrollable natural phenomena, like the earth's rotation. As a system of voluntary activities it is, at least in principle, amenable to internal or external control.

Neutrality and the Subjectivity of Norms and Values

What I shall call the neutrality thesis becomes particularly pernicious when it is coupled with a correlative view about the alleged subjectivity of normative and evaluative judgments—and especially so when the realm of rationality is exclusively identified with that of scientific discourse.

Let me give two examples. The well-known biologist Jacques Monod has written:

> Science rests upon a strictly *objective* approach to the analysis and interpretation of the universe, including Man himself and human societies. Science ignores, and must ignore, value judgments. Knowledge (yet) discloses and inevitably suggests new possibilities of action. But to *decide* upon a course of action is to step out of the realm of objectivity into that of values which, by essence, are non-objective and therefore cannot be derived from objective knowledge. There is strictly no way of *objectively* proving that it is BAD to make war, or kill a man, or to rob him, or to sleep with one's own mother.[2]

On the same page Monod says, "Objective knowledge defines the alternatives. By itself, it does not in any way help in solving

the agonizing ethical problems which science only poses." Later on he says, "Science indeed cannot create, derive or propose values," and he speaks about the necessity for the "*choice*" of value systems. Even the decision to become a scientist, says Monod, results from a "deliberate axiomatic choice of a standard of value."[3]

In these remarks, which could be paralleled by any number of similar ones by distinguished scientists and intellectuals, objectivity connotes what is given, regardless of human wishes and valuations, while values are treated as a matter of mere arbitrary choice. So the questionable—but unquestioned—dogma of the gulf between the factual and the normative gets coupled to an equally dubious dogma of the arbitrariness and subjectivity of all norms and values.

My second example is taken from an article published some years ago in *The New York Times Magazine*.[4] Robert L. Heilbroner, a professor of economics at the New School for Social Research, therein considers the question, once raised by Adam Smith, of whether "a man of humanity, given the choice, would prefer the extinction of a million Chinese in order to save his little finger." (Why Chinese, one wonders? Would it have made any difference if the lives of a million Scots had been in question?) One might have supposed the answer to Adam Smith's question was obvious. But listen to Professor Heilbroner's comment. He does indeed say that it would be "outrageous" to attach such fantastic value to one's finger (or, to turn to his more immediate subject, outrageous to sacrifice posterity for the benefit of those alive today). But he thinks that "there is no rational answer" to such terrible questions. And this, according to Heilbroner,

> reveals the limitations—worse, the suicidal dangers—of what we call "rational argument" when we confront questions that can only be decided by *an appeal to an entirely different faculty from that of cool reason*. More than that, I suspect that if there is cause to fear for man's survival it is because the calculus of logic and reason will be applied to problems where they have as little validity, even as little bearing, as the calculus of feeling or senti-

ment applied to the solution of a problem in Euclidean geometry.

Here we have a kind of muddled fantasia on a theme from Hume: "cool reason" is contrasted with "the unbearable anguish we feel if we imagine ourselves as the executioners of mankind." Well, if Heilbronner is prepared to rely on the "unbearable anguish" of responsibility for mortal injury to human beings, he cannot be an attentive reader of the distinguished newspaper in which his article appeared. Can anybody seriously think there is *no good reason* for preferring the loss of one's little finger to the murder of even a *single* human being? Or that people should be regarded as perfectly rational if they were to adhere to such an insane scale of values? One can only marvel at the advocacy of so peculiar a notion of rationality. If prevalent conceptions of rationality confine reason to the exclusive consideration of factual and technical propositions, so much the worse for such restrictive conceptions. Scientific activity, regarded as rational in *this* way, runs the danger of speedily becoming inhuman.

Science as Activity

So far I have been discussing, in a very abstract way, science in its relations to values and norms, as if scientific propositions were simply given like natural objects, and as if the only proper questions arising were the logical ones of the relations between scientific propositions and normative or evaluative ones. It is time to remind ourselves of how much is ignored in this abstract approach and to consider science as *a system of human activities*—a far more than merely contingent or incidental contrivance for the delivery of eternal truth. Accordingly we must raise the question of neutrality again, in connection now with the enormously complex system of activities that deliver what (for the time being, at least) counts as knowledge.

What are we to say about the neutrality of science considered as a complex of actions? Well, the most obvious point is that

every human action is in principle subject to practical and moral evaluation. It always makes sense to ask whether what human beings are doing is prudent or wise, satisfactory, intelligent, right or wrong. All action is directed, correctly or mistakenly, toward ends held by the actors to be right. Accordingly, it always makes sense to ask whether these ends are properly evaluated or whether there is reason to suppose that their choice has been misguided or evil. Scientists, like all other human beings, in their active capacity, are necessarily subject to such normative and evaluative appraisals.

We are therefore entitled—indeed, obliged—to ask such questions as: What are the goals of science? Can they be approved without qualification or reservation? (Are they supreme?) What are the costs of supporting science, or some particular scientific project, rather than of doing something else ("opportunity costs," as economists say)?

These questions could be brushed aside only if the supreme goals of science were so *un*important, humanly speaking, as not to deserve scrutiny. Similar considerations apply to the means used by scientists: issues of moral evaluation could be brushed aside if the consequences of scientific activity were so trivial as to be negligible, or so obviously good as to make criticism pointless. In short, the only way to ignore moral questions about the values of science would be to hold that the goals and procedures of science are either inconsequential or uncontroversially good.

My original question about the neutrality of science, now that we focus on the manifestation of science as human activity, becomes, first, "Is science negligible or harmless in its consequences?" and second, "Are the ends of scientific activity unequivocally and unreservedly good?" We are led therefore, inescapably, to consider the practical importance of science and what might be called its evaluative dignity.

I shall not linger over questions concerning the practical momentousness of scientific activity, since that is too obvious to need discussion. But it is worth raising a question about the ultimate goals of scientific knowledge because, as I shall argue, there is a tendency illicitly to transfer certain important and genuine values, implicit in the disinterested pursuit of knowledge for

its own sake, to the whole of scientific activity, in all its amplitude and variety.

Science as Disinterested Pursuit of Knowledge

Science, as its etymology and that of its German equivalent *Wissenschaft,* indicate, does have some essential connection with knowledge. Considered as an activity, it needs to be regarded as the *pursuit* of knowledge, and it has of course been so defined on innumerable occasions. It is perhaps a pity that, instead of the relatively new word "scientist," we do not speak, in the old style, of a scientific investigator as a "natural philosopher," engaged in the pursuit of knowledge and insight.

If we are to consider whether the pursuit and love of knowledge is a supreme and unqualified good, we shall have to make an obvious distinction between knowledge that issues in the form of information and knowledge that yields understanding—between the collection of items of fact and the search for explanatory principles (at all levels of depth, from crude empirical generalizations to the most basic theories of physics).

It might seem obvious that knowledge in the form of mere information is not unqualifiedly a good thing. There are probably more "bits" of information in a new telephone directory than in the whole of Von Frisch's beautiful work on bees; but it would be perverse to prefer the former to the latter, or to advocate "scientific" research on the contents of telephone directories. If the gathering of information were an unqualified good, we would have to salute the perpetrators of the Watergate break-in and the compilers of the FBI files for behaving like true scientists.

Apart from such obvious points, there is the consideration that the effective and successful handling of useful knowledge demands the constant suppression and elimination of *irrelevant* information. To see anything well is necessarily to ignore what you don't want to notice: if you look at individual trees, you will be unable to see the woods, and if you look at the woods, you will necessarily ignore details about the trees. The collection and

systematic accumulation of knowledge is inescapably selective. Because it involves human choices, it inescapably commits the observer to making value judgments. Therefore, even in its most prosaic function of fact collecting, scientific activity is bound up with judgments of value. There is no particular virtue in collecting trivial information: the pebbles on a million beaches are waiting to be counted, but one would have to be a fool (or an advanced thinker) to identify that kind of pointless fact gathering with the exercise of significant scientific activity.

When we turn to the other supreme product of scientific activity, *understanding,* we might suppose that the implicated values are obvious. Isn't it always better to understand than to be blinded by illusion or prejudice? Well, the same elementary points apply: it all depends on the objects of understanding; it all depends on their value and the uses, if any, that the understanding is to serve. Is it to be seriously argued that it is a good thing to understand the most painful ways of killing and torturing innocent people, just because it is knowledge? Would it not be better to be and to remain ignorant about such matters?

The chief point that I am making is that the worth and undeniable dignity of science—viewed as the disinterested pursuit of knowledge, whether of facts or of explanatory principles—depends on value choices; choices of significance and relevance; choices of importance and fruitfulness; and, above all, choices concerned with *ultimate worth to human beings.* Paul Goodman once said, in a striking phrase, that "technology is a branch of moral philosophy, not of science."[5] I am not sure I agree; moral philosophy is difficult enough without incorporating engineering. But one can see what he meant. I suggest that the same insight also applies to science, even in its purest manifestation. Considered in the round, the "purest," the most abstract, and the apparently "useless" branches of science are inescapably value-dependent and value-laden.

I do not wish to be misunderstood as an enemy of basic science or of the pursuit of science for its own sake. There are plainly great values in disinterested science, properly enlisting the devotion and sacrifice of some of the most gifted members of our species. The achievements of Galileo and Newton, of Darwin

and Einstein, to name no others, add to the stature of all of us. And the disinterested pursuit of truth about the physical universe, about the biological realm, and about political and social institutions is of great value, strenuously won, constantly under attack, and to be secured only by constant vigilance.

I share the view that Julien Benda expressed in his famous book, *La trahison des clercs* (translated as *The Great Betrayal*), that we shall always urgently need those who will try to pursue and promulgate the truth as they see it—not only scientists, but scholars and artists and poets as well—without being influenced by factional and national interests or by vanity, envy, and greed.

The autonomy and the freedom from all interference or control that are part of the ethos of the disinterested truth seeker can be granted to an individual only on one condition: that he or she be genuinely interested in knowledge and understanding for its own sake. And that presupposes that contemplative knowledge of the universe be indeed a common good, accessible to all, shared and enjoyed by all, and not something deleterious to human welfare.

The great dilemma posed by the extraordinary and unforeseeable efflorescence of science and technology has been the increasing impossibility of insulating such an elevated conception of science from its practical consequences. Now that science has become a terrible power for great evil as well as for great good, assumptions about the neutrality of science urgently need reconsideration.

Science Today

What might be called the classical ideal of the scientist as a disinterested pursuer and lover of truth can be wholeheartedly followed by solitary scientists or solitary scholars—Darwin in his study or Gilbert White in the countryside around Selborne, or even Einstein in the Patent Office at Berne—calmly discovering and contemplating truth, with no ulterior motive, for the sake of the radiant pleasure of the pursuit itself, isolated from the mar-

ketplace and the war room, sharing their discoveries only with a small group of like-minded devotees. But this idyllic picture, which can hardly have been a reality except occasionally, has long been overtaken by the Baconian ideal of science for the sake of "dominion over nature" and "the effecting of all things possible."

We have come a long way from the classical ideal of the independent scholar communicating only with a small group of like-minded researchers. The latest official figures that I have been able to obtain show that in the United States alone there were in 1978 over a million persons identified as practicing natural scientists, and about 190,000 social scientists—altogether some 1,223,000 to the nearest thousand. If we lump together engineers and natural scientists engaged in research and development, the total rises to something like 2,508,000. For the cost of this extraordinary effort we have an estimate in 1981 of $70 billion. (The Department of Defense itself spent, in that year, nearly $14 billion on research and development.)[6]

Considering now the intellectual product alone, we find such statistics as these: that the Chemistry Abstracts for 1981 had 450,587 entries. (I estimate that anybody who undertook to read, even superficially, the volume of work produced, say in some branch of physics, in a given year would have to spend anything from three to six months doing nothing else. Of course, active scientists usually ignore this glut, keeping up to date instead by informal conversation and preprints.)

Society cannot be expected to spend such enormous sums, and to promote such intensive training and employment of high talent, in the services of a noble ideal of the disinterested pursuit of truth. Indeed, considering the very high proportion of financial support for science that comes from industry and from the defense establishment, one would expect the effective goals of this vast enterprise to be at least strongly compatible with the ruling goals of industry and the military; that is to say, providing ways of increasing profit and destroying human beings. This intimate coupling of scientific activity with the industrial-military complex has cast a shadow over the noblest aspect of Bacon's vision: the use of science to promote human welfare and happiness.

Fortunately, welfare and happiness can often be profitable, so a more sanguine view of the situation might emphasize the undoubted and enormous contributions of applied science and technology to the control of disease, improvement in communication, and the pleasure and comfort of, say, 20 percent of all human beings alive today.

The Responsibility of Science

The plain lesson of the fantastic explosion of science in modern times is that the separation of pure science from its applications is no longer feasible within a context of moral evaluation. The "purest" and most recondite discoveries of disinterested scientists can now, in a relatively short time, be harnessed to the production of new ways of transforming individual lives and of reshaping society itself.

This means that the ground has been cut away from the classical claim for the autonomy of pure science. So long as the scientist could be regarded as a harmless discoverer and contemplator of aesthetically satisfying truth, the rest of us could admire and applaud (even when we could not understand) and could consider humanity dignified by such devotion—as it is by any spectacular realization of human excellence. But the case is altered when the consequences of such admirable activity can have such momentous consequences that the bare survival of human beings becomes problematic. Science is even more terrible in its potentiality for evil than are the nuclear bombs for which it is responsible. Over every scientific laboratory should be pasted the warning: "Danger! Scientist at work."

Whether one takes an optimistic or pessimistic view of the probable consequences of the massive transformation of the natural and manmade environment, and the conditions of social life that are traceable in the end to the innocent classical delight in knowledge, the conclusion is inescapable: the claim of moral innocence must now be dismissed as a mere vestigial survival of an earlier and more benign era. It is one thing for a pure geometer

to insist on being left alone, without impertinent demands to prove that the research has practical application. It is quite another thing to hear the same plea of moral blamelessness and neutrality, in the sense now of moral innocuousness, made by contemporary scientists working, say, for the military establishment. For when the immediate consequences of scientific activity are so plain and so obvious, a plea of scientific neutrality can be properly characterized only as an expression of deliberate myopia or, to put it bluntly, moral irresponsibility.

An Example of Moral Irresponsibility

The *New York Times* for December 27, 1967 (p. 8), reported an interview with Louis Frederick Fieser, who during World War II was in charge of a team of Harvard University scientists developing napalm. Dr. Fieser is reported as having said that he felt free of "any guilt." He is also quoted as remarking, "You don't know what's coming. That wasn't my business. That is for other people. I was working on a technical problem that was considered pressing." He went on to declare: "I distinguish between developing a munition of some kind and using it. You can't blame the outfit that put out the rifle that killed the President. I'd do it again, if called upon, in defense of the country." When asked about the use of napalm in Vietnam, Dr. Fieser said: "I don't know enough about the situation in Vietnam. It's not my business to deal with political or moral questions. That is a very involved thing. Just because I played a role in the technological development of napalm doesn't mean I'm any more qualified to comment on the moral aspects of it."

The moral callousness of these remarks is matched by their confusion of thought. The idea, for example, that one needs moral "qualifications" in order to consider the consequences of one's action is very odd. But it is unfair to saddle Dr. Fieser alone with this attitude, which continues to be widely held. The social scientist George A. Lundberg said it plainly nearly forty years earlier:

> It is not the business of a chemist who invents a high explosive to be influenced in his task by considerations as to whether his product will be used to blow up cathedrals or to build tunnels through the mountains. Nor is it the business of the social scientist in arriving at laws of group behavior to permit himself to be influenced by considerations of how his conclusions will coincide with existing notions, or what the effect of his findings on the social order will be.[7]

Against such misplaced high-mindedness, we ought to say that in the case of research plainly intended to produce flame throwers and incendiary bombs, the moral responsibility of the chief inventors and artisans cannot be evaded by protestations of innocence. (That the persons most directly involved can have a good conscience only adds to the irony—one might even say the horror—of the situation.) Responsibility for the invention and use of such "scientific devices" as napalm is of course shared by many—and belongs, no doubt, primarily to those who have the final power to commission and to use such inhuman means of destruction. But the inhumanity of statesmen does not absolve technologists, scientists, and all of us from complicity. Moreover, given the key role played by scientists at all levels of involvement, it is particularly hard to absolve them from fairly direct responsibility. My purpose is not to preach. Nor is my purpose here to suggest practical measures by which to discharge the undoubted responsibility that is borne by science and every society that supports it. My purpose has been the logically prior one of arguing that the responsibility exists and will be discharged or ignored one way or another. Some help can be expected from the determined effort of scientists and laymen who are not content to transfer the ideals of an earlier age to the present situation or to regard scientists as mere hired help fit for any purpose, good or evil.

An optimist can hope for a time when science will have returned to its ancient and nobler functions of enlarging human understanding and relieving human misery. But as Groucho Marx must surely have said on some occasion, "It takes an optimist to be an optimist nowadays."

Envoi

I may reasonably be expected, in conclusion, to answer my original question. *Is* scientific neutrality a myth? I must answer, I am afraid, in the irritating style of a philosopher: "It all depends on how you conceive of science" (and, of course, on how you conceive of neutrality).

If you are satisfied to abstract from the extraordinarily ramified complex of actions, institutions, and traditions that we comprehensively embrace as "science"—together with the supporting educational, economic, industrial, and political structures—that aspect of it that consists of the disinterested pursuit of knowledge for its own sake, scientific neutrality is *not* a myth, but a justified *claim*.

But if you think such emphasis on the disinterested investigator is too partial and abstract a view of the whole of science to be serviceable, if you think of science in its totality of present-day activities, you will need to reject the neutrality thesis as a myth, and a pernicious one; for a particularly dangerous kind of myth is one that, under the guise of valuing truth, insinuates a claim to moral irresponsibility.

Part III

Beyond Rationality

5/

Humaneness

Preamble

There is a well-founded belief that a humanist, unlike such a specialist as a mathematician or an economist, should help ordinary people to become truly or admirably human. Indeed, the history of the subjects we now call humanities shows that they have long been expected to have this kind of application. On the other hand, the task of delineating what ordinary people would recognize and accept as truly human approaches to interpersonal relations has proved to be surprisingly difficult.

Having become interested in the history of such words as "humane," "inhuman," and their near relatives, I was surprised to learn that the present-day connotations of these words are relatively new. Until the beginning of the eighteenth century, according to the *Oxford English Dictionary,* "humane" was simply an alternative spelling of the word "human," and was so pronounced. Two centuries ago, the noun "humanity" was commonly used in a sense that has since almost disappeared from English usage (though it is still attached to the German *Menschlichkeit*). Boswell said somewhere that when he and Johnson were guests in a country house they were received with great "humanity." I don't suppose he meant that ointment was applied to their bruises. In those days "humanity" had not yet been specialized to mean much the same as "humanitarianism," and I

think Boswell would have been surprised to hear "humane killer" used to refer to a device for slaughtering cattle. He could assume that his readers would take "humanity" to mean a certain positive attitude—roughly speaking, the attitude of treating other persons as befits human beings *qua* human.

So the question I am trying to answer takes the form: What could we reasonably and properly mean in describing someone as behaving humanely—I know of no better word—when interacting with other persons? Also, what philosophical assumptions would we need in order to recommend and justify humane attitudes? I have found, rather to my surprise, that while a great many philosophers—perhaps especially Kant, but not he alone—have been keenly interested in this question, none has produced a systematic discussion in the form of a monograph or book. Writers on ethics often supply considerations that are suggestive and helpful. But no major philosopher, so far as I know, has ever addressed himself to answering explicitly the questions: What would it be like to treat other persons in a truly humane fashion? How is such treatment justified?

The Distinctiveness of Human Beings

My formulation of these questions is intended to presuppose that there is something distinctive about human beings. If human beings are only a relatively interesting species of animal (as writers urge nowadays) so that a human being is really nothing but a "naked ape" or something of the sort, the point of my questions must largely disappear; the explication and justification of humane treatment for persons can have no more basic theoretical interest than the corresponding issues about appropriate treatment for cats and dogs. Or, if some of our contemporaries are right in regarding human beings as simply material systems, no different in principle from complex machines, the questions I have been raising must radically change their import.

I shall assume, as a matter of common sense, that in calling human beings *persons,* we are rightly ascribing to them impor-

tant properties that cannot, even in principle, apply to other animals or to inanimate material beings. It seems to me plainly absurd to suppose that digital computers feel pain, understand jokes, or protest the abrogation of their rights; and almost as absurd to suppose that a chimpanzee or a dolphin can feel stabs of conscience, prove a mathematical theorem, plan for the future of the species—or compose essays about the moral obligations of human beings.

When I say such suppositions are absurd, I mean of course not simply that they are empirically absurd, because the performances are unlikely to happen, but rather that they are conceptually excluded, by our knowledge that machines and animals are not persons. Of course, I do not deny the obvious point that all of us have bodies with physical properties in common with animals. Indeed, I like to think of a human being as a polybian (by analogy with "amphibian") subject to physical, chemical, and biological constraints—and no doubt also to psychological, economic, political, and sociological forces. But I believe there are features of human personality that are outside the purview of any of the natural or social sciences, and that there is something therefore conceptually—or, if you like, ontologically—special about human beings. If the distinctive status of human beings as *persons* is acknowledged (as I hope it will be), we are left with the task of providing a positive account of what we understand by "being a person." We have to decide what are the relevant and distinctively personal features that need to be considered in human relationships.

Befitting Conduct to Others

I shall now both broaden and narrow our topic by asking the following question: What is it to act toward another human being as befits a human being—and as befits the humanity of another person? In order to keep our discussion within manageable bounds, I propose to ignore, from now on, everything but face-to-face interaction, deliberately setting aside our respon-

sibilities to posterity—and, for that matter, to animals or to our environment.

A first answer to the question "What is it to interact humanely with another?" might take the form I have already suggested: to behave as befits the humanity of both persons. But this trial formula is not very helpful, and may, indeed, be based on a weak analogy. The word "befitting" suggests the familiar expression "conduct unbefitting an officer and a gentleman," and teachers have been known to be dismissed with contumely for "misfeasance, malfeasance, or nonfeasance"—conduct unsuitable or improperly performed *as a teacher*. Questions about what is "befitting" arise naturally in a context referring to some specialized role—that of a judge, a doctor, or a traffic policeman, say—associated with corresponding sets of required or forbidden acts that partially define that role. Each such role is partially constituted, one might say, by a normative principle having the following form: In order to conduct oneself properly as an *X*, one must do such-and-such. This may explain why certain gross departures from the relevant norms are viewed as perversions of essence: it would be absurdly euphemistic to call some medical practitioner who did not even try to cure his patients a poor doctor; we would be likely, indeed, to exclaim in outrage, "He's no doctor!" Similar remarks apply to other roles, such as that of being a father, which are partly natural and partly conventional (cf. the expressions "He's no father!" or "What an unnatural father!").

While I have no difficulty in understanding these cases, I find it difficult to relate them to human beings *as such*. According to Aristotle, a human being *qua* human being does have a function: ". . . the function of man is a certain form of life . . . the exercise of the soul's faculties and activities in association with the rational principle."[1] But Aristotle's only argument for this ascription of a universal human function seems to have been that it would be strange for the shipbuilder, the doctor, and every other specialized craftsman to have a function while the human as human did not. In the absence of a religious belief in a divine purpose behind our creation, however, there is no plausibility in the notion that human beings are teleological entities having

some identifiable function or goal. Specifically, I can see no force in the contention that a human being has a function analogous to that of a natural organ such as a heart. To accept this view we would have to conceive of individual persons as parts of some superindividual entity—the nation, the *Volk,* or the like—and such a conception tends to legitimize what I regard as typically *in*humane treatment of persons.

So let us return to our original question for a fresh look. What actions and attitudes vis-à-vis another person are required merely by virtue of shared humanity? What, for instance, would be my moral obligations toward another person if we found ourselves alone in a lifeboat, while sharing no ties of family, nationality, or religion? This might be called the moral issue of *humane interaction.*

Bracketing Religious Commitments

I would like to develop an acceptable view of humane interaction without invoking any religious or theological premises. Theologians and divines have had much to say about the moral obligations of human beings to one another. But one should be able, I suggest, to conceive of proper humane interaction without appealing to specific religious commitments. All theologies discriminate between believers and infidels (sometimes most inhumanly), while humaneness, as I would like to conceive of it, should bind religious and nonreligious alike.

If it could be shown that a purely secular justification of obligations to other human beings as such must fail, that negative result would be of the first importance. But I am sufficiently optimistic to believe that atheists, no less than religious believers, have good and sufficient reasons for acting humanely.

Much harm has been done by the slogan that figures prominently in *The Brothers Karamazov:* "If God is dead, everything is permitted." With apologies to Dr. Johnson, I call such complacent amoralism the last resort of a scoundrel. A principled atheist, if I am not mistaken, is required, like everybody else, to comply

with well-defined obligations grounded in shared humanity, and cannot evade such obligations by virtue of disbelief in God's existence. Against Dostoevsky's Ivan, we ought to say: "Some actions toward our fellows are absolutely forbidden, and others are unconditionally required."

Views about Human Nature

Our reaction to the problem of humane interaction will depend strongly on whether we accept a view of human nature that is on the whole optimistic rather than pessimistic. Nowadays, denigrators of human nature are in full cry. William Golding, for example, in an essay about his famous novel, *The Lord of the Flies,* says that "man produces evil as a bee produces honey." His novel is indeed a chillingly plausible depiction of the reality of childhood evil. Golding believed that "the condition of man was to be a morally diseased creation" and that human beings suffer "from the terrible disease of being human."[2]

A reader who shares that view will have a ready answer to my question about humane interaction. If human beings do suffer from the ineradicable "disease of being human," they are dangerous animals and should be treated as such. What are our moral obligations with respect to bacteria? To avoid them if we can, and if we can't to destroy them. What are our duties toward the diseased animals that we call our fellow humans? Is it, similarly, to avoid them if we can and to destroy them if we must? (Isn't this in fact how many people think about communists or capitalists, whites or blacks, Arabs or Jews?) But surely, while admitting the sickening reality of human evil, one doesn't need to be a sentimentalist to recognize and regard Golding's position—or Jonathan Swift's savage condemnations of human beings—as exaggerated. To believe in original sin is not necessarily to view human beings as wholly incapable of goodness. Human beings are not saints, but neither are they Yahoos.

In saying so, I am of course making a moral evaluation of human capacities. Indeed, an enterprise such as mine is inescapably committed to a moral evaluation of human nature. Any

defensible conception of humane interaction must be supported by some tenable view of the nature of human nature—and especially by some adequate analysis of human *needs*.

Suppose one were considering the question, sometimes raised today, about the moral obligations we might have toward plants. One's answer will clearly depend on one's views about the nature of plants. If it is true that plants cannot feel kindness and cannot be affected by it, they do not need it and we have no duty to treat them kindly. Similarly for human beings. Recognition of certain fundamental facts about human nature, notably those that concern basic human needs, is a prerequisite to any reasoned conception of the moral obligations subsumed under the rubric of "human interaction."

Questions about the nature of human nature have exercised theologians, philosophers, and scientists since at least Aristotle's time, and remain highly controversial. Today there are respected authorities who claim that there is no such thing as "human nature," and others who offer incompatible analyses of what human nature is. To judge by the thousands of pages that have been written on the subject, determining the nature of human nature is an extremely difficult enterprise.

I do not regard the problem as quite so difficult as the current literature would lead one to suppose it is, for each of us has one advantage that cannot arise in connection with parallel questions about plants or animals. As human beings ourselves, with some direct awareness of our own aspirations, needs, and limitations, we can use this knowledge and the testimony of our fellow creatures (with suitable precautions against rash generalizations) to formulate some truths about human nature. That, at any rate, is what I now propose to do in a sketchy and schematic way, thus engaging in what used to be called "philosophical anthropology."

Aspects of Human Nature

Well, what do we know about human nature and basic human needs? First, that persons are *sui generis,* a unique type of entity. I

agree with the spirit of the following remarks by G. K. Chesterton:

> Man is an exception, whatever else he is. If he is not the image of God, then he is a disease of the dust. If it is not true that a divine being fell, then we can only say that one of the animals went entirely off its head . . . . Man is always something worse or something better than an animal; and a mere argument from animal perfection never touches him at all.[3]

Shorn of its theology (an action that Chesterton would have furiously protested), what is said here is of profound importance. I reject as thoroughly misguided any account that starts off with: "Man is just an animal that has such-and-such" or "Man is merely a gene carrier" or "Man is only a dressed-up computer." The presence in such reductionist stories of the word "only" or "merely" betrays the theorist's own suppressed awareness of the outrageousness of what he is advocating.

Well, what is distinctively and uniquely characteristic of a human being? I would reply (to start with the obvious): not merely consciousness, which is no doubt also present in lower animals, but critical *self*-consciousness—the capacity to know what we are doing, thinking, or feeling and to reflect critically about what we have done and propose to do.

Among the familiar aspects of such reflective consciousness, as it might be called, is the capacity, intermittently exercised, to understand and misunderstand what we are doing and to *pass judgment* on ourselves and others. Such critical and evaluative self-consciousness allows us some freedom to mold ourselves in our own images—to help ourselves become what we wish to be.

These reflexively critical powers are only sporadically exercised: we reveal ourselves as all too human in behaving impulsively, blindly, heedlessly, uncritically—and we are still human when we fantasize or when we sleep and dream. Indeed, everything that I have said or shall be saying about human nature must be qualified in a similar way, by recognition of the presence of mutually opposed tendencies in human nature. We can make free choices? Yes, but we can also behave like mindless robots. Similarly for other desires and powers.

The point deserves further emphasis: one thing that is unique about human beings, I suggest, is precisely this *paradoxical* character of our condition, the constant co-presence in us of contraries and oppositions. We are preeminently paradoxical creatures. As Alexander Pope put it in Epistle II of his *Essay on Man:*

> Created half to rise and half to fall;
> Great lord of all things, yet a prey to all;
> Sole judge of truth, in endless error hurled;
> The glory, jest, and riddle of the world.

This general point about the "paradoxical" nature of human beings must be constantly borne in mind when we try to frame an acceptable paradigm of human interactions. Every statement of basic human need—whether for food, shelter, security, friendship, sex, work, or esteem—needs to be qualified by "But not too much!" Every human need is balanced or offset by an opposite need. That makes the proper determination of human relationships a delicate and intriguing problem.

The basic needs that I have mentioned presuppose another and still more fundamental need that many thinkers strangely neglect; I mean the constant and enduring need to continue to exist: the need for *bare survival.* Nowadays, writers who do recognize this need label it somewhat deprecatingly as "the biological imperative," regarding it as arising from our animal nature. That seems to me implausible. I should have thought that the "Life-Force," as Bernard Shaw liked to call it, has very little importance for an *individual,* whose biological function is merely to contribute marginally to the reshuffling of the gene pool of the species. But tell a human being that this is his prime function and—unless he happens to be a certain kind of tough-minded professional biologist—he will be profoundly offended, and rightly so.

Each of us has the conviction that our continued existence—simply *our own* individual existence, not as a member of the human species or as some variety of it or as a mother, doctor, poet, or anything else—is, at least *to ourselves,* supremely important. This conviction issues as a declaration of a *right to presence* in

the universe. Each of us wants to say, "I have a right to be here in the world, just like you—and nothing that you say can shake that conviction."

This is, of course, not a claim to bare existence, on any terms whatever, though human beings will go to extraordinary lengths—whether in concentration camps, on the battlefield, or in hospital wards for the terminally ill—to struggle even for that. Typically the demand is for survival under conditions that maintain the individual's integrity: survival on terms that will permit the continued possession of those properties and the exercise of those powers that one counts as particuarly one's own. Spinoza said, "Every thing, as far as in it lies, strives to persevere in its existence."[4]

Value "For Oneself"

Each human being wants to be recognized and valued as a unique individual—not merely as a physical object, an animal, a member of some collective, or a performer of some social function, not merely as an instance of some category.

To be valued for oneself; and, more generally, to be in a relationship in which one's identity counts, not simply one's possession of some property, capacity, function, or role that might belong to someone else: such a demand may well sound obscure and in need of philosophical clarification. Yet *particularizing relationships* are not hard to find. Somebody who waits in a queue to hear Rubinstein play wants to hear *him* and nobody else. Suppose Rubinstein were replaced in a concert by some substitute who looked just like him, had the same mannerisms, and could play in a fashion indistinguishable from that of the great pianist himself (a sufficiently unlikely hypothesis, to be sure): would not the hearer feel cheated on discovering the deception? Or consider the case of a women in love with one of a pair of truly "identical" or indistinguishable twins. If the one she loved were to die, would it be a consolation for her to know that his indiscernibly different brother was awaiting her devotion?

Admiring, loving, and many other relationships can be particularizing relationships, tied to the existence of some uniquely identified person. But we can easily find humbler examples: I may value a particular copy, still in mint condition, of Descartes's *Meditations,* because it was the first philosophical text I was able to buy; if I lost it I should not be consoled by knowing that the edition was still in print, and I could, accordingly, easily obtain an exact replica. In a particularizing relationship, the relevant entity, one might say, is indispensable—and in a doubly particularizing relationship, such as love or friendship, *both* persons are indispensable. Such relationships might be called *maximally personal.*

Desire and need for maximally personal relationships may well be important features of human nature, but we should not regard them as a necessary condition for human interaction. Normally, that implies a desire for intimacy that is balanced by an equally legitimate and natural desire for distance: I do not necessarily desire the surgeon who operates on me to be, as we say, personally involved with me—if only because that might interfere with his skill as a surgeon. (But I would resent his treating me *only* as a "vile body," a mere subject for the exercise or display of his skill.)

The Claim for Attention

The most immediate demand that flows from the claim to self-maintenance is that others should at least be aware of one's presence and existence—should *pay attention.* It is wounding to ignore somebody or to "cut" him; and it is a terrible punishment to be ostracized or "sent to Coventry." There are crude ways of rendering somebody, as it were, invisible or nonexistent; but there are also more subtle devices for ignoring another person and implicitly denying his or her worth.

In Dostoevsky's *Notes from Underground* the fictional narrator, suffering acutely from self-contempt and alienation, nourishes a pathological hatred of "a certain officer":

> Once, passing a cheap tavern, I saw through the lighted window some gentlemen fighting with billiard cues, and one of them was thrown out of the window. At another time I might have felt revulsion, but at that moment I was suddenly overcome with envy of the man who had been thrown out, an envy so acute that I even stepped into the tavern, into the billiard room: Perhaps I too could get into a fight, and also get tossed out of the window.
>
> I wasn't drunk—yet there you have it!—that's what misery can drive a man to. But it came to nothing. It turned out that I wasn't even capable of jumping out of the window, and I left without a fight.
>
> I was put in my place at the very outset by a certain officer. I stood near the billiard table unwittingly barring the way and he had to pass. He took me by the shoulders and silently—without warning and without apology—moved me from where I stood to another spot, and then walked by as though he had never noticed me. I could have forgiven blows, but how could I forgive just being moved like that and being so completely ignored?
>
> The devil knows what I wouldn't have given at that moment for a proper, regular quarrel, a more decent, more *literary* one, as it were! I had been treated as if I were a fly.[5]

Later in the story, the narrator's desire for revenge on this stranger who had ignored him so insultingly festers in the most fantastic way. He concocts various ways of getting his own back—fighting a duel and so on. The opportunity for action comes at last:

> And suddenly, three steps away from my enemy, I made up my mind in an instant, shut my eyes, and—we collided firmly, shoulder against shoulder! I did not yield an inch and passed him by entirely on an equal footing! He did not even glance back and pretended he hadn't noticed anything; but he was only pretending, I am convinced of that! I am convinced of it to this day. Naturally, I got the worst of the collision, for he was stronger, but that was not the point. The point was that I had achieved my goal, I had sustained my dignity, I had not yielded a step and had publicly set myself on an equal social footing with him. I came home fully avenged for everything.[6]

Dostoeveky's antihero was demanding of others not love or friendliness or even elementary decency and kindness but something more fundamental, the recognition of individual presence—and, by implication, of individual worth. To be tossed out of the window may be extremely disagreeable, but it implies being taken seriously; while to be ignored, to be treated as if one didn't exist—and by implication, shouldn't exist—is in some ways even more dreadful.

A passage from the first act of Arthur Miller's *Death of a Salesman* is in point. Willy Loman, the salesman of the title, is in serious personal trouble, and his wife, Linda, is pleading with their son Biff to show some compassion:

> I don't say he's a great man. Willy Loman never made a lot of money. His name was never in the paper. He's not the finest character that ever lived. But he's a human being, and a terrible thing is happening to him. So attention must be paid. He's not to be allowed to fall into his grave like an old dog. Attention, attention must be finally paid to such a person.[7]

"Attention must be paid"—that is the basic thing. That is perhaps what people mean when they complain of lack of communication. (I am not sure that what they really want is communication—if they achieved it, they might be surprised at what would actually be communicated.) What people quite rightly want is attention, recognition of presence, and the implied recognition of worth. Hence the widespread opposition to categorization and stereotyping—to standing in line, being submitted to computer routines, and the like—indeed endemic resentment against all procedures for treating people en masse, as if the individual counted for no more than an instance of some statistical configuration.

"Attention must be paid!" There are two distinguishable components in this basic injunction. First the demand that one shall pay attention, take account of, and not simply ignore another's presence and existence. This presupposes a minimal interest in the existence of someone else—which may, however, be compatible with indifference, dislike or even controlled hatred. (One

of the verses of Arthur Hugh Clough's *Latest Decalogue* runs: "Thou shall not kill; but needst not strive / Officiously to keep alive.") But more is needed.

Concern for Privacy

Paying attention, if it is to be a satisfactory basis for a morally approvable attitude of humane interaction, must imply some degree of positive concern for the other's personal values. At the very least, one should be committed to regarding the other's survival as good. But what more?

One way of approaching an answer to this question is to introduce the notion of a person's *private world.* Each of us has a domain that we like to regard as peculiarly our own. I am not thinking merely of one's own body, and such possessions as clothes, books, and an apartment or house, but rather of a domain also containing the persons, institutions, and things with which we particularly identify. We might think of our private world as a central part of our own life space: a field of perception and potential action, suffused with feeling tone, expanding and contracting with fluctuating attention and changing interest. My private world is the domain that I identify as *mine.* (One might speak, by biological analogy, of one's spiritual territory.) Within it are the things I love and cherish—and also those that I wish to weaken or destroy.

I would like to control my life space by preventing anybody from entering it or changing it without my permission. Trespass on my private world is what we call invasion of privacy. Preservation of one's privacy and the concomitant demand that such privacy be respected by others are fundamental features of the human condition. Every human being values his privacy and justifiably resents intrusions upon it.

Anthropologists and sociologists have taught us that notions of privacy are strongly conditioned by cultural and social factors: parts of the body that are private and hidden in one culture are public objects in another; and the notion of private property

varies from one society to another. Yet I feel that such diversity is still the manifestation of an underlying human invariant. I should be very skeptical of any report to the effect that some human society has no use at all for the notion of what is private. I cannot imagine a society where everything is public, including the very body itself. Each of us, I think, does and must have the conception, attenuated or inflated, of a realm in which he or she is the sole arbiter. (Am I here merely betraying the prejudices of an unregenerate individualist who sets an unreasonable value on privacy? I think not.)

The Englishman's home is proverbially his castle; and I presume that the master of a castle has the right to repel invaders. We all probably are, or would like to be, masters of our own castle. Certainly it is characteristically human to resent intrusions into or interferences with or limitations on such privacy.

Consideration for Others' Values

I said earlier that the component of humane interaction that is responsible to an individual's sense of the worth of his continued existence might be called *attention*. (I may remind you that I am treating attention as involving a positive evaluation of the worth of the other's existence.) But how shall we label the attitudes demanded by recognition of the character and especially the embodied *values* of the other's private world? "Respect," which suggests itself as fitting, is, to my mind, not quite right, because it suggests a positive endorsement of the other's values—of the way he sees the framework of his life.[8] What we need to identify is an attitude of admitting the other person's right both to organize his experience and to distribute his value judgments *as he or she does*—even in cases where we may have strong convictions about the distortions, errors, and mistaken perceptions of the other's life space. It is admittedly a difficult attitude to identify and a still more difficult one to maintain.

The classic formulation of the Golden Rule is "Do unto others as you would be done by." But what if the others have different

preferences? Are you to impose your moral views on them? The version of the rule that I recommend for present purposes—which might (with apologies to the late T. V. Smith) perhaps be called the Silver Rule—is: "Do unto another as you would be done by if you were he or she." (I repeat that this rule does not commit one to endorsing or sharing the other's moral evaluations: it is merely a first step, but an important one, in satisfying the necessary conditions for genuinely humane interaction.) As a label for the elusive attitude that I am recommending, I would suggest *consideration,* defined as the provisional acceptance of the salient features of the other's private world.

Consideration, thus conceived, already demands a degree of imaginative insight into the other's condition, which is hard to achieve at the best of times. (Here literature and other humanistic arts can be vitally important aids to achieving the understanding that is needed, if consideration is not to be restricted to the decorous ties, important as they are, of civilized and polite intercourse.) The point I wish especially to emphasize is that consideration, as I am conceiving it, implies a kind of concern. In admitting that other persons have a right to see the world and to judge its elements *as they see it*—whether they are hallucinating or fantasizing, and whether in our sober judgment they are deluded, morally corrupt, or insane—we are granting others an equal right to consideration.

Fellow Feeling

If humane interaction were restricted to attention and consideration it might still be somewhat calculating and frigid. Attentive consideration is compatible with principled hostility, with repugnance, with the desire to manipulate, if one can, another merely as an instrument. What we need is an infusion of *sympathy* that will convert what might otherwise, for all our underlying understanding of the other person and our provisional acceptance of his or her values, still be an antagonistic relationship into something warmer and more fruitful—a transition from the

mere "you and I" to the position of "we together." The best label I can think of for this quality is "fellow feeling"—Aristotle's *filia,* sometimes translated as "friendship" (perhaps too strong a word), which he claimed to be a "natural instinct" between members of the same species, "especially strong in the human race."[9]

As a paradigm of manifested fellow feeling I suggest the characteristic behavior of a hiker meeting another on some Alpine trail: the pause, the smiling responses, the exchange of greetings (possibly in hilariously broken versions of each other's languages), inquiries about itineraries and destinations, parting expressions of good luck, and so on. (Contrast the behavior of a suspicious and misanthropic wanderer, who averts his eyes, moves aside, and proceeds in surly silence.)

A natural expression of fellow feeling is "*There's* somebody of my own kind *(mon frère, mon semblable).*" It might be called a "we feeling"; it is grounded in recognition or anticipation of something shared, a feeling of kinship and affinity (animated perhaps by the pleasure of self-recognition—reflection in a human mirror). Accordingly, everything in fact shared—whether family relationship, previous association, common nationality or religion, occupation, hobbies, interests—promotes fellow feeling; and each revelation of divergence of properties, life histories, or personal traits tends to inhibit fellow feeling and to render its achievement more difficult. Perception of community and its anticipation are naturally pleasurable: fellow feeling, even in its most rudimentary form, anticipates, however optimistically, the pleasures of mutuality and community—the pleasures of *congeniality.* However, given the actual dependence of intensity of fellow feeling on the strength of the community of shared interests and shared values between interacting persons—the richness of what they have in common—the practical problem for those who want to foster humaneness is to strengthen feelings of kinship, rooted merely in shared humanity.

Is it really the most important thing about two interacting persons that they are both human (at least in a biological sense)? Is there a corresponding "sentiment for humanity" that can be genuinely held? Is it needed for the exercise of humaneness?

(What of a doctor who is treating a thoroughly disagreeable and cynical old codger who thinks of human beings as Golding does? Does that exempt the physician from working toward a "we relationship" and from perhaps proceeding, on the basis of a useful fiction, as if it had already been established? I should think not.) And what is the ultimate moral justification of the cultivation of fellow feeling? Is there some categorical imperative behind it—or is it, as I think, a matter ultimately of well-grounded choice? I leave these questions to the reader.

6/

The Prevalence of Humbug

> It would be a wonderful thing for mankind if some philosophic Yankee would contrive some kind of "ometer" that would measure the infusion of humbug in anything. A "Humbugometer" he might call it. I would warrant him a good sale.
>
> —*P. T. Barnum*

Introduction

Humbug has the peculiar property of being always committed by others, never by oneself. This is one reason why it is universally condemned. No doubt we can agree that humbug is a Bad Thing; but what are we agreeing about? It proves astonishingly hard to say. In trying to understand what humbug is, which is my main objective, one soon finds that no satisfactory definitions are available. I therefore propose to use for the most part an inductive approach. I shall offer a number of varied examples for consideration, hoping that we can eventually arrive at some reasonably satisfying analysis of this elusive concept. There should be time, also, to reflect on the mischief that humbug can work, and to consider some ways of curbing the disposition to produce it.

Chekhov's Lady

My first exhibit is drawn from Maxim Gorky's reminiscences of Chekhov:

> Once a plump, healthy, handsome, well-dressed lady came to him and began to speak *à la Chekhov:* "Life is so boring, Anton Pavlovich. Everything is so grey: people, the sea, even the flowers seem to me grey. . . . And I have no desires . . . my soul is in pain . . . it is like a disease."
>
> "It is a disease," said Anton Pavlovich with conviction, "it is a disease; in Latin it is called *morbus fraudulentus.*"
>
> Fortunately, the lady did not seem to know Latin, or, perhaps, she pretended not to know it.[1]

Morbus fraudulentus—literally, "the fraudulent disease"—is not listed in manuals of pathology, although the disorder is endemic, infectious, and seriously injurious to thought, feeling, and action. (A medical friend has compared it to rheumatism.) The Latin label is too opaque for common use, but "humbug" serves nicely. Chekhov's lady provides us with a clear example of humbug.

Bernard Shaw on Disarmament Conferences

My next example is taken from an interview granted by George Bernard Shaw to an American journalist (M. E. Wisehart) in 1930, on the eve of a naval conference. When the interviewer called the coming meeting a "disarmament conference," Shaw strenuously objected: "'Don't!' exclaimed Mr. Shaw. 'Everyone knows it's an armament conference! . . . The question is not 'Shall we do away with armament?' but 'How much armament?'"

The interviewer referred to the preliminary conversations between the British prime minister and President Hoover as "an event of great historical importance," and went on to say: "It is a harbinger of international understanding and good will. It has brought the English-speaking peoples together as never before and shown them that in sentiment, friendship, respect and good will they are united."

Shaw exploded: "Do you really believe that? Humbug."

When the interviewer said, "Why do you say 'humbug'?," Shaw replied:

> "Because, generally speaking, Englishmen and Americans do not like one another. Now they are asked to pretend that they do. And this pretense of being affectionate cousins is as dangerous as poison. Better to confess our dislike—our hatred, if you please—and ask ourselves what it is all about. Then there would be the possibility of ridding ourselves of it."[2]

The Shavian Probe

Bernard Shaw's formula, "Do you *really* believe that?" is a useful device; but it needs to be generalized into "Do you really mean that?" in order to fit cases involving something other than belief. In this form, it is a useful blunt instrument that deserves a label. I propose to call it the *Shavian probe.*

Unfortunately, it won't always work. No doubt, a journalist who actually thought in terms of bringing nations "together as never before" so that they become "united" in "sentiment, friendship, respect and good will" would be well advised to change his occupation—perhaps to that of a speech writer for presidential candidates. So in the case I have cited, Shaw's accusation of humbug seems justified.

But what are we to make of the following episode? On January 25, 1980, Mary McCarthy said, in an interview with Dick Cavett on Public Broadcasting, that Lillian Hellman was "a bad writer, overrated, a dishonest writer." Well, true or false, justified or not, there was no humbug about that. But on being asked by Mr. Cavett what was dishonest about Miss Hellman's writing, Miss McCarthy continued: "Everything. I once said in an interview that every word she writes is a lie, including 'and' and 'the.'" Well, did she *really believe* that one could lie by using the words "and" and "the"? Hardly—unless she was using "lie" in some extraordinary and unusual sense. But no doubt Mary McCarthy

was in earnest, and did *mean* what she said, was using just the words she wanted. That leaves the question of whether McCarthy committed humbug still unsettled: we have to undertake a difficult and controversial evaluation of the speaker's feeling and attitude.

According to Miss Hellman's lawyer, his client may get damages for defamation if she can show that "the person making the allegedly defamatory remarks either knew them to be untrue or uttered them without caring whether they were untrue." A philosophical logician might object that truth, in the sense of conformity with ascertainable matter of fact, is not in point. Yet a writer can be dishonest while saying something that is neither true nor untrue: common sense would regard some caricatures as libelous even if they were uncaptioned. When the lawsuit is ultimately heard, it will be interesting to see how judge and jury will cope with the possibly extenuating effects of context—and with such controversial issues as the applicable legal constraints on emotive or offensive language.

Some Preliminary Comments

Chekhov's lady and Shaw's interviewer provide clear cases of humbug. I think of them as touchstones of usage—paradigm cases. For me, the two examples are cases of humbug if anything is; if you disagree, then your usage of the key word probably differs in certain respects from my own. These paradigm cases have some readily discernible features that are worth noting, for future reference in struggles with more controversial examples.

One reason why "Humbug!" is so offensive an exclamation is that it charges the speaker with some kind of falseness. But in neither of our cases was the speaker supposed to be lying. For Chekhov's lady was not necessarily lying when she said that everything looked and felt "grey" to her: perhaps she used that very word in her private thoughts. Humbug need not entail lying in the strict sense of that word—even though humbug can be akin to outright lying.

We can usefully distinguish between the speaker's *message,* as I shall call it, and his or her *stance.* By the message I mean whatever is explicitly or implicitly said about the topic in question; while I reserve the term "stance" for the speaker's beliefs, attitudes, and evaluations, insofar as they are relevant to the verbal episode in question. To illustrate: If you say to me in a confident way, "The plane leaves at four o'clock," you are not *saying* that you believe what you say, because that is not the topic on which you are supplying information. But of course, by speaking as you do, in a standard situation in which trustworthy information is normally expected, you are giving me reason to believe that you are not deliberately misleading me.

Similarly, Mary McCarthy, to return to the earlier example, was not *saying* that she despised the subject of her scathing comments, nor was she overtly claiming to be sincere, but clearly she spoke as one who expected her remarks to be taken as a sincere expression of contempt. (Try saying: "She's disgusting—but of course I don't feel disgust." That would be a paradox: we could make some sense of it, but not without hard work.)

Now, the pejorative implication of a charge of humbug is commonly leveled against the content of a message (a remark or a text) rather than at what I have called the speaker's stance: then it usually has the force of "Stuff and nonsense," denigrating the message without necessarily imputing falseness or insincerity.[3]

Consider the following minidialogue:

> First speaker: As McLuhan has taught us, the medium is the message.
> Second speaker: Humbug!

Here I take the second speaker to be rejecting McLuhan's absurd slogan, that is, rejecting the *substance* of what is being said: there need be no imputation about the sincerity of the speaker's stance. For he or she may genuinely regard McLuhan's widely quoted fragment of pseudo-wisdom as profoundly illuminating. No matter: without impunging a speaker's stance, we can sometimes condemn what is being said as balderdash, claptrap, rubbish, cliché, hokum, drivel, buncombe, nonsense, gibberish, or

tautology. With so rich a vocabulary for dismissing the substance of what is said, we could dispense with this use of "humbug." That useful word might well be reserved for criticism of a speaker's stance—to discredit the message's provenance rather than its content. I shall respect current usage, however, by sometimes using "humbug" in the sense of "piece of humbug."

What then is the prima facie charge against a speaker accused of humbug? Well, some of the words that immediately suggest themselves are pretense, pretentiousness, affectation, insincerity, and deception. Often there is also a detectable whiff of self-satisfaction and self-complacency: humbug goes well with a smirk. A common symptom is clever-me-ism, as in Jack Horner's case. In this respect, it resembles *cant,* which Dr. Johnson memorably defined as "a whining pretension to goodness, in formal and affected terms."

To say that humbug has something to do with insincerity and deception is to point in the right direction, but does not sufficiently identify the word's meaning. Let us see whether the history of the word's changing uses can provide a more specific analysis.

A Short History of the Word's Shifting Meanings

I used to think that the word "humbug" came into general use in the nineteenth century—possibly because I took the Victorians to be especially prone to hypocrisy. To my surprise, I discovered that its career dates from the middle of the eighteenth century, when it seems to have entered the language as "a piece of fashionable slang" *(Century Dictionary)* of unknown origins. It may have been used originally in the restricted sense of a false alarm, a hoax, or a practical joke. But its meaning was uncertain even from the start. In 1751 a writer (quoted in the *Century Dictionary*) complained about it in the following terms:

> There is a word very much in vogue with the people of taste and fashion, which though it has not even the 'penumbra' of a meaning yet makes up the sum total of the wit, sense and judgement

> of the aforesaid people of taste and fashion! [He gives quotations] *Humbug* is neither an English word nor a derivative from any other language. It is indeed a blackguard sound made use of by most people of distinction! It is a fine make-weight in conversation, and some men deceive themselves so egregiously as to think they mean something by it!

Dr. Johnson did not include the word in his dictionary (1775), possibly because he thought it too coarse or vulgar to be noticed.

By 1828, the first edition of Webster's dictionary treated "humbug" as an approximate synonym for "swindle" or "fraud." As a noun, Webster says, it refers to "an imposition under false pretences"; and as a verb it means "to deceive; to impose upon"—or, as we might nowadays say, "to con." So the relatively innocent old sense of a practical joke had made way by then for something more obnoxious.

An explicit example of this use occurs in *The Pickwick Papers,* when Mr. Pickwick furiously upbraids Mr. Winkle for pretending to be able to skate:

> Mr. Pickwick . . . uttered in a low but distinct and emphatic tone, these remarkable words,—
>
> "You're a humbug, sir."
>
> "A what!" said Mr. Winkle, starting.
>
> "A humbug, sir. I will speak plainer if you wish. An impostor, sir."

Readers familiar with *The Wizard of Oz* may remember the unmasking scene, in which the Great Wizard is finally revealed as a timid old man, who confesses to having been making believe.

> "I'm supposed to be a Great Wizard."
>
> "And aren't you?" [Dorothy] asked.
>
> "Not a bit of it, my dear; I'm just a common man."
>
> "You're more than that," said the Scarecrow, in a grieved tone; "you're a humbug."

Webster's *Third New International Dictionary* of 1966 contains the following entries: for the noun's head sense, "something designed to deceive and mislead" (with cross-references to

quackery, hoax, fraud, and imposture); for the verb, "impose on" (with cross-references to deceive, cajole, and hoax); for its application to a humbugger, "a person who usually willfully deceives or misleads others as to his true condition, qualities, or attitudes, one who passes himself off as something that he is not" (with cross references to sham, hypocrite, and impostor). The general impression left by this rather confused set of definitions is of adherence to the strong nineteenth-century equation of a humbug with an impostor or swindler. But that identification fails to reflect present usuage. If a main sense of "humbug" were that of something designed to deceive and mislead, a skillfully constructed wig would have to count as a prime example. If a humbugger is a person who willfully deceives others, then the pseudo-Arabs lately used by the FBI to "sting" congressmen would be properly described as humbugs.[4] Something is plainly wrong.

By relying too much on the entries in earlier editions, the makers of Webster's Third have overlooked the present dilutions of the old intensely pejorative implications of "humbug." I hope that in the end we can do somewhat better, by considering some further examples of clear cases and test cases.

Russell's Tirade

Consider now the following glimpse of the private behavior of a famous philosopher. In a charming book of reminiscences, Rupert Crawshay-Williams tells of accompanying Bertrand Russell on a trip to inspect a house that was being remodeled for Russell's use. When the two friends arrived, they had to suspend a lively discussion of a new book by the pragmatist F. C. S. Schiller.

On meeting the builder and the architect, who were awaiting his arrival, Russell at once launched a furious denunciation of their supposed delinquency. He brushed aside their excuses, called the architect a liar, and hardly allowed the others to finish a sentence.

The tirade rolled over them until both of them were left floundering and gasping.

Russell ended off by demanding a complete change in their future behaviour. He stopped talking and walked smartly out to the car; we got in; I started the engine;

'So Schiller was really making the context of the statement part of its meaning', said Russell.

Elizabeth and I were still stunned.

'But Bertie', we said, 'you seem quite calm!'

'I *am* quite calm', he said. 'That's taught them a lesson I think, hasn't it?'

'We certainly think they were impressed. Do you mean to say', we asked, 'that the whole explosion was deliberate and contrived?'

'Yes indeed', said Russell, 'it was the only thing to do—the only way of making an effect.'

'Well, I suppose it may work', I said. 'But I did think you were being just a little bit unfair at times.'

'Unfair!' Of *course* I was being unfair.'[5]

Crawshay-Williams, reverting to a previous conversation, then said: "There you are . . . it's what we were saying last night: you're an aristocrat, and I'm merely a gentleman." Fair comment, although a proper English gentleman may not call himself a gentleman, except ironically.

Ghotbzadeh's Indignation

At the end of January 1980, the Canadian government announced the escape of six American diplomats who had been sheltered in the Canadian embassy in Teheran.

According to the news reports, one of the so-called embassy militants responded by crying, "That's illegal! That's illegal!" When the Iranian foreign minister, Sadegh Ghotbzadeh, later met reporters, he took the same line, calling the secret operation a "flagrant violation" of international law. " 'They have violated the laws they claim to defend,' Mr. Ghotbzadeh said of the Cana-

dians. He denounced 'so-called international laws' as having been made only 'for the suppression of the small nations by the big ones.'"[6] To this egregious nonsense Ghotbzadeh added a veiled incitement to violence—"Canada will pay" and "Everybody is free to do whatever they want"—apparently intending "an open invitation to Iranians around the world to take action against Canada or Canadians"; also an allegation that he must have known to be a lie, to the effect that he had received an apology from the Canadian prime minister, with an accompanying explanation that "the action had been for political reasons in Canada" (both immediately denied by Joe Clark).

In this farrago, I am particularly interested for present purposes in the role played by what is surely a prime case of humbug, the implicit presentation of the speaker (what I have previously called his "stance") as one who, himself respecting international law, is therefore entitled to complain of an alleged violation. Ghotbzadeh's own explicit denunciation of "so-called international law" merely adds to the confusion of what an editorial in the *New York Times* called his "flagrant logic."

Emerson's Friendship

Ralph Waldo Emerson says the following about friendship, in his celebrated essay with that title:

> The moment we indulge our affections, the earth is metamorphosed: there is no winter and no night: all tragedies, all ennuis vanish,—all duties even; nothing fills the proceeding eternity but the forms all radiant of beloved persons. Let the soul be assured that somewhere in the universe it should rejoin its friend, and it would be content and cheerful alone for a thousand years.[7]

Well, all of us have sometimes been kept waiting for a good friend, but a delay of a thousand years is, as the English say, a bit much. But Emerson is relentlessly enthusiastic about friendship:

> Happy is the house that shelters a friend! It might well be built, like a festal bower or arch, to entertain him a single day. Happier, if he know the solemnity of that relation and honor its law! It is no idle bond, no holiday engagement. He who offers himself a candidate for that covenant comes up, like an olympian, to the great games where the first-born of the world are the competitors. He proposes himself for contest where Time, Want, Danger, are in the lists, and he alone is victor who has truth enough in his constitution to preserve the delicacy of his beauty from the wear and tear of all these.[8]

Surely there is something suspicious about this exaggerated rhapsodizing. Indeed, the very last paragraph of the essay suggests that Emerson might really have preferred the kind of friend that need never be met:

> It has seemed to me lately more possible than I knew, to carry a friendship greatly on one side, without due correspondence on the other. Why should I cumber myself with the poor fact that the receiver is not capacious? It never troubles the sun that some of his rays fall wide and vain unto ungrateful space, and only a small part on the reflecting planet. Let your greatness educate the crude and cold companion. If he is unequal he will presently pass away; but thou art enlarged by thy own shining, and no longer a mate for frogs and worms, dost soar and burn with the gods of empyrean.[9]

"It never troubles the sun that some of his rays fall wide and vain"; "thou art enlarged by thy own shining." Enlarged or puffed up? Emerson sometimes reminds me of Mr. Pecksniff, who is described on his first appearance in *Martin Chuzzlewit* as "a moral man, a grave man, a man of noble sentiments, and speech." Dickens says of him (in connection with his calling his daughter Mercy a "playful warbler"):

> "Playful—playful warbler," said Mr. Pecksniff. It may be observed in connexion with his calling his daughter "a warbler," that she was not at all vocal, but that Mr. Pecksniff was in the frequent habit of using any word that occurred to him as having

> a good sound and rounding a sentence well, without much care for its meaning. And he did this so boldly and in such an imposing manner that he would sometimes stagger the wisest people with his eloquence and make them gasp again.

The Czar's Vodka

The back cover of *The New Yorker*'s issue of January 21, 1980, displays a richly colored photograph captioned "The spirit of the Czar lives on." We see an impressively bearded man, head tossed back, with a smidgeon of a smile, dressed in full regimentals, with scarlet jacket, white sash, and enough medals to start a collection. In one hand he holds a wineglass, with the other he is fondling the neck of a fine borzoi. Meanwhile his czarina, with an equally lavish display of evening dress and jewels, squats on the imperial carpet to play with a couple of borzoi puppies. And the message? This, in part:

> It was the Golden Age of Russia, and the Czar reigned supreme. Europe, Asia: all the empire was his. Regal coaches carried him in elegance, but with his Cossacks he rode like thunder. Hunting wild boar in the northern forests, hosting feasts for a thousand guests in the Great Palace, no man could match the Czar's thirst for life. And his drink? The toast of St. Petersburg. Genuine Vodka.

(I omit the brand name of what I think of as Humbug Vodka.)

It would be a waste of time to criticize this text by asking such questions as why the pseudo-czar is wearing all those medals, and why he "rode like thunder," and whether he really did thirst for life as nobody else could, and what all this has to do with the barely perceptible difference between one vodka and another. We know that good sense and relevance have nothing to do with the case, the desired effect on the impressionable reader being achieved if favorable associations are created.

Zaftig Bedfellows

Consider now the following item from the "Personal" columns of *The New York Review of Books* (February 21, 1980):

> ZAFTIG FEMALE WANTED. NYC male, 35, editor/author, lean, reasonably good-looking, financially secure. Seeks woman Renoir would have painted, 25–35, with pretty face, stable personality. Excellent opportunity to share museums, movies, affluence, quiet conversations, caring, maybe marriage. NYR, Box 2956.

If you think this hard to beat for vulgarity, listen to the journal's own puff, on the same page, for its new English affiliate:

> AMERICAN INTELLECTUALS too warm and bloody open for you? Prefer treacle in your tarts, not your mail? Tired of being mashed by provincial American bangers? Woo your very own little gooseberry fool abroad merely by running an accurate but winning Personal Ad in the columns of *The London Review of Books*.

A rare example, this, of puffing a go-between. It might be called pimping for a pimp—or metapimping.

The Cornell Ship

Ceremonial and political occasions invite humbug. Here is a prime example from the early history of Cornell University:

On Inauguration Day, October 7, 1868, the new university had so many more students than it could handle that "the department of geology was confined to a single room adjoining the coal cellars, and demonstrations in natural history were conducted in the vacant space next to a furnace." The campus consisted partly of "a ravine six or eight feet deep, bridged by two dirt causeways."[10] Against this kind of backdrop, George William Curtis delivered an elaborate address with the following peroration:

> Here is our university, our Cornell, like the man-of-war, all its sails set, its rigging full and complete from stem to stern, its crew embarked, its passengers all ready and aboard; and even as I speak to you, even as the autumn sun sets in the west, it begins to glide over the waves as it goes forth rejoicing, every stitch of canvas spread, all its colors flying, its musical bells ringing, its heartstrings beating with hope and joy.[11]

"Complete from stem to stern"—and students in the coal cellars! The university's cofounder and president, Andrew Dickson White, "looking out over the ragged cornfield and the rough pasture land and noticing the unfinished buildings and the piled-up rubbish," felt that no words "could fail more completely to express the reality," and Ezra Cornell confessed that there was "not a single thing finished."[12] Curtis, one might say, was operating on the principle of "Take care of the sounds and the nonsense will take care of itself." The sentiments were appropriately edifying, and the elaborately studied and rehearsed phrasing evoked the expected applause.

Academic Humbug (Veblen)

Any survey of the varieties of humbug should include specimens of the pretentious verbiage that infests scholarly writing. I reproduce the following from Thorstein Veblen's book *The Theory of the Leisure Class* (part only of an extraordinarily long paragraph) for the pleasure of resuscitating H. L. Mencken's commentary on it.

> In an increasing proportion as time goes on, the anthropomorphic cult, with its code of devout observances, suffers a progressive disintegration through the stress of economic exigencies and the decay of the system of status. As this disintegration proceeds, there come to be associated and blended with the devout attitude certain other motives and impulses that are not always of an anthropomorphic origin, nor traceable to the habit of personal subservience. Not all of these subsidiary

> impulses that blend with the bait of devoutness in the later devotional life are altogether congruous with the devout attitude or with the anthropomorphic apprehension of sequence of phenomena. Their origin being not the same, their action upon the scheme of devout life is also not in the same direction. In many ways they traverse the underlying norm of subservience of vicarious life to which the code of devout observances and the ecclesiastical and sacerdotal institutions are to be traced as their substantial basis.[13]

Here we have garrulity laced with jargon. Mencken says that Veblen "achieves the effect, perhaps without employing the means, of thinking in some unearthly foreign language—say Swahili, Sumerian or Old Bulgarian—and then painfully clawing his thoughts into a copious but uncertain and book-learned English." As to the long passage he quotes, Mencken concludes that Veblen is trying to say "that many people go to church, not because they are afraid of the devil but because they enjoy the music, and like to look at the stained glass, the potted lilies and the rev. pastor." Mencken says that "this highly profound and highly original observation" might have been made on a postage stamp, thereby saving a good deal of wasted paper.[14]

Misfires and Violations

With these varied examples of ostensible humbug before us, we can ask what it is about such episodes that inclines us to regard all of them, in spite of their obvious differences, as instances of the same complex phenomenon. Do we mean the same thing each time, or are we perhaps applying the pejorative label to cases connected only by shifting similarities, rather than by the presence of some detectable common property?

Let us first recall the great amount and variety of information normally transmitted in even the simplest and most familiar kind of conversation. Suppose the driver of a stationary automobile asks me the way to, say, Route 13. I would normally take for granted much concerning the speaker's situation and compe-

tence that is unsaid, indeed much that would mark the episode as perplexing if it were said: for instance, that the driver wants to get to the highway in question, that he is en route to some other destination, and that he doesn't know how to proceed. Also, on the evidence of his question, that he is a native English speaker who knows what he is saying and hasn't made a slip of the tongue. Correspondingly, I assume that he himself is making parallel assumptions about my own understanding of his problem and willingness to help. (I shall ignore any further information possibly conveyed by signs of anxiety, distress, and the like.)

I propose to speak in such cases of the *framing presuppositions* of the initiated verbal transaction—or, more briefly, of the conversation's *framework*. Establishing the framework—an operation so commonplace that we normally fail to notice it—determines the character of the initiated conversation in a way that is crucially important for the possibility of a successful outcome. (Of course much talk has little discernible purpose, amounting to no more than friendly chatter or cocktail-party babble.)

The centrality of the role played by what I have called the conversation's framework can be highlighted by cases of willful falsification of the presupposed understandings. Suppose that on being asked by a stranger, "Do you know the way to the campus from here?" I simply reply, "Yes." That will probably get me a look of resentful incomprehension, especially if I respond to the further question "Would you like to tell me how to get there?" by saying, "No." Please notice that, far from lying, I may be literally—yet quite inappropriately—telling the truth. In such a case I would of course be willfully violating the conditions that normally enable the kind of conversational exchange in question to proceed. No doubt I would be resentfully regarded as "trying to be funny." If I then suffer a change of heart, pursue the departing stranger, and, having caught up with him, say, "Would you like me to tell you the way?" he may play the same down-putting game by saying, "No!"

I shall now contrast two different kinds of ways in which intended exchange of information may fail. The first type of

case, which I shall call a *misfire,* results from ignorance or incompetence on the respondent's part: I might mishear the number of the highway in question, or get the number right but not know how to get there, or I might even be suffering from some painful personal anxiety that made me unable to help. In the absence of such impediments I might be simply inept in giving intelligible and useful instructions. Such misfires—or as our president recently called them, in his usual euphemistic mode when caught blundering, "failures of communication"—are sufficiently common to induce caution in relying on testimony or authority, however generally reliable and useful. But the risk of misfire is no ground for radical skepticism about the feasibility of successful communication in relatively unproblematic cases. If one stranger doesn't know the way, perhaps another does; if some passer-by is too stupid to understand my problem or too selfish to help, perhaps another will.

Far more serious than such occasional hitches in communication (misfires, in my terminology) are breakdowns in communicative interaction induced by deliberate falsifications of the constitutive framework. To start with relatively innocuous but still potentially pernicious abuses of this sort: a prankster might perversely pretend not to understand the motorist's question, or pretend to be unable to speak (pointing meaningfully at his own throat), or even deliberately act as if he were a lunatic. (The case I previously considered of an absurdly literal interpretation of a polite formula would also fall under this heading.) In such a case I propose to speak of a *violation* of the standard framework. Violations, unlike misfires, are not the predictable and excusable consequences of human ignorance or incompetence. They maliciously trade on and undermine the implicit understandings that underpin successful communication and cooperation, and hence erode the foundations of social existence. (Imagine a society in which joking was so common that one could never be certain whether communications were serious or maliciously disruptive.) Violations of the understandings that sustain communication must be regarded as perversions of verbal interaction, animated by deliberate deceit.

How Humbug Differs from Lying

We have already seen that violations of the communicative framework need not consist in the utterance of falsehoods. If I reply on the telephone to the question "Have you got any sausages today?" by saying, "No," and continue in the same vein, saying that I won't have any in the foreseeable future, and the like, everything I say may be literally true, but I shall deceive the other as if I were deliberately lying. As William Blake said (in *Auguries of Innocence*):

> A truth that's told with bad intent
> Beats all the lies you can invent.

There is good reason, however, to regard conscious and deliberate falsity—what Immanuel Kant calls "intentional untruthful declaration"—as having primary theoretical importance. Sissela Bok, in her valuable book *Lying,* is following a long-established tradition when she confines her discussion to explicit lying, defined as the production of an utterance expressing what the speaker disbelieves.[15] Let us call such an utterance, with acknowledgment to Shakespeare, the *lie direct.* Until further notice, when I speak of lying I shall mean the utterance, *as if believed,* of an assertion disbelieved by the utterer.

Moralists have long regarded brazen lying (the deliberate assertion of "the thing that is not," or at least "the thing thought to be not") as meriting the severest reprobation. Montaigne said, "Lying is indeed an accursed vice. We are human beings, and hold together only by speech. If we knew the horror of it, and the gravity, we should pursue it with fire, and more justly so than other crimes."[16] In such statements as this, lying is regarded as a cardinal vice, an unforgivable sin against humanity. Such a view led Immanuel Kant to claim—implausibly, I think—that we have an absolute duty not to lie, even when a truthful reply would lead an intending murderer to his victim. If this rigoristic condemnation of lying and the corresponding elevation of truthfulness to a supreme virtue were justified, it would be hard to understand the absence of references to lying in the Ten Com-

mandments (except in the special prohibition against "bearing false witness") or in other religious compendia of vices and virtues.

Traditional defenses of such rigor are unconvincing. Kant, like other writers on the subject, notices the damage inflicted on the liar himself, but reserves his most earnest condemnation for the damage inflicted by a lie on all humankind. But as much can be said about all vice: every criminal injury to an individual also damages the moral fabric of society. Nor is a lie the worst injury: violence and willful cruelty may be regarded as at least equally damaging to the moral character of the perpetrator and as a violation of general obligations needed for social life.

A plausible justification of the prominence assigned to the vice of lying is suggested by a gloss of Montaigne's contention that as human beings "we . . . hold together only by speech." It is obvious that children could not learn to speak if they were reared only by adults who lied to them irresponsibly and at random. For in such an environment a child could not even learn the common names of familiar objects. Closely connected with this point is the familiar observation that the liar is parasitic on general, though not universal, veracity: lying, as a species of deceit, would be futile in the absence of general efforts to be truthful. It seems reasonable to conclude that a liar is, in a radical way, sapping the foundations of social institutions, all of which depend on the general effectiveness of speech. The liar is indeed an enemy of society, who tends to undermine all possibility of civilized intercourse. Universal lying would destroy intelligible speech.

Still, an endorsement of Montaigne's emphasis on the gravity of shamelessly explicit lying needs some supplementation, if it is not to be misleadingly one-sided. The immediate harm done by a successful lie direct—the deceitful generation of a false belief by concealed violation of the standard framework—can often be achieved more efficiently, and with less fear of detection or reprisal, by indirect means. One can intimate "the thing that is not" by implication, by significant silence, or even by the double bluff of pretending to lie while actually speaking the truth (as in the classical Minsk-Pinsk joke). (The annals of espionage are a

rich source for this kind of deception.) Such maneuvers, a standard resource of advertising and diplomacy, are secure against the accusation of explicit and knowing mendacity: the offender can always plead that he didn't literally *say* anything that he himself disbelieved. Given the prevalence and effectiveness of such indirect ways of achieving the disreputable benefits of lying, it is surprising that we have no better label for indirect verbal deception than the lawyer's tag of *suggestio falsi*. We might perhaps speak in such cases of *virtual lying*. (Webster defines the relevant sense of "virtual" as "being functionally or effectively but formally not of its kind.")

With virtual lying, we are at last in the close neighborhood of the kind of humbug that "functionally or effectively" implants false belief. For in such cases there is characteristically a *conscious* discrepancy between the utterer's beliefs and the false beliefs to be implanted. Such cases cannot properly be regarded as cases of outright lying, but are all the more pernicious for that reason. The person who composed the vodka advertisement probably believed that the drink he was puffing was virtually indistinguishable in taste and sedative power from any of the competing brands on the market. In eschewing direct lies or even, for the most part, virtual lies for which he might be accountable, he was relying on the powerful forces of suggestion and association—with all the flummery of the czar's legendary court and so on—to implant what would have been a naked lie if it had been explicitly stated. The difference between such cases of humbugging deception and outright or even virtual lying is not in the content of the communicated message, or in the intention to deceive by implanting false beliefs, but rather in the sophistication of the means used to achieve the purpose.

The continuities between explicit lying, virtual lying, and what I now propose to call *falsidical humbug* (I borrow the term "falsidical" from W. V. Quine) have tempted many writers to regard the conventional distinction between lying and humbug as superficial and ultimately misleading. Indeed, some writers will assimilate to lying even relatively harmless efforts to make a good public showing. Thus Adrienne Rich, in her notes on lying, says: "We have expected *to lie with our bodies:* to bleach,

redden, unkink or curl our hair, to pluck eyebrows, shave armpits, wear padding in various places or lace ourselves, take little steps, glaze finger and toe nails, wear clothes that emphasized our helplessness."[17] (One supporting myth regards only nudity as genuinely natural and "truthful," all concealment or clothing being counted as hypocritical.)

The tendency here illustrated to convert similarities into supposedly profound underlying identities (dressing and personal adornment as "really" the same thing as lying) is a constant temptation for philosophically inclined scholars in search of excitingly paradoxical insight.

For a splendid example one might turn to George Steiner and his startling rediscovery of "the creativity of falsehood," which he characterizes as "a seminal, profound intuition" of the Greeks.[18]

> Linguists and psychologists (Nietzsche excepted) have done little to explore the ubiquitous, many-branched genus of lies. . . . Constrained as they are by moral disapproval or psychological malaise, these inquiries have remained thin. We will see deeper only when we break free of a purely negative classification of 'untruth', only when we recognize the compulsion to say 'the thing which is not' as being central to language and mind. We must come to grasp what Nietzsche meant when he proclaimed that 'the Lie—and *not* the Truth is divine!'[19]

Steiner quotes Nietzsche again, approvingly, as saying, in *The Will to Power:* "There is only *one* world . . . and that world is false, cruel, contradictory, misleading, senseless. . . . We need lies to vanquish this reality, this truth, we need lies *in order to live.* Steiner seems himself to endorse the view that "lying is a necessity of life" by which "man violates an absurd confining reality" in a way that "is at every point artistic [and] creative."[20]

This confused and shoddy defense of lying is what Jeremy Bentham would have called "nonsense on stilts." Steiner here seems to be emulating that kind of German metaphysician than whom, according to Carlyle, none could dive deeper or emerge muddier. It is a prime example of the kind of academic or schol-

arly humbug that consists of saying more than you can reasonably mean, for the sake of the booming sound of your periods (what the older rhetoricians called "bomphoiologia"). If you fail to make the distinctions that I have been proposing between plain lying, virtual lying, pretentious inflation of belief, and so on, proceeding so far, in Steiner's case, as to regard hypothetical if–then statements as cases of lying, you will end with a conceptual gruel in which everything looks like everything else and all intellectual distinctions have vanished in the service of grandiose obfuscation.

Still, a decent respect for the conceptual distinctions between plain and fancy lying and the allied but distinguishable varieties of deceptive humbug, such as I am here advocating, leaves as yet undiscussed the question of relative harm. The subversive effects of the brazen liar in undermining the foundations of linguistic institutions might be compared to outrageous violations of the constitutive bases of other institutions. There is something peculiarly monstrous about a judge who accepts bribes, a farmer who adds poison to his corn, or a doctor who infects patients in order to ensure a steady income. Yet the adulteration of food, to stay with that example, may, in the not very long run, be even more harmful than outright poisoning. The most serious indictment of falsidical humbug is that, without directly striking at the roots of linguistic institutions, it tends progressively to adulterate speech and thought. As a recent writer has well said, the "distortion of values, this insidious numbing of what we once knew without question as true or false, can be blamed, in part, on the language we hear and read every day and night."[21]

The Complexities of Self-Deception

When humbuggers say what they themselves disbelieve, evading the risks of lying while reaping its benefits, the gross discrepancy between utterance and actual belief (the speaker's stance) can sometimes be established beyond all reasonable doubt. If the perpetrator rebuffs the Shavian probe—"Do you

really believe that?"—by insisting that he or she really did believe it, bolstering the original humbug by a brazen lie, perhaps tone, facial expression, or actions will expose the fraud. I call such conscious deception first-order humbug.

Humbug is often less obvious and forthright. Suppose a college student told Vladimir Nabokov that he was the greatest writer since Gogol (an imaginary but plausible episode). Any eavesdropper who knew the student's shaky standing in Nabokov's course on Russian literature might question the flatterer's sincerity; but how is the imputed bad faith to be established against reiterated protestations of sincerity? If we suppose the flatterer to be subjectively honest, we might still impute *self*-deception. If so, we shall have a good example of a self-humbugged humbugger producing what I shall call second-order humbug.

Although self-deception is perhaps as common as lying, there is a difficulty in understanding how it can possibly occur. Consider the conditions for successful deception of one person by another. If somebody else is to be successfully deceived, what I say must seem initially plausible and my assertion of it must provide *some* reason for the other's acceptance. Hence my own disbelief must be concealed. Should any of these conditions be violated, the attempted deceit will fail: if you say that you are the illegitimate son of the monarch, as one of the British spies used to do, your hearer will probably think that you must be joking (the intended effect); if you show by a wink that you don't believe what you say, your hearer will not succumb to the intended deception. The deceptive appearance must masquerade as reality.

How can one hide one's own disbelief in an intended act of private deception? Is it not absurd to say to oneself, "I don't believe such-and-such and yet I am going to believe it?" One cannot be an authority for oneself, and nothing that I know that I disbelieve can be a reason for me to believe it. And how can I fail to know my own disbelief? Ivy Compton-Burnett once said in an interview, "I don't think there is such a thing as self-deception. When people say they do things unconsciously and subconsciously, I am quite sure they do them consciously."[22]

The following argument for the impossibility of self-deception seems to be conclusive: Humbug requires concealment of a deceptive intent; but if the speaker and the audience are identical, as in soliloquy, there can be no such concealment; so there can be no such thing as self-deception.

One might respond by pointing to clear cases of what we call self-deception—as when a woman shows by words and actions that she still believes in her son's survival, although she possesses proof of his death in battle. Anybody persuaded by the impossibility argument would presumably retort that what happens in such an instance is misdescribed as deception and ought properly to be called something else. But this would amount to an arbitrary change in language, motivated by nothing better than obstinate defense of a dubious argument.

The impossibility argument is underpinned by the following conceptions: "Either you know that you believe what you say or else you don't. And in either case you can't be mistaken." Knowledge of one's own belief is immediately accessible; and there is no middle term between belief and disbelief. Both contentions are wrong: knowing one's momentary belief is not, like a sneeze, a hit-or-miss affair; and various degrees of awareness may be involved.

Consider the following typical example. Before meeting my doctor to hear the latest report on some chronic affliction, I resolve to take a cheerful view. Then, while the doctor talks, I withdraw attention from, blank out, anything that begins to sound like bad news, while attending closely to encouraging remarks. In this way I end by genuinely believing a comfortable but wrong conclusion, based on deliberately selected and distorted evidence. When I close my ears in this way against a sentence that starts "Unfortunately your blood pressure . . ." so that I barely hear the rest of the sentence, do I know what the doctor says? Do I believe what he says? Am I aware of using a strategy of selective attention in support of a predetermined verdict? In each case the answer has to be "Yes and no."

Do I hear the bad news about the blood pressure? Unless I am an unusually talented self-deceiver, I probably do—as shown by the fact that I may reluctantly be able to dredge it up into full

consciousness later. But the censored news is relegated to the back of my mind; I know it as I know that I am writing a letter when I am attending to something other than the act of writing. One might say that I have twilight awareness of the suppressed material. (There is no need yet to invoke the unconscious.) Do I believe what I hear in this twilight way? Again, yes and no. Yes, because unless I do believe it, on my doctor's unquestioned authority, I have no need to suppress it; no, since I manage to prevent the very question of belief from rising into full consciousness. Parallel verdicts apply to the overall program of selective attention and wishful distortion that I execute. I know what I am doing as I know that I am walking even while I am thinking about something else; but having a normal distaste for distortion of evidence, I need to mask my disreputable strategy.

The foregoing analysis of what might be called, in Jean-Paul Sartre's terminology, "bad faith" *(mauvaise foi)* seems to fit the most familiar cases of self-deception. But it does not fit severe cases of the repression of unwelcome thoughts or tidings, where the strategy of selective attention and rejection can induce neurotic symptoms. Yet it does fit such testimony as the following, from a recent discussion of obesity: "Even while on a binge [of gobbling] one vows to start a diet tomorrow and emerge from it miraculously transformed."[23]

A further feature of the process of selective attention and repression of unwelcome input in the service of what might be called wishful acceptance deserves notice. In brushing aside the bad news of which he is at least partly aware, the self-deceiver makes the welcome good news part of his consciousness, part of himself, as it were, while doing his best to pretend that what he partly hears and would like to forget simply has not happened. But it has happened and he knows that it has. So there is a kind of *dissociation* at work, as in the familiar cases of motives that are suppressed as unworthy. So long as the self-deceiver is in the initial stages of the process sketched above, he has a "divided self," a state of strain that is disagreeable to all but accomplished hypocrites.

The constant practice of self-deception may produce a character that cheats as effortlessly as a bird sings: the mask eventually

becomes ingrown, fits the face as closely as a death mask. Exposure of such inveterate self-deception is difficult, since it requires a critical judgment of a whole way of life. When we are coping with confirmed hypocrisy rather than momentary self-deception, a verdict of humbug seems euphemistic.

Yet the exposure of episodic self-deception is sometimes both practicable and useful. Second-order humbug can sometimes be detected even if the producer is unshakably convinced of subjective honesty and sincerity. It would itself be an act of humbug, however, to suppose that the critic can pride himself on being free of self-deception. Selective and differential attention, repression, and dissociation are features of all perception and thought. And even wishful acceptance is not necessarily reprehensible when it leads to beneficial results, at least in the short run. Should we denigrate the wishful belief that one is going to win a contest when self-verifying predictions are involved?

Is Humbug Ineradicable?

Ought implies can, say the moralists. So before considering antihumbug remedies, we ought first to hear the vehement objections of those who consider such a project dangerously quixotic because, as one writer has put it, "deception is an inevitable aspect of human action." That quotation comes from a critical review of Sissela Bok's book on lying—the only adverse notice of the book that I have come upon—by David Bazelon.[24]

Bazelon reproaches Bok for exclusive attention to outright and explicit lying—which he regards as a "disastrous" limitation—and he heartily dislikes her recommended maxim that "no one should lie except on the rarest occasions." Bazelon thinks such concentration on the impracticable "best" will "unavoidably assure that lesser forms of deception will be used more frequently, accompanied by a greater sense of virtue." He himself (a "lumper" rather than a "splitter") wants attention to be paid to "the phenomenon of deception in full, rich context." Lying should always be regarded, he claims, as "the major form of

modern power," short of actual violence, and hence "an actual or potential aspect of all action involving two or more people." Convinced as Bazelon is, then, of the ubiquity and inevitability of lying as an exercise of power over the deceived, he ends with the following remark and conclusion: "Once the ubiquity of deception is appreciated, and also its central relation to power, the need of power to achieve one's purpose, the issue clearly becomes—*which lie to tell, when, to whom, and for how long.*"

Since Bazelon is operating with so broad and comprehensive a definition of lying as a mode of deception, his argument, if sound, ought to apply equally well to humbug. No doubt Bazelon would say: The only sensible questions are *how to humbug, when, to whom, and for how long*. This conclusion might appeal to anybody impressed by the prevalence of humbug; but before succumbing to its attractive cynicism, we ought to notice that the implied argument is invalid.

Think of parallel reflections about, say, dirt or infection: to the effect that since perfect cleanliness is a chimera, the only questions sufficiently practical to be worth considering are how much dirt to tolerate and in which circumstances. As a correction to an impracticable perfectionism, the position is acceptable. But surely indefensible is the suggestion that there is no harm in dirt, or that nothing needs to be done about it. It would be a gross example of the fallacy, for which there is no special label, of arguing from ideal impracticability to permissive laissez-faire.

We need not be committed to the utopian project of a society completely free from humbug in order to hold, as I do, that the evil should be combated to the best of our ability. Humbug may well be as ineradicable as degeneration and death, but that is a poor reason for indifference or complacency.

Coping with Humbug

I shall now try to keep my original promise to say something useful about how to cope with humbug. I hope you will agree that while humbug has the short-term advantages of devious

hypocrisy over naked felony, it is indeed an insidious and detestable evil.

In order to cope with humbug we need, of course, to be sensitive to its occurrence. One soon develops a nose for it. Indeed, there is a danger of becoming tiresomely overzealous in its exposure.

For short-term remedies, I recommend first the ploy that I earlier called the Shavian probe—the deliberately naive and rather impolite challenges expressed by the questions "Do you *really* believe that?" and "Do you really *mean* that?" (If the answer is yes, one might then use one of G. E. Moore's favorite expressions: "How extraordinary!") A more elaborate maneuver is to take the humbugging formula literally in order to reveal its latent exaggerations and absurdities. Thus if somebody solemnly delivers the shoddy bit of proverbial wisdom that "the exception proves the rule," one might trump it by saying, "Quite so. The more exceptions, the better the rule!" But a more useful therapy is to translate humbug into plain and clear English. Such translation is especially effective in coping with learned humbug. (The abuse involved is a kind of converse of the emperor's clothes—too many clothes and no emperor.) Strongly to be recommended also are humor, parody, and satire. (The glorious response, for instance, of the philosopher Samuel Alexander, in his deaf old age, shaking his ear trumpet with laughter on being introduced to a Harvard professor: "I must be getting *very* deaf—I thought you said he was a professor of business ethics!") Required reading might well include Flaubert's dictionary of received opinions,[25] Frank Sullivan's interviews with Mr. Arbuthnot, the cliché expert,[26] and some of the splendid parodies of Russell Baker.[27]

Fortunately, literature provides wonderful portraits and caricatures of accomplished humbugs—Dickens' Pecksniff, Uriah Heep, Podsnap, and many more; Molière's Tartuffe and Alceste; and the confidence men of Herman Melville and Thomas Mann. And much can be learned from a long line of exemplary antihumbuggers, among whom I include Dr. Johnson, Samuel Butler, Sydney Smith, Anton Chekhov, George Orwell, Vladimer Nabokov, E. B. White, and Adlai Stevenson.

Coda

It would be satisfying to be able to end with some concise and accurate definition. The best I can now provide is the following formula:

> HUMBUG: deceptive misrepresentation, short of lying, especially by pretentious word or deed, of somebody's own thoughts, feelings, or attitudes.

This definition covers only first-degree humbug. For second-degree humbug, produced by a self-deluded speaker or thinker, the unsatisfactory reference to thoughts and so on would need to be replaced by something like "thoughts . . . that might be revealed by candid and rational self-examination." I must leave the problems concealed in this all too brief formulation for another occasion.

Plenty of work remains to be done. For, as Miss Mowcher liked to say: "What a world of gammon and spinnage it is!" and "What a refreshing set of humbugs we are, to be sure!"[28]

Part IV

Some Puzzles

7/

The Rationality of Voting

Introduction

A central problem for political theory is to determine how far private interest can be reconciled with public good. It is important to understand what rational motives an individual citizen can have for contributing, even at the sacrifice of immediate self-interest, to indivisible services that only a government or a state can provide. Since the time of Hobbes, it has been widely held that the requisite social cooperation can be ensured only by force and the threat of force. Without questioning this general proposition, I shall try to show that a case can sometimes be made for *free* and rational individual contributions to a public good.

In the following discussion, I confine myself to one facet of these problems, by considering whether and to what degree an individual voter in a large democratic election can have good reasons for voting, from the standpoint of immediate self-interest. That there are such reasons is not immediately obvious. Indeed, a number of plausible arguments can be produced, as will be seen, for the contention that short-term self-interest cannot rationally justify the expense involved in voting in large elections. I shall try to demonstrate the unsoundness of these arguments and will supply some *self-interested* reasons for voting, even in large elections. (Many of the issues that arise are relevant to the broader questions discussed by economists and other social scientists under the heading of the rationality of "free riding.")

The First Argument against Voting

Before setting off for the polling station, a prospective voter, say Joe Doakes, might stop to consider whether there was any point in his taking the trouble to vote. For he might think that his vote would make no difference to the outcome. Now, since the act of voting would cost him something, in time, trouble, or expense, rationality would seem to dictate that he stay at home.

This line of thought in favor of absention from voting might be countered by the following reply. A single vote can make *some* difference, because that vote might break a tie, or produce one. (To simplify the following discussion, I shall simply ignore the second possibility.) Suppose the chance of a tie, if Joe Doakes does not vote, is p; suppose the value to him of the election of his preferred candidate is V; and finally suppose that the cost to him of voting is k. Then the net expected benefit, b, satisfies the equation

$$b = pV - k$$

For a given cost, k, whether the benefit b will be positive or negative obviously depends on the size of the *product* of the probability p and the value V. Thus p might well be very small when a large number of persons can be expected to vote, but multiplying p by a sufficiently high value, V, may generate an expected value that is higher than the cost. Indeed, sufficiently high values of V might yield benefits of *any* magnitude. Thus the voter may or may not have a rational motive for voting, depending on circumstances.

Secondary Reasons for Voting

A number of writers have thought that the probability p of a tie in a large election is so small that the individual voter will have after all no good reason to vote. Some of these writers then rely on such secondary benefits as contributing to the size of the vote,

fulfilling one's civic duty, satisfying one's conscience, or enjoying the pleasures of manifesting solidarity with like-minded partisans. Such considerations may indeed contribute enough additional (expected) value to make *b* turn out to be positive.

I shall not consider such secondary reasons for voting, however, important as they are in actual elections. Some of these secondary reasons depend for their value on the rationality of the primary purpose of contributing toward the election of the preferred candidate. Considerations that tend to show that this primary purpose cannot be achieved by an individual vote would impugn the value of satisfying one's conscience or enjoying the ritual of voting. Somebody who still thinks it makes sense to vote, for these supplementary reasons, even though he thinks that he is directly achieving nothing by his own vote, is like somebody who goes to church to enjoy the company of other members of the congregation, even though he thinks prayer ineffective and pointless. (We can eliminate some of the secondary values by supposing that the size of the vote is unknown and that individual votes are delivered secretly—say, by a mail ballot.)

Is Power to Break a Tie the Decisive Consideration?

The first reply to Joe Doakes' misgivings does establish conclusively that in his original train of thought he erred in supposing that there could be *no* relevant primary benefit in voting. The reply, however, shares a troublesome assumption with the original argument in taking it for granted that the only relevant contribution of a single vote could be its power to break a tie.

We might analyze the contribution of Joe Doakes' vote in the following way. If there should be a tie that could be broken by his vote, his proposed action would be both sufficient and necessary to achieve his purpose. Let us say that the single vote would then be *conditionally necessary and sufficient* for winning the election. Of course, it would be a mistake for Joe Doakes to think that *he* elected the candidate—even if he were the very last person to cast a vote. For I have been assuming that the election is a democratic

one, in which each vote counts as much as but no more than any other. And though Joe Doakes might pride himself on having provided a conditionally necessary and sufficient condition for his candidate to be elected, after voting in an election decided by a single vote, exactly the same claim could properly be made, and for the same reason, by each person who voted on the winning side.

Cases in Which Ties Are Ruled Out

But what are we to say in cases, such as those of presidential elections, in which a decision by one vote is not merely improbable but in some sense impossible? A case, that is to say, in which the probability of a tie can be known in advance to be zero? Then it would seem that the term pV in the calculation of b is reduced to zero, and the primary value of b (ignoring as before what I have called the secondary values) must equal the cost, or negative benefit, $-k$.

There are two cases to be considered: the first, one in which the voters know in advance that the election cannot be won by a single vote because the regulations covering the election forbid such a result; the second, a case in which there are supposed to be conclusive empirical grounds for knowing that such a result will in fact not occur.

Suppose there is a regulation that the margin for a victory must be at least 100 votes. Then, presumably, an election in which one candidate has a favorable margin of 99 votes will be counted as indecisive, with some provision for a runoff. But then our previous argument about the possibility of the individual voter's action being conditionally necessary and sufficient will still apply with obvious modifications. Joe Doakes can say to himself that the margin, apart from his contribution, would have been 99, but not 100. So he would have truly made a decisive contribution by converting an indecisive outcome into a victory for his side. (And if the other side would, but for his own vote, have had a margin of 100 votes, his vote would convert their margin to 99

and so have prevented the other side from winning on this occasion.)

The Second Argument against Voting

The case of a supposed empirical impossibility of casting a tie-breaking vote (apparently quite realistic for large elections) is more troublesome. Imagine Joe Doakes to argue as follows:

1. I know, on the basis of adequate empirical evidence of the outcome of similar elections in the past, that this election will not be decided by a single vote.

2. Therefore, I know in advance that my vote cannot break a tie.

3. Therefore, my vote will be ineffective—will make no difference: I do not even have a chance of acting in a way that would be conditionally necessary to elect the person I want to win.

4. Accordingly, the expected benefit, as previously calculated, must be negative (given that I must incur some cost in voting).

5. Therefore, it is irrational for me to vote.

I shall call the first premise, that the election will not in fact be decided by a single vote, the *disparity premise* (*DP,* for short).

This new argument has the disturbing feature that if enough of the other voters considered it, regarded it as sound, and acted on its conclusion by abstaining from voting, the disparity premise would be rendered false. So the argument is one whose soundness would have a tendency to make one of its premises false *(DP)* and hence to make it an unsound argument. (This makes it akin to the much-discussed phenomenon of the so-called self-refuting prediction.) So if the argument is to be considered as soundly derived from true premises, it must be assumed that enough other voters will *not* consider the argument, or if they do consider it will not see that it is sound, or if they do take it to be sound will still vote in sufficient numbers to make *DP* true (possibly because they are swayed by the secondary motives for voting). So Joe Doakes can still persist in viewing his argument as sound, although he recognizes that his abstention has to be justi-

fied by the nonabstention of others. (Colloquially: "If they vote, the more fools they! But given that they *will* be foolish, there's good reason for me *not* to vote.")

The Doakesian argument (as I shall call it) for not voting in an unbalanced election (one for which *DP* can be known to be true in advance of voting) can be dramatized and rendered more persuasive than it otherwise might be in the following way. Suppose Joe Doakes is so late in deciding whether to vote that he knows that if he does vote he will be the last person to do so. Then he might reasonably say to himself that the election has already been decided, one way or the other. It is then, as it were, too late for his vote to make any difference. So the rational thing in such a situation, it seems, would be not to expend energy in doing something pointless. On the other hand, I have assumed throughout that we are dealing with a fair democratic election, in which a vote has the same effect or lack of effect no matter when it is cast. So, if the vote were to make no difference if it were the very last vote, it ought also to make no difference *whenever* it was delivered.

Rebutting the Doakesian Argument

There must be *something* wrong about the Doakesian argument. This can be seen as follows. The total vote of those on the winning side of the election certainly *does* accomplish something, by electing the successful candidate. Suppose n votes were cast for him. Then one might say that n voters between them exerted a certain political power (call it P) manifested in the election of the preferred candidate. But this effective power can be the resultant only of the power exerted equally by each of the n individual voters who between them produced the resultant power, P. If each of them literally contributed nothing to the outcome, then P would be nothing also: a sum or resultant of zeroes can be no more than zero. So let us think again of how we can understand the contribution that Doakes makes, even in an election that

cannot be decided by a single vote—an election for which the disparity premise turns out to be true.

A Second Look at the Effect of a Single Vote

We can profitably invoke probability considerations once more. Surely one consequence of Doakes' voting is to increase the probability that his side will win. Suppose that when the election starts, p_1 is the probability that Doakes' preferred candidate will win if Doakes does not vote and p_2 the probability that he will win if Doakes does vote. Then p_2 is greater than p_1, and the difference, q, or $p_2 - p_1$, will be positive. So our previous equation for the benefit to be produced by Doakes' action can be replaced by

$$b = qV - k$$

which, as before, may be positive.

We may notice that this argument works independently of knowledge of what the values of p_1 and p_2 are, or of how they ought theoretically to be calculated. If I am playing a game in which a number of heads and tails are produced without my knowledge, with the understanding that I get a payoff only if the number of heads exceeds the number of tails, then it is rational for me to avail myself of a chance of adding another head to the pot, even if I am completely ignorant of the other relevant details, because I thereby increase the odds in favor of my winning.

We have now come to see how an individual vote contributes to the outcome of an election, even when the result is settled by a substantial margin. According to our analysis, each vote augments the probability of the desired result, and in this way provides the voter with enough rational incentive to offset the cost of voting—when the incremental increase in the expected value of the outcome is greater than the cost (in time, energy, expense) of voting. Each such increment in value may be so small as to

make the individual contribution seem trivial or negligible. But triviality is not the same thing as nullity: the sum total of all these small contributions may constitute enough electoral power to decide the election. And no contribution to such a result, however small, is literally negligible.

Puzzles Connected with Overdetermination

The foregoing analysis seems correct, but needs reconsideration, since further puzzles, almost as troublesome as those we have resolved, may still arise. In an unbalanced election, one that is predictably to be decided by a substantial number of votes on the winning side, we have what might be called *overdetermination of the outcome.* For if m votes are in fact cast on the winning side, and n on the losing side, any number of votes from m to $n + 1$ would have been enough to achieve the same result. So a prospective voter, uncertain whether it is rationally worth his while to vote, may still remain unconvinced by the considerations I have offered.

Consider the following analogy. Suppose I see three men about to lift a heavy weight by jointly hauling on three ropes, and I know that their combined effort will be enough to raise that weight. Suppose also that the outcome matters to me, would be a good thing from my standpoint. If I am considering whether to take hold of a fourth rope and help in the task, I can at once see that the work I would thereby do would not be wasted, since the four of us together would then be raising the weight, all taking an equal share in the labor. But if the other three, as I suppose, *would* have raised the weight by themselves, my prospective contribution might still be regarded as unnecessary. Ignoring supplementary considerations about the damage that might be done to my reputation by selfishly standing by as a free rider, it would seem, after all, that the rational thing to do, from the standpoint of egoistic self-interest, would be not to bear the cost of hauling the weight, but to leave the job to the others. A parallel argument

would be in point in any case of overdetermination of the desired outcome.

I think the analogy with the case of the overdetermined election is strong. Although the hesitant voter would, by voting, make a genuine contribution to the outcome (as we have seen), the election's outcome is supposed to be predictably overdetermined. So if the prospective voter knows this and is willing to endure any odium or self-reproach that may attach to behaving as a free rider, it would seem, after all, rational for him not to vote.

It is no answer, I think, to point out, correctly, that the collective "work" to be done by the voters is not the tangible one of producing a change in the physical or political world (raising a weight, electing a candidate) but rather that of progressively augmenting an initial *probability*. (Even the voters on the losing side may be said to do this job of increasing the probability mass, as it were, and should not regard their efforts as wasted, even if their side loses.) For it remains true that in an overdetermined election, all the votes on one side less one would still, according to our hypothesis, render the accumulated probability sufficient to achieve the desired end.

We may notice that this last argument against an individual's voting in an unequal election generalizes only if a certain precaution is observed. It would be fallacious to conclude, if one regarded the argument against voting as correct, that *each* rational voter should not vote. For that would lead to obvious paradox, as before. The proper generalizing conclusion should be that each voter has rational grounds for not voting *provided enough of the others on his own side of the issue do vote.* Thus each individual voter who rationally refrains from voting makes it more risky for the others of his party to refrain. The recommended policy of abstention is rational if and only if a sufficient number of like-minded voters do *not* act on that policy.

I am uncomfortable with the conclusion, now seemingly forced upon me, that free riding, whether in the election case or in others where a sharable good is to be achieved by collective action, is not rational from the standpoint of egoistic self-in-

terest, moral considerations about profiting from the actions of others set aside. I still have the feeling, though possibly on mistaken grounds, that it *is* rational to vote, even in an election in which one has good reason to expect that the outcome will be settled by a substantial margin—that is, good reasons to think that the disparity premise is true.

Uncertainty about the Disparity Premise

Perhaps the best way to escape this unpalatable conclusion is to reexamine the assumption that available empirical evidence is strong enough for the hesitant voter to *know* that the disparity premise is true. If this *is* known, it does begin to look as if it is rational not to vote. To revert to the case of the last person to vote: if, when his turn comes, he *knows* that the election has already been settled, it really would be pointless for him to take the trouble. One might say that by the time his turn comes, the point of adding to the chances of electing his candidate has become irrelevant, since by that time it is a *certainty* that his candidate has either won or lost the election. (It is irrational to bet on a *known* certainty.) But in realistic situations, the assumption that the election is known to be of the sort that makes the disparity premise true may well be too strong. Given that the stakes are sufficiently high and that there is *any* uncertainty about the behavior of other voters and the nature of the outcome, it may after all be rational to take the chance of contributing to what *may* turn out to be overdetermination.

Imagine that a surgeon is conducting a critical operation in the course of which his patient may die and has already taken all reasonable precautions against that outcome. Suppose now that a colleague suggests some further measure that may well be unnecessary but *might* help to save the patient's life, if some unlikely contingency were to be realized. Then, if the cost of that additional safety measure were not excessive, I should think it would be altogether reasonable to overdetermine the precautions. It would be reasonable, I think, to argue that on balance nothing

would be lost by extra precautions, and something might be gained. (Similarly, somebody who thought that the chance that being strapped into a car would make any difference to his survival in a crash was very small indeed might still feel that since his life might be at stake, it would be rational to suffer the annoyance of being strapped in.)

To the question "Why add further precautions when enough have already been taken?" the answer should be "Because we can't be *sure* that they are enough." Similarly, the answer to "Why vote in a safe election?" is "Because we can't be *sure* that it is safe."

If I were ever to have a vote in a presidential election in which a Nixon *redivivus* were a candidate and I had reasonable grounds for being very sure that he would be defeated without my help, I might for all that still feel that the consequences of being wrong would be so monstrous that I and every like-minded voter ought to face with equanimity the prospect of contributing to an overdetermined outcome. In such a case, I cannot believe that I would retroactively have cause for rational regret at contributing to a one-sided result. In general, I believe that contribution to a probably overdetermined desired outcome should count as rational except in cases where there is ironclad assurance *in advance* that the outcome *will* be overdetermined. There might be rational if selfish reasons—moral considerations aside—for not helping others to raise a weight from a wounded person if one really did *know* that the task would be accomplished by others. But if another's life were at stake, one might hesitate to assert that one *knew* one's assistance would be pointless.

Another Argument for Voting

Suppose that I favor a certain candidate in an "unbalanced" election and am wondering whether it is rational to take the trouble to vote for him. Let *B* be the set or bloc of like-minded electors, defined as those registered voters who would vote for my candidate if they voted at all. My primary purpose of electing

my candidate can be achieved only through the *collective* or aggregative electoral power of *B*. The best that *I* can achieve is to strengthen the power of *B*, in the hope that *B's* power will in fact prove sufficient to produce the desired outcome.

The *interest* of *B* (if I may speak in this style for the present) is to further the election of the very candidate whom I favor. That, however, is precisely my own interest. There is, accordingly, a *perfect coincidence* between my individual interest and the interest of *B*.

Consider the electoral task now from the standpoint of *B*. Clearly, the rational way to achieve *its* goal (of doing everything possible to elect the candidate) is to take all steps in its power to get a maximum turnout of the members of *B*.* (*B* cannot know that it will be successful, although it knows that the outcome, whatever it turns out to be, will be "unbalanced," not settled by a single vote.) Hence it is in *B's* interest that as many members of B as possible shall vote, exceptions being permissible only in cases of special hardship. It would be plainly contrary to *B's* interest to adopt the rule that every member of *B* should vote *unless he or she thinks otherwise*. For this would promote abstentions, contrary to the interest of producing a maximal turnout. But *my* interest is, as we have seen, precisely the same as *B's* interest: so I must rationally agree with the rule and hence apply it to myself.

Objection. It might be retorted that *B* does not have the sociological reality of a corporation, labor union, or some other constituted body with power to deliver a block vote, but is only a notional unity, merely assembled in thought for the purpose of the argument. Hence a rule that would be binding on the members of a corporate body who had contracted or promised to serve the general good of that body has not been shown to be binding on me as a member of the set *B*.

Answer. I can grant that there is something fictitious in speaking of "*B's* interests" and of "its" requiring such-and-such to further "its interest." Yet I also think that the language I have

*It follows that members of *B* should also take other steps, such as canvassing or providing transport, that are likely to increase the turnout of *B*'s members.

used can be demystified, without detriment to the argument. Simply replace "*B's* interest" by "the common interest of all the members of *B*"; replace "*B's* contribution" by "the resultant effect of the votes of all the voting members of *B*"; and so on. It still remains true, so far as I can see, that *my* interest is correctly identified as an interest, given uncertainty of the outcome, in having as many members of *B* as possible vote. And the same conclusion follows.

In order to appreciate the force of the argument, one needs to resist the temptation of seeing *my* project as an individualistic one that concerns only myself and to recognize the obvious point that the task is a collective one. Impartial consideration of the nature of that collective task and the demands consequently placed on all the members of the relevant collectivity suffices to show that my own interest, mediated by the common interest of the collectivity, requires me to vote, if I can do so without exceptional hardship.

Formally, the argument might look approximately as follows:

1. I want *X* to be elected.
2. The best chance of achieving X's election is for as many voters as possible who would vote for *X* to vote.
3. Hence, I want every such voter to vote if possible.
4. But I am such a voter (and it is possible for me to vote).
5. Hence, I should vote.

(This argument could also be presented in terms of interests rather than wants.) There is an assumption made in passing from 2 to 3 which, if fully elaborated, would produce a cost-benefit equation like the one I have previously used. So if the argument as now presented is sound, it will not yet follow that I have a categorical reason for voting, all things considered, since the cost may be too high in my special situation. But it will follow, if I am not mistaken, that I *can* have a rational primary reason for voting, which is all that I have been trying to establish.

8/

Newcomb's Problem Demystified

Hard cases make bad law.

A curious feature of the fanciful conundrum known as Newcomb's Problem[1] is the failure of a decade of copious and sophisticated discussion to produce an agreed solution. I shall provide, though with little hope of converting dissenters, what seems to me a satisfactory analysis.

The Original Version

Imagine, if you can, the following choice in what I shall call a Newcomb situation:

You are offered the contents of one or two boxes. The first is open and is seen to contain $1,000; the other is closed and is said to contain either $1 million or nothing. You may take whatever the closed box contains (call this the *single* choice) or, if you prefer, the contents of both boxes (the *double* choice).

You are supposed to have the following information:

1. The setup has been arranged by some super-predictor *(SP)* who has always correctly predicted your choice in similar situations and has nearly always correctly predicted the choices of other persons who resemble you. You therefore confidently expect that *SP* has correctly predicted your choice on this occasion.

2. *SP* has already placed $1 million in the closed box, if he predicted a single choice. However, if he expected you to make the double choice, or to evade a reasoned choice by tossing a coin

(or randomizing in some other way), he has left the closed box empty.

Given such a situation, how should you choose?

Conflicting Solutions

There are two conflicting and "equally plausible and highly intuitive" responses to the extraordinary situation just described.[2] (According to some commentators, a radical antinomy of rational choice is to blame.)

A perplexed chooser in the imagined predicament might plausibly argue as follows: "By taking the open box, I am certain to get $1,000. As far as the closed box is concerned, the predictor has either already inserted $1 million or has left it empty. In either case, by choosing to take both boxes, I have nothing to lose and possibly something to gain (an extra $1 million!). I should therefore take both boxes."

This kind of argument—colloquially formulated as "nothing to lose and possibly something to gain, no matter what the case may be"—is known in game theory as a (weakly) dominant strategy. It has a powerful appeal to common sense.

I shall call this line of thought the *conservative argument,* because it recommends a safe choice, supposedly secure against all contingencies.

On the other hand, our imagined chooser might well soliloquize as follows: "I have enormous confidence in the predictor's power to anticipate my choice and therefore to have loaded the closed box, or to have left it empty, as the case may be. So I ought to be willing to bet at high odds—say 9 to 1—that the closed box contains $1 million if I take it alone. The position, then, is that the expected value of a gamble on the closed box is at least $900,000. It would be foolish to renounce such a profitable option for the sake of getting $1,000 for certain. I should therefore take only the closed box."

Unlike the conservative argument, this *adventurous argument* supports a risky choice. It is of a type that strongly recommends

itself to common sense if the expected gain is sufficiently high and the stake sufficiently low.

One of the two arguments must be wrong: some preliminary analysis of the original Newcomb situation will be helpful in deciding which choice to endorse.

Some Features of the Newcomb Situation

The piquant embellishments of this story are fantastic enough to distort our normal intuitions. For the present, I shall suspend disbelief in the fiction of a super-predictor who has the uncanny power to foresee by ultimate choice. It is worth emphasizing, however, that nothing turns on the character of the predictor, provided he does have the imputed powers: we can think of *SP* as a visitor from outer space, as a human being endowed with second sight, or as some inanimate computer program.

I shall not be in a Newcomb situation unless *SP,* whatever he or it may be, arranges the setup independently of my ultimate choice: trickery or chicanery *after* I choose are excluded. That is why I stipulated that *SP* has already, and irrevocably, made his arrangements before I choose.

If *SP* were to be substantially inaccurate in prediction, the adventurous argument for taking only the closed box would be weakened. But the alternative conservative argument would be unaffected. There will be no clash between two plausible solutions unless *SP* is as accurate as alleged; that is, unless he or it does have the imputed predictive power.

Moreover, the chooser is not entitled to make the adventurous choice unless he believes that the *SP* is not only accurate but veracious. (An accurate predictor might well have a strong motive to misrepresent his prediction and hence to cheat before the choice occurs.)

Finally, the foregoing assumptions about *SP's* accuracy and veracity must be based on *justified* beliefs: the chooser must have good reasons for trusting all the details of the presented description of his situation. Much depends on the credibility of that

description: the bizarre nature of a Newcomb situation is ample reason for suspecting deception. (When I see a magician saw his assistant in half, there is much that I can and should doubt.) Only if the chooser has good reasons for believing in the existence of a super-predictor who has honestly and competently behaved as alleged is the adventurous solution attractive.

A Demystified Version

An intelligent but literal-minded child, listening to a fairy tale that begins: "Once upon a time there was a wonderful magician who stopped passing travelers and made them choose between boxes" and so forth, as in our original version, might well ask: "What was he up to? Why did he offer so much money for nothing? Was it all a trick? How did he know what you were going to choose?" Well, in Fairyland anything may happen, but in the real world such questions deserve to be answered. Let us therefore consider how the Newcomb story might look when purged of some of its fantastic embellishments.

Imagine that while visiting some up-to-date museum of science and technology you encounter the following notice:

AN INVITATION
Play an interesting game
and help the Museum

> Would you like to risk a dollar in order to play a simple and amusing game in which you might win as much as $10? Most players in fact end by losing 50 cents, which they think worth the fun of participating. All proceeds are assigned to the Museum's Development Fund.

Anybody who pays the $1 entry fee is directed to a booth, in which he is asked to answer some simple questions about choice and decision (including some designed to show how strongly he relies on dominance). He is then handed the following supplementary instructions:

> In a few minutes you will have a chance to choose a closed box and, in addition, an open one that can be seen to contain 50 cents. If you take only the closed box, you will almost certainly find $10 inside, which you can keep. If instead you choose both boxes, you will almost certainly find nothing in the closed box and so you will get only the 50 cents.
>
> The boxes have already been loaded and will not be tampered with after you have chosen. We should warn you, however, that on the basis of your answers to our questionnaire, and our wide experience with others who have played the game, we have been able to predict your ultimate decision with high accuracy and have loaded the boxes accordingly; that is, to conform with the above statements of your expected losses and gains.
>
> After you have left this booth, you will receive an explanation of how we use the answers to the questionnaire in loading the boxes.
>
> Now, please take your time and CHOOSE.

This version is free of some of the meretricious details of the original version.

The museum game, as I shall call it, differs from the original version in the following respects:

1. The museum's spokesman does not claim *infallible* predictive power (note the presence in the notice of the phrase "you will *almost certainly* . . .").

2. The chooser's experience of answering the questionnaire lends some plausibility to the imputed predictability of the choice.

3. Reference to the consequences of random choice is absent.

4. Only modest and hence credible gains and losses are promised.

5. The museum's alleged interest in raising support funds is a reasonable explanation of the game's existence.

6. The levied entrance fee of $1 can reasonably motivate the chooser to play responsibly.

Altogether, then, this revised version has better claims than the original one to be regarded as posing a realistic choice situation.

Coping with the Museum Game

If I were faced with the museum game, what ought I to do?

I could take a skeptical attitude toward the instructions, thinking that there must be some catch in the pretentious talk about predicting my final choice. (But if I reject part of the notice as deceitful, why should I accept the rest?) Or I could respond trustingly, assuming that in some mysterious way my choice will have been anticipated, as claimed.

On the first approach, I shall regard talk about the probabilities of reward as irrelevant or misleading. For I may reasonably think that the point of the setup (possibly an experiment on the psychology of choice?) is merely to test the strength of my commitment to the dominance principle in a case with seductive irrelevancies. But if I reject the information about probabilities as deceptive, I am left with a blind choice, and can do no better than use the conservative argument.

On the second, trusting approach, I shall be embarrassed by not knowing how certain the "almost certain" consequences of my choice may be. I may plausibly argue as follows: If the museum hopes for long-run profit, it may reasonably expect that for each person who chooses the closed box alone and gets $10 (with a net loss to the museum of $9), at least 18 will choose both boxes and get only 50 cents each (with a net gain to the museum of $9). So the museum may well have arranged that the chance of my getting $10 if I choose the closed box alone is less than one in nineteen. Assuming that the "game" is fair or only slightly biased in the museum's favor, there is then really little to choose between getting 50 cents for sure (and taking a loss of 50 cents) and gambling on a one-nineteenth chance of winning $10. Of course, this train of thought treats the museum's notice as deceptive—as it may well be if the concealed purpose of the game is really to collect data on the vagaries of human choice or merely to entertain its visitors. But how can I tell what the museum's purpose really is? And how can I choose if I don't know?

The case here imagined involves too many unknowns for any decisive judgment in favor of the adventurous choice to be at-

tractive. The chooser must cope with as many uncertainties as a field commander ignorant of his enemy's disposition and intentions. Given the small stakes involved in the museum game and the good purpose to be possibly achieved by playing, it would not be unreasonable to gamble on the uncertain prize, partly for the fun of the gambling and partly out of curiosity. (One might compete several times in order to see how long one would have to wait to recoup the entrance fee.) But the more realistic a revised version becomes, the more its uncertainties erode the attraction of an adventurous choice.

The previously highlighted features of my imagined museum game would be paralleled in any realistic choice situation that might look Newcombish. A rational chooser so situated ought to be attracted by the risky gamble to the degree that he could justifiably treat his instructions as veridical and could therefore confidently believe in the prior prediction of his own choice (call this the *prediction hypothesis,* or *PH*). In a limiting case, in which he thought he knew for certain that *PH* was true, adventurous choice might well seem overwhelmingly attractive. But there are normally such strong objections to the truth of *PH*—as I shall argue below—as to render its acceptance highly questionable.

An adventurous choice in a supposedly Newcombish situation would, accordingly, be very risky at best, since alternative and initially more credible hypotheses would always be available. If you think that the presented description of your situation may be deceptive—because you are really the subject of some investigation of human credulity, for example, or are the butt of a practical joke—you should, as a rational chooser, be correspondingly skeptical of the truth of *PH,* given what we know about the variability of human choice and the difficulty of anticipating its outcome. (Think of jury trials, or the betting behavior of stock market investors!) To rely on the supposed truth of *PH* as a basis for rational choice in any situation that might arise outside science fiction would be as reprehensible as believing in UFOs, the sagacity of the notorious talking horse of Ebersfeldt, or, to take a contemporary example, the supposed prowess of Uri Geller.[3]

Conditions for a Credible Version

By a Newcombish choice situation I shall understand one that resembles the situation in the museum game in the following respects:

Somebody, say Eligo, is supplied with a narrative of the following form:

1. Eligo is to have a single choice between two alternatives, either *C* or *C*-together-with-*O*.

2. A highly accurate but not infallible predictive device, *SP*, has already predicted whether Eligo will make a reasoned choice or a randomized one and what that choice will be.

3. For a predicted reasoned choice, Eligo will receive rewards alpha or beta, the former more valuable than the latter, according to the scheme presented in our original version.

4. An anticipated randomized choice has been treated as if Eligo had chosen both *C* and *O*.

5. Some evidence is presented in support of *SP's* alleged predictive powers.

If Eligo is to take the narrative seriously, he must assign some plausible motivation to the narrator, weigh the evidence for *SP's* existence and extraordinary powers, and decide whether the narrative deserves to be believed. I have suggested some plausible motivation in the last section, but the issues of evidential support and, above all, the narrative's veracity need further consideration.

The following evidence might be available to Eligo: he sits before a computer, allegedly programmed to make highly reliable predictions of his choice in the Newcombish situation in question, and is allowed any number of trial runs. Specifically: in each trial, he decides what his choice in the Newcombish situation is going to be, writes his decision on a piece of paper that he hides in his hand, and then signals the choice to the computer. The computer then displays the gain he would have received. Suppose that during, say, 100 trials in which Eligo sometimes chooses *C* and sometimes *C-plus-O*, the computer always signals $1,000 in the first case and $1 million in the second. (Finally, he

will be allowed *one and only one* real try.) Should that suffice to convince Eligo of *SP's* power? Such supposedly strongly counterfactual evidence would be extraordinary enough to justify caution. For a computer capable of performing as supposed, even to the extent of anticipating random choices, might well be programmed precisely to thwart the last "real" choice. (That suspicion would be reinforced by the thought that the narrator's dominant strategy, if his motive dictated minimizing the payoff, would be *never* to reward the choice of *C!*) So we are back to the question of the narrator's veracity.

My chief contention is that in any realistic Newcombish situation, it is more rational to view the fantastic narrative as deceitful than to believe it.

Can Human Choice Be Predicted?

I shall now reconsider whether I might have reasonable grounds in a Newcombish situation—say, the museum game, in order to fix our ideas—for supposing my choice to have been already predicted (that is, for supposing *PH* to be true in such a case).

Reasons for thinking one's choice had in fact been predicted might be based on the supposed existence of (1) behavioral evidence, (2) what might be called interpretive evidence, or (3) a mixture of both.

By *behavioral* evidence about how a given person will play the museum game, I mean such generalizations as "Blondes take both boxes," based on sheer statistical correlations between observable features, unsupported by an explanatory theory. For instance: the loading of the museum boxes might be determined—implausibly, to be sure—by the simple induction that players who answer the seventh item of the questionnaire affirmatively do in point of fact choose only the closed box. (Cf. the well-known phenomenon of the "personal equation" in the choice of the last significant digit in a scientific measurement.)

By *interpretive*[4] evidence, I mean knowledge of the chooser's

own reasons for his final decision. (Thus somebody reading the present essay might well rely "interpretively" on my choosing both boxes in the museum game.)

By *mixed* evidence, I mean inductive evidence, strengthened by some supporting theory about the influence of known character or temperament, and bolstered by relevant interpretive evidence. (Thus one might reasonably expect an adventurous chess master such as Tal to find intelligible reasons for making a risky sacrifice, which a more cautious player such as Petrosian would reject.)

Behavioral evidence might be called opaque: it tends to be unreliable for prediction, because we don't understand why the apparent uniformity holds. Interpretive evidence might be called transparent, because we can understand, without necessarily accepting, another's reasons and see the point of his using them. It may be highly reliable—as when one predicts so confidently that one's chess opponent will see a mate in two that one resigns in anticipation.

Could *PH* be true? I believe that it might be, provided that the subject of the prediction had no access to the detailed evidence used by the predictor. With some stretch of the imagination, I can conceive of some unknown "behavioral" trait of mine (myopia? high blood pressure?) giving the box loader a tipoff. But it would normally be much more plausible to surmise that the predictor was relying on intelligible evidence, revealed in my answers to the questionnaire.

Suppose I had good reasons to suspect this, before answering the museum questionnaire. Then it would pay me to misrepresent, if I could, a true belief in the dominance strategy. It seems to me rather unplausible that the other "player" could have *intelligible* evidence for such a bluff and very implausible that he could have relevant behavioral evidence for it. But if I am wrong, then I am once again faced with ignorance and uncertainty, while the case for the dominance strategy remains as strong as ever. (It might sometimes be more advantageous to choose irrationally; but it's no use *pretending* to be doing so when playing against somebody who can see through the pretense.)

Does predictability impugn freedom of choice? I should think

the answer is clearly no. Does the fact that I can predict my chess opponent's mating me in two moves show that he won't be acting freely when he does so? Of course not. I know and he knows that he could have done otherwise—for example, in order to humiliate me by prolonging a hopeless struggle.

The question to be addressed to a putative predictor of my performance at the museum game is not whether he will predict my move this time, but rather whether he *would have* predicted a deliberate deviation by me from a conclusion based on interpretive evidence.[5] I shall believe *this* kind of behavior predictable when somebody can predict when I shall misrepresent my genuine choice and what item I choose from a table of random numbers, when I decide to randomize.

For the reasons offered above, I believe that in some situations, such as choosing a chess move leading to an immediate checkmate, or drawing a conclusion from the premises of an elementary valid syllogism, or adding up two and two, the chooser's choice can be predicted, without any question of his freedom to choose. Such predictability is predicated on its subject's having no prior knowledge of the particular prediction and no legitimate interest in defeating it. These general observations, however, lend little comfort to an advocate of an adventurous solution to a Newcombish situation. For in such a case, the chooser has reasonable grounds to be skeptical about the reported nature of his situation and its expected consequences. He has, therefore, only shaky reasons for believing the predictor to be candid, let alone accurate.

Adventurous versus Conservative: Pro and Con

In the following dialogue, *A* is an adventurer, *C* is conservative. References are mainly to the original version of Newcomb's problem.

A. Don't you agree that if there were such a being, able to predict your choice with almost perfect reliability, it would be rational to gamble on the million-dollar prize?

C. You take the fable more seriously than I can. There are too many ifs involved for comfort. *If* there were such a being or some other super-predictive device, and *if* I had good reason to believe in his powers and his truthfulness, and *if* I could take seriously his motivation in offering me such an inordinate reward for credulity—why, then I might believe almost anything.

A. Why be so prosy? If you are willing to entertain seriously Descartes's fable of a malicious demon, why not free your imagination to consider Newcomb's less preposterous fable, for the sake of seeing what we might learn about our position as would-be rational choosers?

C. I expect as little illumination as a pianist would by wondering how to play, if his instrument were in fact a player piano.

A. Well, I can sympathize with your wish to bring the puzzle down to earth, for instance by replacing the original version by what you call the museum game. But aren't you exaggerating your warranted skepticism about the possibility of a super-predictor? I think we might have strong evidence for his existence.

C. Why do you think so? *I* think it would be foolishly credulous to believe it.

A. Well, suppose you were allowed a number of trial runs. Might they not reasonably convince you that what you stigmatize as credulity would in fact have paid better than your rational skepticism? More generally, might not experience show that it sometimes pays better to be a believer than a skeptic?

C. There you go again, imagining events that beggar belief! If, *per impossibile,* I seemed to be confronted with, say, a computer program such as you postulate, I would emphatically refuse to make the choice that you recommend, if anything serious—say, somebody's life—turned upon it. For such a fantastic computer could be programmed precisely to defeat my final, real choice. (And how could I know that? By further trials?) In any case, I think such examples, however much you vary or embroider them, are really beside the point.

A. I can guess what you are going to say.

C. I am glad you said "guess" rather than "predict." If you had, I would have enjoyed exposing you as a false prophet.
Consider the situation in the first Newcomb version: before I make the final choice, the being (or computer or whatever) has set up the boxes irrevocably. If the back of the closed box were open and I were allowed to take a look before choosing, I would see it to be either empty or else loaded with a million dollars, as the case might be. If the second case holds, then by taking both boxes, I shall get the million and an extra thousand to boot; if the first, there is no harm in taking the two boxes, for I still get the thousand in the open box—all that there is to take. But I know that one or other of the two cases is present if your story is to be believed, so *I don't need to look*. Whatever the true situation may be, I get everything available by taking the two boxes. I have *nothing to lose* by taking both, and nothing to gain by looking first—except perhaps dispelling any anxiety caused by your obdurate refusal to see the light.

A. But you *do* have something to lose—what economists call the "opportunity cost" of the forgone adventurous choice. You throw away the chance of winning a large prize, if there is after all some coupling between your choice and the contents of the closed box. Suppose I, as a confirmed risk taker, were allowed to choose and picked up the million, while you, as a confirmed conservative, tried and got only the pitiful thousand. What would you say then?

C. I'd say that you didn't deserve it! It is no news that poor reasoners are sometimes luckier than rational choosers; but that's no argument for trusting weak arguments.

A. And if I repeated the trials and I constantly won, while you always lost, what then?

C. I shall believe that possible when pigs can fly. If pigs could fly, it might be sensible to trade in one's bicycle.

A. You called me obdurate just now. If we were to trade insults, I might call you myopic. You are so impressed by the apparent impregnability of the dominance principle that you can't attend to anything else. But wouldn't you agree that additional information might make an argument for using the dominance strategy *in the face of ignorance* no longer applicable?

C. Certainly. If I see a shop assistant inadvertently add an extra book to the parcel of books I have bought, appeal to dominance (morality aside) might suggest that I have nothing to lose by walking off with the parcel. But not if I happen to know that a detective is waiting at the exit to catch shoplifters.

A. Well, then, why not concede that, whatever the merits of the dominance argument in conditions of uncertainty, the extra information supplied in the Newcomb situation rationally justifies rejecting a dominance strategy as no longer applicable?

C. I can only repeat myself now. I remain loyal to my conservative strategy, not because it is always applicable, but because in the fantastic situation you imagine, the postulated extra information is too implausible to be believed. There never was a genuine opportunity in the situation you envisage. At the moment of choice, the boxes are what they are and will be what they will be; then why should we be deceived?

A. You are too stubborn to be persuaded. Let's go and play some backgammon.

Summary of Main Findings

1. The original version of Newcomb's Problem strains credulity: (a) by postulating the existence of a super-predictor who has forecast whether the chooser will decide to randomize and, if he does not, whether he will opt for the single box; and (b) by failing to motivate the supposed offer of gratuitous reward.

2. A rational chooser, apparently finding himself in so fantastic a situation, would be well advised to suspect fraud, and to play safe by taking both boxes.

3. In any realistic variation of the Newcomb Problem (such as the museum game), with credibly modest gains and losses in prospect, the narrator's motives will remain questionable.

4. A wise choice in a Newcombish situation calls for weighing the relative likelihood of genuine predictability against that of the narrator's mendacity.

5. Although the conservative choice, based on a dominance

argument, is safe against all known contingencies, it might conceivably be superseded by additional information about the predictability of choice.

6. Powerful general arguments discredit the postulated predictability: the considerations favoring the adventurous choice are too weak to be trusted in real life.

7. A rational chooser in a Newcombish situation should assume the narrator's mendacity and choose conservatively.

Notes

Introduction

1. C. Day Lewis, *The Poetic Image,* Clark Lectures (London, 1947), p. 18.

2. H. L. Mencken, *A Mencken Chrestomathy* (New York, 1949), p. 13. He also compared a working philosopher to a chimpanzee hopelessly chasing fleas: "Both suffer damnably and neither can win" (p. 17).

3. Anybody who thinks it easy to say what "we mean by two" should take a look at Gottlob Frege's splendid little book, *The Foundations of Arithmetic,* trans. J. L. Austin (Oxford, 1950).

4. Lewis Carroll, *Alice in Wonderland,* chap. 8.

5. W. B. Gallie, *Philosophy and the Historical Understanding* (New York, 1964), chap. 8.

6. More about this in Essay II.

7. F. M. Cornford, *From Religion to Philosophy* (London, 1912), p. 138. A celebrated philosopher made the point more tersely when he said: "Metaphysics is the finding of bad reasons for what we believe upon instinct, but to find these reasons is no less an instinct" (F. H. Bradley, *Appearance and Reality,* 2d ed. [Oxford, 1897], p. xii). I would replace "believe upon instinct" with "unshakably believe."

1. *"Why Should I Be Rational?"*

Originally delivered, in part, at an international colloquium in Biel, Switzerland, May 1980. A version substantially different from this one has been published in *Dialectica* 36 (1982). I am much indebted to discussions of earlier drafts at Wolfson College, Oxford, the universities of Lancaster, Norwich, and Sussex, and the London School of Economics, during the fall of 1980.

1. A questionnaire, widely distributed among students and university teachers in England, Canada, and the United States, has revealed extensive disagreement as to whether backing a horse to win and to lose at the same odds, or crying over spilled milk (to mention two examples), should count as rational, irrational, or neither. (The results of this inquiry are as yet unpublished.)

2. *Collected Papers* (Cambridge, Mass., 1931–58), 5:232, 278.

3. P. T. Geach, *Reason and Argument* (Oxford, 1976), p. 1.

4. Peter Singer, *Practical Ethics* (Cambridge, 1979), p. 203.

5. A. J. Ayer, *The Problem of Knowledge,* Penguin ed. (Baltimore, 1956), p. 75.

6. It is tempting, but mistaken, to suppose that a standard cannot itself be tested for accuracy. If we had evidence of a break-in to the chamber containing the standard meter rod in Paris and could see that its end had been removed, we would have excellent reason to believe that the official standard had changed its length.

7. Karl Popper, "Utopia and Violence," in *Conjectures and Refutations* (London, 1963), pp. 356–57.

8. Popper has elsewhere provided far more adequate and persuasive conceptions of the rationality of which he has been a lifelong advocate. See, for instance, Paul Schilpp, ed., *The Philosophy of Karl Popper* (La Salle, Ill., 1974), 2:1085–86.

9. There is a valuable discussion titled "The Infinite Regress" in John Passmore, *Philosophical Reasoning* (London, 1961), chap. 2.

10. Adriaan D. de Groot, *Thought and Choice in Chess* (The Hague, 1965).

11. In a natural habitat, of course. Our knowledge of an extensive municipal effort to poison pigeons might give us an advantage.

12. According to Diogenes Laërtius, the Greek skeptic Pyrrho of Elis, who believed "that custom and convention govern human action," was indifferent to common hazards, "but was kept out of harm's way by his friends who . . . used to follow closely after him"—and he lived to be nearly ninety! (*Lives of Eminent Philosophers,* ed. R. D. Hicks [London, 1925], 2:245, quoted from Arthur W. Burks, *Chance, Cause, Reason* [Chicago, 1977], p. 137).

13. I am using here a technique explained and defended in my "More about Metaphor," *Dialectica* 31 (1977), especially pp. 443–45.

2. *Reasonableness*

Originally published in *Education and the Development of Reason,* ed. R. F. Dearden, P. H. Hirst, and R. S. Peters (London: Routledge & Kegan Paul, 1972); reprinted in my previous collection of essays *Caveats and Critiques* (Ithaca: Cornell University Press, 1975).

1. "The Function of Reason in Morals," *Proceedings of the Aristotelian Society* 39 (1938–39):251.

2. *Men and Society* (London, 1940), p. 53.

3. I use "dianoetic appraisal" as shorthand for "appraisal as reasonable or unreasonable."

4. This term is further explained in my book *Models and Metaphors* (Ithaca, N.Y., 1962), pp. 156–59, where there is a detailed application to the analysis of causation. See also the article "Paradigm-Case Argument" by Keith S. Donnellan in *The Encyclopedia of Philosophy*, ed. Paul Edwards (New York, 1967), 6:39–44. Philosophers sometimes appeal to paradigm cases of the use of some word *W* to show that there can be no doubt that instances of *W* exist. I do not need this type of argument in the present essay, since I take it as obvious that some actions are reasonable and some are not.

5. This conjecture is supported by the familiar sense of "reasonable," otherwise ignored in this essay, to mean "moderate, not excessive" (cf. "a reasonable price"). This may derive from Aristotle's influential doctrine of right action as a "mean" between extremes. Even if the "extremes" of excess are well defined, there is usually plenty of room for argument about what would be neither "too much" nor "too little." According to this view, the reasonable might well have a partially indeterminate nature. See also the remarks in note 8 about combining probabilities with utilities.

6. Is this irrational, rather than unreasonable? I think not. If I badly wanted to speak to the man whose number I was looking up, I might justify myself by saying that he might have a telephone after all—even if I "knew" that he hadn't! But perhaps this kind of case does verge upon the irrational.

7. Ideally, several people who understand the paradigm case procedure should be asked to construct large and varied lists for comparison and discussion. One man's choices may show unconscious bias or inattention to potential diversity. My examples have, in fact, not been chosen altogether at random.

8. There is an extensive technical literature about alternative ways of combining probabilities and "utilities" to calculate most-preferred actions under uncertainty. A disconcerting outcome is the wide range of choice of plausible solutions. It looks as if even in theory and in highly simplified situations, there can be no *determinate* answer to "What is the most reasonable way to act?" Here we have perhaps another illustration of the indefiniteness of the notion of reasonableness, already hinted at earlier.

9. But what if *A* does cost more than *B*? Here we encounter the type of difficulty mentioned in the last note. Is it reasonable to pay one surgeon ten times as much as another in order to raise your chance of surviving an operation from 5/6 to 7/8? It is too facile to reply that it depends on how much you value your life. I believe that the question has no determinate answer.

10. "Some Questions Concerning Validity," in *Essays in Conceptual Analysis*, ed. A. Flew (London, 1956), p. 128.

3. *Scientific Objectivity*

Based on a talk given at Drexel University, Philadelphia, April 1975. A shorter version appeared in *The Bulletin of the Atomic Scientists* 33 (February 1977):55–60. Reprinted by permission of *The Bulletin of the Atomic Scientists,* a magazine of science and public affairs. Copyright © 1977 by the Educational Foundation for Nuclear Science, Chicago, Ill., 60637.

1. Leon Chwistek, *The Limits of Science* (New York, 1949), p. 1.
2. Theodore Roszak, *Where the Wasteland Ends* (Garden City, N.Y., 1972), pp. 179–81, 184, 188.
3. André Lalande, ed., *Vocabulaire de la philosophie,* 5th ed. (Paris, 1947), p. 934.
4. Ernst Mach, *Popular Scientific Lectures* (Chicago, 1943), pp. 214–15.
5. Ibid., p. 217; emphasis in original.
6. J. Ravetz, *Scientific Knowledge and Social Problems* (Oxford, 1971), pp. 20, 21, 28.
7. Ibid., p. 29.
8. D. Runes, ed., *Dictionary of Philosophy* (New York, 1942), p. 217.
9. G. H. Hardy, *A Mathematician's Apology* (Cambridge, Eng., 1940), pp. 63–64.
10. Max Planck, *Scientific Autobiography* (New York, 1949), p. 13.
11. Roszak, *Where the Wasteland Ends,* p. 188; emphasis in original.
12. T. A. Roberts, ed., *Butler's Fifteen Sermons* (London, 1970), p. 71.
13. Jean Piaget, *The Child's Conception of the World* (Totowa, N.J., 1975), p. 42.
14. Jean Piaget, *The Child's Construction of Reality* (New York, 1973), p. 14.
15. Ibid., p. 16.
16. Sigmund Freud, *Introductory Lectures on Psychoanalysis,* Penguin ed. (Harmondsworth, 1974), p. 133.
17. Thomas H. Huxley, *Science and Education* (New York, 1895), p. 45.
18. Ibid., pp. 45, 46.
19. Hudson Hoagland, "Some Comments on Science and Faith," in *Conference on Science, Philosophy, and Faith: Second Symposium,* ed. Lyman Bryson (New York, 1942), p. 35.

4. *Is Scientific Neutrality a Myth?*

An invited address to the annual meeting of the American Association for the Advancement of Science, New York City, January 1975. A shorter version titled "Scientific Neutrality," appeared in *Encounter* 51 (August 1978):56–62.

1. David Hume, *A Treatise of Human Nature,* bk. 3, pt. 1, sec. 1. Those who wish to pursue the philosophical issues further might refer to my essay "The Gap between 'Is' and 'Should,'" chap. 3 of *Margins of Precision* (Ithaca, N.Y.,

1970). A useful collection of related papers is W. D. Hudson, ed., *The Is/Ought Question* (New York, 1969).

2. Jacques Monod, "On the Logical Relationships between Knowledge and Values," in *The Social Impact of Modern Biology,* ed. Watson Fuller (London, 1971), p. 12.

3. Ibid., p. 15.

4. Robert L. Heilbroner, "What Has Posterity Ever Done for Me?" *New York Times Magazine,* January 19, 1975, pp. 14–15.

5. Paul Goodman, "Can Technology be Humane?" in *The Social Responsibility of the Scientist,* ed. Martin Brown (New York, 1971), p. 251. The quotation continues: "It [technology] aims at prudent goals for the commonweal and to provide efficient means for these goods."

6. *Statistical Abstract of the United States* (Washington, D.C., 1981).

7. G. A. Lundberg, R. Bain, and N. Anderson, eds., *Trends in American Sociology* (New York, 1929), pp. 404–5, quoted in Robert K. Merton, *Social Theory and Social Structure* (Glencoe, Ill., 1953), p. 543.

5. *Humaneness*

Based on talks given at Brown University, the Massachusetts Institute of Technology, and the India International Centre, New Delhi. An earlier version appeared in *India International Centre Quarterly* 5 (1979):83–95.

1. J. A. K. Thomson, trans., *The Ethics of Aristotle,* Penguin ed. (Harmondsworth, 1955), pp. 37–38.

2. William Golding, "Fable," in *The Hot Gates* (London, 1965), pp. 87, 89.

3. G. K. Chesterton, "Wine When It is Red," in his *Stories, Essays and Poems* (London, 1935), p. 149.

4. Spinoza, *Ethica* (1677), Part III, Prop. VI. I have used George Eliot's translation, from *Ethics by Spinoza, Translated by George Eliot,* ed. T. Deegan (Salzburg, 1981), p. 100.

5. Fyodor Dostoevsky, *Notes from Underground,* trans. Mirra Ginsburg (New York, 1974), pp. 56–57.

6. *Ibid.,* pp. 64–65.

7. Arthur Miller, *Death of a Salesman* (New York, 1967), p. 56.

8. See in this connection Carl Cranor's useful analysis in his essay "Toward a Theory of Respect for Persons," *American Philosophical Quarterly* 12 (1975):309–19.

6. *The Prevalence of Humbug*

A greatly expanded version of the Stanton Griffis Lecture at the Cornell Medical School, New York City, March 1980. A variety of versions have been

delivered at Cornell University, the University of New Mexico, Dartmouth College, and elsewhere.

1. Maxim Gorky, *Reminiscences of Tolstoy, Chekhov, and Andreev* (New York, 1959), p. 79.

2. *The Mentor* (Springfield, O.) 18 (1930):16.

3. Of course, criticism of the substance of a remark will often imply censure of its producer. When Scrooge, in Dickens' *Christmas Carol,* responded to his nephew's cheerful greeting of "Merry Christmas!" with the unforgettably scornful "Bah! Humbug!" he reinforced his disgust by the rhetorical question "[W]hat right have you to be merry? what reason have you to be merry?" A little later he said: "If I could work my will, . . . every idiot who goes about with 'Merry Christmas' on his lips, should be boiled with his own pudding, and buried with a stake of holly through his heart. He should!"

Similarly, in the wonderful first chapter of *The Pickwick Papers,* Mr. Blotton (the haberdasher from Aldgate), in calling Mr. Pickwick a humbug, plainly intended a personal insult, later expunged by the frank declaration "that, personally, he entertained the highest regard and esteem for the honorable gentleman; he merely considered him a humbug in a Pickwickian point of view" (another tantalizing glimpse of our elusive quarry).

4. The point is well made in the second chapter of P. T. Barnum's *Humbugs of the World* (New York, 1866), the only book, to my knowledge, devoted to examination of humbug and its practitioners. He there says: "[A]s generally understood, 'humbug' consists in putting on glittering appearances—outside show—novel expedients, by which to suddenly arrest public attention, and attract the public eye and ear" (p. 20); and "An honest man who thus arrests public attention will be called a 'humbug,' but he is not a swindler or an impostor" (p. 21). Phineas Taylor Barnum (1810–91), often called the Prince of Humbugs, deserves to be better remembered than as the co-founder of the Barnum & Bailey circus. An international celebrity in his day, a favorite of Queen Victoria and a friend of Horace Greeley, Mark Twain, and Thackeray, he was an honest and harmless humbug, a magnificent advertiser, showman, and entertainer. Fortunately, his extraordinary career can now be enjoyably followed in Neil Harris' fine biography, *Humbug: The Art of P. T. Barnum* (Boston, 1973).

5. Rupert Crawshay-Williams, *Russell Remembered* (London, 1970), pp. 18–19.

6. *New York Times,* January 31, 1980, p. A10.

7. *Essays by Ralph Waldo Emerson* (New York, 1926), p. 139.

8. Ibid., p. 145.

9. Ibid., p. 146.

10. Carl L. Becker, *Cornell University: Founders and the Founding* (Ithaca, N.Y., 1944), pp. 132, 134.

11. Ibid., p. 135.

12. Ibid.

13. The opening section of the first paragraph of chap. 12 of Veblen's book, quoted from H. L. Mencken, *Prejudices: First Series* (New York, 1919), pp. 67–68.

14. Ibid., pp. 66–67, 69.

15. Sissela Bok, *Lying: Moral Choice in Public and Private Life* (New York, 1978), p. 14.

16. From the essay "Of Liars," in *The Essays of Montaigne,* trans. E. J. Trechmann (New York, 1946), 1:30.

17. Adrienne Rich, "Women and Honor: Some Notes on Lying," in *On Lies, Secrets, and Silence: Selected Prose* (New York, 1979), p. 188; my italics.

18. George Steiner, *After Babel* (New York, 1975), p. 220.

19. Ibid., pp. 221–22.

20. Ibid., p. 227.

21. M. F. K. Fisher, "As the Lingo Languishes," in *The State of the Language,* ed. L. Michaels and C. Ricks (Berkeley, 1980), p. 270.

22. Kay Dick, *Ivy and Stevie* (London, 1971), p. 10.

23. Sylvia Robinow, review of Marcia Millman, *Such a Pretty Face: Being Fat in America,* in *New Republic,* April 12, 1980, p. 35.

24. *Times Literary Supplement,* August 11, 1978, pp. 908–10.

25. *Flaubert's Dictionary of Accepted Ideas,* trans. and ed. Jacques Barzun (Norfolk, Conn., 1954).

26. See, for instance, the fine examples collected in George Oppenheimer, *Well, There's No Harm in Laughing* (Garden City, N.Y., 1970), formerly titled *Frank Sullivan through the Looking Glass.*

27. Russell Baker, *So This Is Depravity* (New York, 1980).

28. Charles Dickens, *David Copperfield,* chap. 20.

7. *The Rationality of Voting*

Originally a talk to the Philosophy Discussion Club at Cornell University. Read at Dalhousie University, 1977. A revised version appeared in *Reason, Action, and Experience,* ed. Helmut Kohlenberger (Hamburg: Felix Meiner Verlag, 1979), pp. 61–70.

8. *Newcomb's Problem Demystified*

Presented and discussed at Dartmouth College, 1982. Published in a somewhat different version in "Essays in Philosophical Analysis" ed. Ingmar Pörn, *Acta Philosophica Fennica* (Helsinki) 32 (1981):30–43.

1. First published in Robert Nozick, "Newcomb's Problem and Two Prin-

ciples of Choice," in N. Rescher et al., *Essays in Honor of Carl G. Hempel* (Dordrecht, 1969), pp. 114–46. The "original" version presented here is based on Nozick's presentation, with minor changes. See also Martin Gardner, "Mathematical Games," *Scientific American* 229, no. 1 (1973):104–9, and 230, no. 3 (1974):102–8. The latter article contains Nozick's reflections on the flood of letters elicited by the first article.

2. Nozick, "Newcomb's Problem."

3. See, for instance, James Randi, *The Magic of Uri Geller* (New York, 1975).

4. Suggested by, but not identical in application with, Max Weber's *Verstehen*.

5. Samuel Butler, though a confirmed atheist, used to attend a church service occasionally, so as not to settle into a rut—a policy to be recommended.

Index

Alexander, Samuel, 142
Aristotle, 49, 100, 103, 113
Austin, J. L., 49
Ayer, A. J., 25–26

Bacon, Francis, 90
Baker, Russell, 142, 181
Barnum, P. T., 115, 143, 180
Bazelon, David, 140–141
Becker, Carl L., 127–128, 180
Benda, Julien, 89
Bentham, Jeremy, 135
Blake, William, 132
Bok, Sissela, 132, 140
Boswell, James, 97–98
Bradley, F. H., 175
Brown, Martin, 179
Butler, Bishop Joseph, 69
Butler, Samuel, 182

Carlyle, Thomas, 135
Carroll, Lewis, 175
Cavett, Dick, 117
Chekhov, Anton, 115–116, 118, 142, 180
Chesterton, G. K., 104
Chwistek, Leon, 59–60
Clark, Joe, 124
Clough, Arthur Hugh, 109
Cornell, Ezra, 128
Cornford, F. M., 175
Cranor, Carl, 111
Crawshay-Williams, Rupert, 122, 180
Curtis, George William, 127–128

Darwin, Charles, 62, 88–89
DeGroot, Adriaan D., 31, 176
Descartes, René, 107, 171
Dickens, Charles, 125, 142, 181
Diogenes Laërtius, 176
Donnellan, Keith, 49
Dostoevsky, Fëdor, 102, 107–109

Einstein, Albert, 88–89
Emerson, Ralph Waldo, 124–125, 180
Euclid, 16, 29

Fieser, Louis Frederick, 92
Fisher, M. F. K., 136
Flaubert, Gustave, 142, 181

Franklin, Benjamin, 64
Fraser, A. C., 20
Freud, Sigmund, 72–73
Fuller, Watson, 179

Galileo, 62, 88
Gallie, W. B., 175
Gardner, Martin, 182
Geach, P. T., 22–23
Geller, Uri, 166
Ghotbzadeh, Sadegh, 123–124
Ginsberg, Morris, 41
Gogol, Nikolai, 137
Golding, William, 102, 114
Goodman, Paul, 88
Gorky, Maxim, 115, 180
Greeley, Horace, 180

Hardy, G. H., 68
Harris, Neil, 180
Heilbroner, Robert L., 84–85
Hellman, Lillian, 117–118
Hoagland, Hudson, 75
Hoover, Herbert, 116
Hudson, W. D., 179
Hume, David, 21–22, 55, 79–81, 85, 178
Huxley, Thomas H., 73–74

Johnson, Samuel, 97, 101, 120–121, 142

Kant, Immanuel, 98, 132–133
Kuhn, Thomas, 75

Lewis, C. Day, 175
Locke, John, 19
Lundberg, George A., 92

McCarthy, Mary, 117–119
Mach, Ernst, 62
McLuhan, Marshall, 119
Mann, Thomas, 142
Mannheim, Karl, 42
Marx, Groucho, 93
Melville, Herman, 142
Mencken, H. L., 14, 129, 175, 181
Michaels, L., 136
Miller, Arthur, 109
Millman, Marcia, 139
Molière, 142
Monod, Jacques, 83–84
Montaigne, Michel de, 132–133
Moore, G. E., 142

Nabokov, Vladimir, 137, 142
Newton, Isaac, 60–61, 88
Nietzsche, Friedrich, 135
Nozick, Robert, 161–162, 181–182

Oppenheimer, George, 181
Orwell, George, 142

Passmore, John, 176
Peirce, Charles Sanders, 21, 42
Piaget, Jean, 70–71
Planck, Max, 68
Pope, Alexander, 105
Popper, Karl, 26–27, 29, 176
Pyrrho of Elis, 176

Quine, W. V., 134

Randi, James, 182
Ranke, Leopold von, 68
Ravetz, J., 63–64
Reagan, Ronald, 66
Rescher, N., 182
Rich, Adrienne, 134
Ricks, C., 136
Rabinow, Sylvia, 139
Roszak, Theodore, 60–61, 64, 68
Rubinstein, Arthur, 106
Russell, Bertrand, 122–123

Sartre, Jean-Paul, 139
Schiller, F. C. S., 122
Shakespeare, William, 132
Shaw, George Bernard, 105, 116–118
Singer, Peter, 23
Smith, Adam, 84
Smith, Sydney, 142
Smith, T. V., 112
Spinoza, Baruch, 106
Steiner, George, 135–136
Stevenson, Adlai, 142
Sullivan, Frank, 142, 181
Swift, Jonathan, 102

Teichmann, E. J., 181
Thackeray, William Makepeace, 180
Twain, Mark, 180

Urmson, J. O., 53

Veblen, Thorstein, 128–129, 181
Victoria, queen of England, 180
Von Frisch, Karl, 87

Weber, Max, 182
Webster, Noah, 121–122, 134
White, Andrew D., 128
White, E. B., 142
White, Gilbert, 89
Wilde, Oscar, 19, 29
Wisehart, M. E., 116
Wittgenstein, Ludwig, 71

Library of Congress Cataloging in Publication Data

Black, Max, 1909–
The prevalence of humbug, and other essays.

Includes bibliographical references and index.
Contents: Why should I be rational?—Reasonableness—Scientific objectivity—[etc.]
1. Philosophy—Addresses, essays, lectures.
I. Title.
B29.B543 1983 100 82-22211
ISBN 0-8014-1514-4

DATE DUE

TRENT UNIVERSITY
0 1164 0525182 2

PHILOSOPHY

The Prevalence of Humbug and Other Essays

MAX BLACK

"Amusing and enlightening, these essays for the general reader deal with such topics as being reasonable, what it is to be humane, and P. T. Barnum's delightful book on humbugs. Reading them, one is reminded of William James's similar essays, not to adulterate philosophy into chit-chat but to elevate common sense."—*Key Reporter*

"[Black's] discussion is original, thorough and interesting, and shows that political philosophy need not always be empty verbiage."—*Times Higher Education Supplement*

"Written in an accessible style, Black's collection of eight essays will be of interest to both professional philosophers and the general reader. . . . Among the questions it asks are: Why should I be rational? What does it mean to behave humanely? What is it to be reasonable? Above all, the book is informed by a sense of humor, a quality as refreshing as it is rare; as Black observes, quoting Hu Hsai, 'life is too serious to be taken altogether seriously.'"—*Studies in the Humanities*

"Very few books by professional philosophers deserve as wide an audience as this one does."—Alasdair MacIntyre, Vanderbilt University

Cornell Paperbacks
Cornell University Press

ISBN 0-8014-9321-8