More praise for

Intuition Pumps and Other Tools for Thinking

"This user-friendly volume by Daniel Dennett distills a career's worth of thinking on evolution, consciousness and free will."

—*New York Times*

"In his long career, philosopher and cognitive scientist Daniel Dennett has made substantial contributions to—and stirred considerable controversy in—our thinking about human behavior and the very nature of the human mind. . . . His intellectual curiosity is deep and wide-ranging. And at his best, when he slows down and takes care with his writing, he can be both entertaining and enlightening in articulating difficult concepts."

—*Boston Globe*

"A master class in philosophical theatre, packed with snappy vignettes. . . . If you've never read Dennett, *Intuition Pumps* is the one to go for."

—*Guardian*

"Read this book carefully and you'll find yourself Jumping Out of the System in all directions. Dennett will lift off the top of your head, and tie your forehead into knots."

—*Daily Beast*

"Highly readable, ebullient, irreverent."

—*Brain*

"Aimed at ordinary civilians like us, and couched in language we can understand. . . . Fresh and fascinating."

—*Humanist*

"Dennett presents compelling expositions of provocative ideas. . . . An intellectual smorgasbord."

—*Publishers Weekly*

Copyrighted material

Intuition Pumps
and other tools for thinking

Daniel C. Dennett

W. W. Norton & Company
New York • London

Copyrighted material

Copyright © 2013 by Daniel C. Dennett

All rights reserved
Printed in the United States of America
First published as a Norton paperback 2014

Since this page cannot legibly accommodate all copyright notices,
page 461 constitutes an extension of the copyright page.

For information about permission to reproduce selections from
this book, write to Permissions, W. W. Norton & Company, Inc.,
500 Fifth Avenue, New York, NY 10110

For information about special discounts for bulk purchases,
please contact W. W. Norton Special Sales at specialsales@wwnorton.com
or 800-233-4830

Manufacturing by RR Donnelley, Harrisonburg
Book design by Marysarah Quinn
Production manager: Devon Zahn

Library of Congress Cataloging-in-Publication Data

Dennett, D. C. (Daniel Clement)
Intuition pumps and other tools for thinking / Daniel C. Dennett. — First edition
pages cm
Includes bibliographical references and index.
ISBN 978-0-393-08206-7 (hardcover)
1. Philosophy—Miscellanea. 2. Thought and thinking—Miscellanea. I. Title.
BD31.D46 2013
121'.6—dc23

2013000930

ISBN 978-0-393-34878-1 pbk.

W. W. Norton & Company, Inc., 500 Fifth Avenue, New York, N.Y. 10110
www.wwnorton.com

W. W. Norton & Company Ltd., Castle House, 75/76 Wells Street, London W1T 3QT

1 2 3 4 5 6 7 8 9 0

Copyrighted material

CONTENTS

Copyrighted material

Copyrighted material

Copyrighted material

Copyrighted material

PREFACE

Tufts University has been my academic home for more than forty years, and for me it has always seemed to be just right, like Goldilocks's porridge: not too burdened, not too pampered, brilliant colleagues to learn from with a minimum of academic prima donnas, good students serious enough to deserve attention without thinking they are entitled to round-the-clock maintenance, an ivory tower with a deep commitment to solving problems in the real world. Since creating the Center for Cognitive Studies in 1986, Tufts has supported my research, largely sparing me the ordeals and obligations of grantsmanship, and given me remarkable freedom to work with folks in many fields, either traveling afar to workshops, labs, and conferences or bringing visiting scholars and others to the Center. This book shows what I've been up to all these years.

In the spring of 2012, I test-flew a first draft of the chapters in a seminar I offered in the Tufts Philosophy Department. That has been my custom for years, but this time I wanted the students to help me make the book as accessible to the uninitiated as possible, so I excluded graduate students and philosophy majors and limited the class to just a dozen intrepid freshmen, the first twelve—actually thirteen, due to a clerical fumble—who volunteered. We led each other on a rollicking trip through the topics, as they learned that they really could stand up to the professor, and I learned that I really could

Copyrighted material

reach back farther and explain it all better. So here's to my young collaborators, with thanks for their courage, imagination, energy, and enthusiasm: Tom Addison, Nick Boswell, Tony Cannistra, Brendan Fleig-Goldstein, Claire Hirschberg, Caleb Malchik, Carter Palmer, Amar Patel, Kumar Ramanathan, Ariel Rascoe, Nikolai Renedo, Mikko Silliman, and Eric Tondreau.

The second draft that emerged from that seminar was then read by my dear friends Bo Dahlbom, Sue Stafford, and Dale Peterson, who provided me with still further usefully candid appraisals and suggestions, most of which I have followed, and by my editor, Drake McFeely, ably assisted by Brendan Curry, at W. W. Norton, who are also responsible for many improvements, for which I am grateful. Special thanks to Teresa Salvato, program coordinator at the Center for Cognitive Studies, who contributed directly to the entire project in innumerable ways and helped indirectly by managing the Center and my travels so effectively that I could devote more time and energy to making and using my thinking tools.

Finally, as always, thanks and love to my wife, Susan. We've been a team for fifty years, and she is as responsible as I am for what we, together, have done.

DANIEL C. DENNETT
Blue Hill, Maine
August 2012

Copyrighted material

Intuition Pumps
AND OTHER TOOLS FOR THINKING

Copyrighted material

Copyrighted material

I. INTRODUCTION: WHAT IS AN INTUITION PUMP?

You can't do much carpentry with your bare hands and you can't do much thinking with your bare brain.

—Bo Dahlbom

Thinking is hard. Thinking about some problems is so hard it can make your head ache just thinking about thinking about them. My colleague the neuropsychologist Marcel Kinsbourne suggests that whenever we find thinking hard, it is because the stony path to truth is competing with seductive, easier paths that turn out to be dead ends. Most of the effort in thinking is a matter of resisting these temptations. We keep getting waylaid and have to steel ourselves for the task at hand. *Ugh.*

There is a famous story about John von Neumann, the mathematician and physicist who turned Alan Turing's idea (what we now call a Turing machine) into an actual electronic computer (what we now call a Von Neumann machine, such as your laptop or smart phone). Von Neumann was a virtuoso thinker, legendary for his lightning capacity for doing prodigious calculations in his head. According to the story—and like most famous stories, this one has many versions—a colleague approached him one day with a puzzle that had two paths to a solution, a laborious, complicated calculation and an elegant, *Aha!*-

Copyrighted material

type solution. This colleague had a theory: in such a case, mathematicians work out the laborious solution while the (lazier, but smarter) physicists pause and find the quick-and-easy solution. Which solution would von Neumann find? You know the sort of puzzle: Two trains, 100 miles apart, are approaching each other on the same track, one going 30 miles per hour, the other going 20 miles per hour. A bird flying 120 miles per hour starts at train A (when they are 100 miles apart), flies to train B, turns around and flies back to the approaching train A, and so forth, until the two trains collide. How far has the bird flown when the collision occurs? "Two hundred and forty miles," von Neumann answered almost instantly. "Darn," replied his colleague, "I predicted you'd do it the hard way, summing the infinite series." "Ay!" von Neumann cried in embarrassment, smiting his forehead. "There's an easy way!" (Hint: How long until the trains collide?)

Some people, like von Neumann, are such natural geniuses that they can breeze through the toughest tangles; others are more plodding but are blessed with a heroic supply of "willpower" that helps them stay the course in their dogged pursuit of truth. Then there are the rest of us, not calculating prodigies and a little bit lazy, but still aspiring to understand whatever confronts us. What can we do? We can use thinking tools, by the dozens. These handy prosthetic imagination-extenders and focus-holders permit us to think reliably and even gracefully about really hard questions. This book is a collection of my favorite thinking tools. I will not just describe them; I intend to *use* them to move your mind gently through uncomfortable territory all the way to a quite radical vision of meaning, mind, and free will. We will begin with some tools that are simple and general, having applications to all sorts of topics. Some of these are familiar, but others have not been much noticed or discussed. Then I will introduce you to some tools that are for very special purposes indeed, designed to explode one specific seductive idea or another, clearing a way out of a deep rut that still traps and flummoxes experts. We will also encounter and dismantle a variety of bad thinking tools, misbegotten persuasion-devices that can lead you astray if you aren't

Copyrighted material

careful. Whether or not you arrive comfortably at my proposed destination—and decide to stay there with me—the journey will equip you with new ways of thinking about the topics, and thinking about thinking.

The physicist Richard Feynman was perhaps an even more legendary genius than von Neumann, and he was certainly endowed with a world-class brain—but he also loved having fun, and we can all be grateful that he particularly enjoyed revealing the tricks of the trade he used to make life easier for himself. No matter how smart you are, you're smarter if you take the easy ways when they are available. His autobiographical books, *"Surely You're Joking, Mr. Feynman!"* and *What Do You Care What Other People Think?*, should be on the required reading list of every aspiring thinker, since they have many hints about how to tame the toughest problems—and even how to dazzle an audience with fakery when nothing better comes to mind. Inspired by the wealth of useful observations in his books, and his candor in revealing how his mind worked, I decided to try my own hand at a similar project, less autobiographical and with the ambitious goal of persuading you to think about these topics *my way*. I will go to considerable lengths to cajole you out of some of your firmly held convictions, but with nothing up my sleeve. One of my main goals is to reveal along the way just what I am doing and why.

Like all artisans, a blacksmith needs tools, but—according to an old (indeed almost extinct) observation—blacksmiths are unique in that they make their own tools. Carpenters don't make their saws and hammers, tailors don't make their scissors and needles, and plumbers don't make their wrenches, but blacksmiths can make their hammers, tongs, anvils, and chisels out of their raw material, iron. What about thinking tools? Who makes them? And what are they made of? Philosophers have made some of the best of them—out of nothing but ideas, useful structures of information. René Descartes gave us *Cartesian coordinates*, the x- and y-axes without which *calculus*—a thinking tool *par excellence* simultaneously invented by Isaac Newton and the philosopher Gottfried Wilhelm Leibniz—would be almost

Copyrighted material

unthinkable. Blaise Pascal gave us *probability theory* so we can easily calculate the odds of various wagers. The Reverend Thomas Bayes was also a talented mathematician, and he gave us *Bayes's theorem*, the backbone of Bayesian statistical thinking. But most of the tools that feature in this book are simpler ones, not the precise, systematic machines of mathematics and science but the hand tools of the mind. Among them are

Labels. Sometimes just creating a vivid name for something helps you keep track of it while you turn it around in your mind trying to understand it. Among the most useful labels, as we shall see, are warning labels or alarms, which alert us to likely sources of error.

Examples. Some philosophers think that using examples in their work is, if not quite cheating, at least uncalled for—rather the way novelists shun illustrations in their novels. The novelists take pride in doing it all with words, and the philosophers take pride in doing it all with carefully crafted abstract generalizations presented in rigorous order, as close to mathematical proofs as they can muster. Good for them, but they can't expect me to recommend their work to any but a few remarkable students. It's just more difficult than it has to be.

Analogies and metaphors. Mapping the features of one complex thing onto the features of another complex thing that you already (think you) understand is a famously powerful thinking tool, but it is so powerful that it often leads thinkers astray when their imaginations get captured by a treacherous analogy.

Staging. You can shingle a roof, paint a house, or fix a chimney with the help of just a ladder, moving it and climbing, moving it and climbing, getting access to only a small part

Copyrighted material

of the job at a time, but it's often a lot easier in the end to take the time at the beginning to erect some sturdy staging that will allow you to move swiftly and safely around the whole project. Several of the most valuable thinking tools in this book are examples of staging that take some time to put in place but then permit a variety of problems to be tackled together—without all the ladder-moving.

And, finally, the sort of thought experiments I have dubbed *intuition pumps*.

Thought experiments are among the favorite tools of philosophers, not surprisingly. Who needs a lab when you can figure out the answer to your question by some ingenious deduction? Scientists, from Galileo to Einstein and beyond, have also used thought experiments to good effect, so these are not just philosophers' tools. Some thought experiments are analyzable as rigorous arguments, often of the form ***reductio ad absurdum***,* in which one takes one's opponents' premises and derives a formal contradiction (an absurd result), showing that they can't all be right. One of my favorites is the proof attributed to Galileo that heavy things don't fall faster than lighter things (when friction is negligible). If they did, he argued, then since heavy stone A would fall faster than light stone B, if we tied B to A, stone B would act as a drag, slowing A down. But A tied to B is heavier than A alone, so the two together should also fall faster than A by itself. We have concluded that tying B to A would make something that fell both faster and slower than A by itself, which is a contradiction.

Other thought experiments are less rigorous but often just as effective: little stories designed to provoke a heartfelt, table-thumping

* Words and phrases in boldface are the names of tools for thinking described and discussed in more detail elsewhere in the book. Look in the index to find them, since some of them do not get a whole piece to themselves.

Copyrighted material

intuition—"Yes, of course, it has to be so!"—about whatever thesis is being defended. I have called these *intuition pumps*. I coined the term in the first of my public critiques of philosopher John Searle's famous **Chinese Room** thought experiment (Searle, 1980; Dennett, 1980), and some thinkers concluded I meant the term to be disparaging or dismissive. On the contrary, I love intuition pumps! That is, some intuition pumps are excellent, some are dubious, and only a few are downright deceptive. Intuition pumps have been a dominant force in philosophy for centuries. They are the philosophers' version of Aesop's fables, which have been recognized as wonderful thinking tools since before there were philosophers.* If you ever studied philosophy in college, you were probably exposed to such classics as Plato's cave, in *The Republic*, in which people are chained and can see only the shadows of real things cast on the cave wall; or his example, in *Meno*, of teaching geometry to the slave boy. Then there is Descartes's evil demon, deceiving Descartes into believing in a world that was entirely illusory—the original Virtual Reality thought experiment—and Hobbes's state of nature, in which life is nasty, brutish, and short. Not as famous as Aesop's "Boy Who Cried Wolf" or "The Ant and the Grasshopper," but still widely known, each is designed to pump some intuitions. Plato's cave purports to enlighten us about the nature of perception and reality, and the slave boy is supposed to illustrate our innate knowledge; the evil demon is the ultimate skepticism-generator, and our improvement over the state of nature when we contract to form a society is the point of Hobbes's parable. These are the enduring melodies of philosophy, with the staying power that ensures that students will remember them, quite vividly and accurately, years after they have forgotten the intricate surrounding arguments and analysis. A good intuition pump is more robust than any one version of it. We will consider a

* Aesop, like Homer, is almost as mythic as his fables, which were transmitted orally for centuries before they were first written down a few hundred years before the era of Plato and Socrates. Aesop may not have been Greek; there is circumstantial evidence that he was Ethiopian.

Copyrighted material

variety of contemporary intuition pumps, including some defective ones, and the goal will be to understand what they are good for, how they work, how to use them, and even how to make them.

Here's a short, simple example: the Whimsical Jailer. Every night he waits until all the prisoners are sound asleep and then he goes around unlocking all the doors, leaving them open for hours on end. Question: Are the prisoners free? Do they have an *opportunity* to leave? Not really. Why not? Here's another example: the Jewels in the Trashcan. There happens to be a fortune in jewelry discarded in the trashcan on the sidewalk that you stroll by one night. It might seem that you have a golden opportunity to become rich, except it isn't golden at all because it is a *bare* opportunity, one that you would be extremely unlikely to recognize and hence act on—or even consider. These two simple scenarios pump intuitions that might not otherwise be obvious: the importance of getting timely information about genuine opportunities, soon enough for the information to cause us to consider it in time to do something about it. In our eagerness to make "free" choices, uncaused—we like to think—by "external forces," we tend to forget that we shouldn't want to be cut off from all such forces; free will does not *abhor* our embedding in a rich causal context; it actually *requires* it.

I hope you feel that there is more to be said on that topic! These tiny intuition pumps raise an issue vividly, but they don't settle anything—yet. (A whole section will concentrate on free will later.) We need to become practiced in the art of treating such tools warily, watching where we step, and checking for pitfalls. If we think of an intuition pump as a carefully designed persuasion tool, we can see that it might repay us to *reverse engineer* the tool, checking out all the moving parts to see what they are doing.

When Doug Hofstadter and I composed *The Mind's I* back in 1982, he came up with just the right advice on this score: consider the intuition pump to be a tool with many settings, and "turn all the knobs" to see if the same intuitions still get pumped when you consider variations.

Copyrighted material

So let's identify, and turn, the knobs on the Whimsical Jailer. Assume—until proved otherwise—that every part has a function, and see what that function is by replacing it with another part, or transforming it slightly.

1. Every night
2. he waits
3. until all the prisoners
4. are sound asleep
5. and then he goes around unlocking
6. all the doors,
7. leaving them open for hours on end.

Here is one of many variations we could consider:

> One night he ordered his guards to drug one of the prisoners and after they had done this they accidentally left the door of that prisoner's cell unlocked for an hour.

It changes the flavor of the scenario quite a lot, doesn't it? How? It still makes the main point (doesn't it?) but not as effectively. The big difference seems to be between being naturally asleep—you *might* wake up any minute—and being drugged or comatose. Another difference—"accidentally"—highlights the role of the intention or inadvertence on the part of the jailor or the guards. The repetition ("every night") seems to change the odds, in favor of the prisoners. When and why do the odds matter? How much would you pay *not* to have to participate in a lottery in which a million people have tickets and the "winner" is shot? How much would you pay *not* to have to play Russian roulette with a six-shooter? (Here we use one intuition pump to illuminate another, a trick to remember.)

Other knobs to turn are less obvious: The Diabolical Host secretly locks the bedroom doors of his houseguests while they sleep. The Hospital Manager, worried about the prospect of a fire, keeps the

Copyrighted material

doors of all the rooms and wards unlocked at night, but she doesn't inform the patients, thinking they will sleep more soundly if they don't know. Or what if the prison is somewhat larger than usual, say, the size of Australia? You can't lock or unlock all the doors to Australia. What difference does that make?

This self-conscious wariness with which we should approach any intuition pump is itself an important tool for thinking, the philosophers' favorite tactic: "*going meta*"—thinking about thinking, talking about talking, reasoning about reasoning. Meta-language is the language we use to talk about another language, and meta-ethics is a bird's-eye view examination of ethical theories. As I once said to Doug, "Anything you can do I can do meta-." This whole book is, of course, an example of going meta: exploring how to think carefully about methods of thinking carefully (about methods of thinking carefully, etc.).* He recently (2007) offered a list of some of his own favorite small hand tools:

wild goose chases
tackiness
dirty tricks
sour grapes
elbow grease
feet of clay
loose cannons
crackpots
lip service
slam dunks
feedback

* The philosopher W. V. O. Quine (1960) called this *semantic ascent*, going *up* from talking about electrons or justice or horses or whatever to talking about *talking about* electrons or justice or horses or whatever. Sometimes people object to this move by philosophers ("With you folks, it's all just semantics!"), and sometimes the move is indeed useless or even bamboozling, but when it's needed, when people are talking past each other, or being fooled by tacit assumptions about what their own words mean, semantic ascent, or going meta, is the key to clarity.

Copyrighted material

If these expressions are familiar to you, they are not "just words" for you; each is an abstract cognitive tool, in the same way that *long division* or *finding-the-average* is a tool; each has a role to play in a broad spectrum of contexts, making it easier to formulate hypotheses to test, making it easier to recognize unnoticed patterns in the world, helping the user look for important similarities, and so forth. Every word in your vocabulary is a simple thinking tool, but some are more useful than others. If any of these expressions are not in your kit, you might want to acquire them; equipped with such tools you will be able to think thoughts that would otherwise be relatively hard to formulate. Of course, as the old saw has it, when your only tool is a hammer, everything looks like a nail, and each of these tools can be overused.

Let's look at just one of these: sour grapes. It comes from Aesop's fable "The Fox and the Grapes" and draws attention to how sometimes people pretend not to care about something they can't have by disparaging it. Look how much you can say about what somebody has just said by asking, simply, "Sour grapes?" It gets her to consider a possibility that might otherwise have gone unnoticed, and this might very effectively inspire her to revise her thinking, or reflect on the issue from a wider perspective—or it might very effectively insult her. (Tools can be used as weapons too.) So familiar is the moral of the story that you may have forgotten the tale leading up to it, and may have lost touch with the subtleties—if they matter, and sometimes they don't.

Acquiring tools and using them wisely are distinct skills, but you have to start by acquiring the tools, or making them yourself. Many of the thinking tools I will present here are my own inventions, but others I have acquired from others, and I will acknowledge their inventors in due course.* None of the tools on Doug's list are his

* Many of the passages in this book have been drawn from books and articles I have previously published, revised to make them more portable and versatile, fit for use in contexts other than the original—a feature of most good tools. For instance, the opening story about von Neumann appeared in my 1995 book *Darwin's Dangerous Idea*, and this discussion of Hofstadter's hand tools appeared in my 2009 *PNAS* paper, "Darwin's 'Strange Inversion of Reasoning.'" Instead of footnoting all of these, I provide a list of sources at the end of the book.

Copyrighted material

inventions, but he has contributed some fine specimens to my kit, such as **jootsing** and **sphexishness**.

Some of the most powerful thinking tools are mathematical, but aside from mentioning them, I will not devote much space to them because this is a book celebrating the power of *non*-mathematical tools, *informal* tools, the tools of prose and poetry, if you like, a power that scientists often underestimate. You can see why. First, there is a culture of scientific writing in research journals that favors—indeed insists on—an impersonal, stripped-down presentation of the issues with a minimum of flourish, rhetoric, and allusion. There is a good reason for the relentless drabness in the pages of our most serious scientific journals. As one of my doctoral examiners, the neuroanatomist J. Z. Young, wrote to me in 1965, in objecting to the somewhat fanciful prose in my dissertation at Oxford (in philosophy, not neuroanatomy), English was becoming the international language of science, and it behooves us native English-speakers to write works that can be read by "a patient Chinee [*sic*] with a good dictionary." The results of this self-imposed discipline speak for themselves: whether you are a Chinese, German, Brazilian—or even a French—scientist, you insist on publishing your most important work in English, bare-bones English, translatable with minimal difficulty, relying as little as possible on cultural allusions, nuances, word-play, and even metaphor. The level of mutual understanding achieved by this international system is invaluable, but there is a price to be paid: some of the thinking that has to be done apparently *requires* informal metaphor-mongering and imagination-tweaking, assaulting the barricades of closed minds with every trick in the book, and if some of this cannot be easily translated, then I will just have to hope for virtuoso translators on the one hand, and the growing fluency in English of the world's scientists on the other.

Another reason why scientists are often suspicious of theoretical discussions conducted in "mere words" is that they recognize that the task of criticizing an argument not formulated in mathematical equations is much trickier, and typically less conclusive. The language

Copyrighted material

of mathematics is a reliable enforcer of cogency. It's like the net on the basketball hoop: it removes sources of disagreement and judgment about whether the ball went in. (Anyone who has played basketball on a playground court with a bare hoop knows how hard it can be to tell an air ball from a basket.) But sometimes the issues are just too slippery and baffling to be tamed by mathematics.

I have always figured that if I can't explain something I'm doing to a group of bright undergraduates, I don't really understand it myself, and that challenge has shaped everything I have written. Some philosophy professors yearn to teach advanced seminars only to graduate students. Not me. Graduate students are often too eager to prove to each other and to themselves that they are savvy operators, wielding the jargon of their trade with deft assurance, baffling outsiders (that's how they assure themselves that what they are doing requires expertise), and showing off their ability to pick their way through the most tortuous (and torturous) technical arguments without getting lost. Philosophy written for one's advanced graduate students and fellow experts is typically all but unreadable—and hence largely unread.

A curious side effect of my policy of trying to write arguments and explanations that can be readily understood by people outside philosophy departments is that there are philosophers who as a matter of "principle" won't take my arguments seriously! When I gave the John Locke Lectures at Oxford many years ago to a standing-room-only audience, a distinguished philosopher was heard to grumble as he left one of them that he was damned if he would learn anything from somebody who could attract non-philosophers to the Locke Lectures! True to his word, he never learned anything from me, so far as I can tell. I did not adjust my style and have never regretted paying the price. There is a time and a place in philosophy for rigorous arguments, with all the premises numbered and the inference rules named, but these do not often need to be paraded in public. We ask our graduate students to prove they can do it in their dissertations, and some never outgrow the habit, unfortunately. And to be fair, the opposite sin of high-flown

Copyrighted material

Continental rhetoric, larded with literary ornament and intimations of profundity, does philosophy no favors either. If I had to choose, I'd take the hard-bitten analytic logic-chopper over the deep purple sage every time. At least you can usually figure out what the logic-chopper is talking about and what would count as being wrong.

The middle ground, roughly halfway between poetry and mathematics, is where philosophers can make their best contributions, I believe, yielding genuine clarifications of deeply puzzling problems. There are no feasible algorithms for doing this kind of work. Since *everything* is up for grabs, one chooses one's fixed points with due caution. As often as not, an "innocent" assumption accepted without notice on all sides turns out to be the culprit. Exploring such treacherous conceptual territories is greatly aided by using thinking tools devised on the spot to clarify the alternative paths and shed light on their prospects.

These thinking tools seldom establish a *fixed* fixed point—a solid "axiom" for all future inquiry—but rather introduce a worthy *candidate* for a fixed point, a likely constraint on future inquiry, but itself subject to revision or jettisoning altogether if somebody can figure out why. No wonder many scientists have no taste at all for philosophy; everything is up for grabs, nothing is take-it-to-the-bank secure, and the intricate webs of argument constructed to connect these "fixed" points hang provisionally in the air, untethered to clear foundations of empirical proof or falsification. So these scientists turn their backs on philosophy and get on with their work, but at the cost of leaving some of the most important and fascinating questions unconsidered. "Don't ask! Don't tell! It's premature to tackle the problem of consciousness, of free will, of morality, of meaning and creativity!" But few can live with such abstemiousness, and in recent years scientists have set out on a gold rush of sorts into these shunned regions. Seduced by sheer curiosity (or, sometimes, perhaps, a yearning for celebrity), they embark on the big questions and soon discover how hard it is to make progress on them. I must confess that one of the delicious, if guilty, pleasures I enjoy is watching eminent scientists, who only a few years

Copyrighted material

ago expressed withering contempt for philosophy,* stumble embarrassingly in their own efforts to set the world straight on these matters with a few briskly argued extrapolations from their own scientific research. Even better is when they request, and acknowledge, a little help from us philosophers.

In the first section that follows, I present a dozen general, all-purpose tools, and then in subsequent sections I group the rest of the entries not by the type of tool but by the topic where the tool works best, turning first to the most fundamental philosophical topic—meaning, or content—followed by evolution, consciousness, and free will. A few of the tools I present are actual software, friendly devices that can do for your naked imagination what telescopes and microscopes can do for your naked eye.

Along the way, I will also introduce some false friends, tools that blow smoke instead of shining light. I needed a term for these hazardous devices, and found *le mot juste* in my sailing experience. Many sailors enjoy the nautical terms that baffle landlubbers: port and starboard, gudgeon and pintle, shrouds and spreaders, cringles and fairleads, and all the rest. A running joke on a boat I once sailed on involved making up false definitions for these terms. So a *binnacle* was a marine growth on compasses, and a *mast tang* was a citrus beverage enjoyed aloft; a *snatch block* was a female defensive maneuver, and a *boom crutch* was an explosive orthopedic device. I've never since been able to think of a boom crutch—a removable wooden stand on which the boom rests when the sail is lowered—without a momentary image of *kapow!* in some poor fellow's armpit. So I chose the term as my name for thinking tools that backfire, the ones that only *seem* to aid in understanding but that actually spread darkness and confusion instead of light. Scattered through these chapters are a variety of boom crutches with suitable warning labels, and examples

* Two of the best: "Philosophy is to science what pigeons are to statues," and "Philosophy is to science as pornography is to sex: it is cheaper, easier and some people prefer it." (I'll leave these unattributed, but their authors can choose to claim them if they wish.)

Copyrighted material

to deplore. And I close with some further reflections on what it is like to be a philosopher, in case anybody wants to know, including some advice from Uncle Dan to any of you who might have discovered a taste for this way of investigating the world and wonder whether you are cut out for a career in the field.

Copyrighted material

Copyrighted material

II.
A DOZEN GENERAL THINKING TOOLS

Most of the thinking tools in this book are quite specialized, made to order for application to a particular topic and even a particular controversy within the topic. But before we turn to these intuition pumps, here are a few general-purpose thinking tools, ideas and practices that have proved themselves in a wide variety of contexts.

Copyrighted material

Copyrighted material

1. MAKING MISTAKES

> He who says "Better to go without belief forever than believe a lie!" merely shows his own preponderant private horror of becoming a dupe. . . . It is like a general informing his soldiers that it is better to keep out of battle forever than to risk a single wound. Not so are victories either over enemies or over nature gained. Our errors are surely not such awfully solemn things. In a world where we are so certain to incur them in spite of all our caution, a certain lightness of heart seems healthier than this excessive nervousness on their behalf.
>
> —William James, "The Will to Believe"

> If you've made up your mind to test a theory, or you want to explain some idea, you should always decide to publish it whichever way it comes out. If we only publish results of a certain kind, we can make the argument look good. We must publish both kinds of results.
>
> —Richard Feynman, *"Surely You're Joking, Mr. Feynman!"*

Scientists often ask me why philosophers devote so much of their effort to teaching and learning the history of their field. Chemists typically get by with only a rudimentary knowledge of the history of chemistry, picked up along the way, and many molecular biologists, it seems, are not even curious about what happened in biology before about 1950. My answer is that the history of philosophy is in large measure the history of very smart people making very tempting mistakes, and if you don't know the history, you are doomed to making the same darn mistakes all over again. That's why we teach the history of the field to our students, and scientists who blithely ignore philosophy do

Copyrighted material

so at their own risk. There is no such thing as philosophy-free science, just science that has been conducted without any consideration of its underlying philosophical assumptions. The smartest or luckiest of the scientists sometimes manage to avoid the pitfalls quite adroitly (perhaps they are "natural born philosophers"—or are as smart as they think they are), but they are the rare exceptions. Not that professional philosophers don't make—and even defend—the old mistakes too. If the questions weren't hard, they wouldn't be worth working on.

Sometimes you don't just want to *risk* making mistakes; you actually want to make them—if only to give you something clear and detailed to fix. Making mistakes is the key to making progress. Of course there are times when it is really important not to make any mistakes—ask any surgeon or airline pilot. But it is less widely appreciated that there are also times when making mistakes is the only way to go. Many of the students who arrive at very competitive universities pride themselves in not making mistakes—after all, that's how they've come so much farther than their classmates, or so they have been led to believe. I often find that I have to encourage them to *cultivate the habit* of making mistakes, the best learning opportunities of all. They get "writer's block" and waste hours forlornly wandering back and forth on the starting line. "Blurt it out!" I urge them. Then they have something on the page to work with.

We philosophers are mistake specialists. (I know, it sounds like a bad joke, but hear me out.) While other disciplines specialize in getting the right answers to their defining questions, we philosophers specialize in all the ways there are of getting things so mixed up, so deeply wrong, that nobody is even sure what the right *questions* are, let alone the answers. Asking the wrong questions risks setting any inquiry off on the wrong foot. Whenever that happens, this is a job for philosophers! Philosophy—in every field of inquiry—is what you have to do until you figure out what questions you should have been asking in the first place. Some people hate it when that happens. They would rather take their questions off the rack, all nicely tailored and pressed and cleaned and ready to answer. Those who feel that way

Copyrighted material

can do physics or mathematics or history or biology. There's plenty of work for everybody. We philosophers have a taste for working on the questions that need to be straightened out before they can be answered. It's not for everyone. But try it, you might like it.

In the course of this book I am going to jump vigorously on what I claim are other people's mistakes, but I want to assure you that I am an experienced mistake-maker myself. I've made some dillies, and hope to make a lot more. One of my goals in this book is to help you make *good* mistakes, the kind that light the way for everybody.

First the theory, and then the practice. Mistakes are not just opportunities for learning; they are, in an important sense, the *only* opportunity for learning or making something truly new. Before there can be learning, there must be learners. There are only two non-miraculous ways for learners to come into existence: they must either evolve or be designed and built by learners that evolved. Biological evolution proceeds by a grand, inexorable process of trial and error—and without the *errors* the trials wouldn't accomplish anything. As Gore Vidal once said, "It is not enough to succeed. Others must fail." Trials can be either *blind* or *foresighted*. You, who know a lot, but not the answer to the question at hand, can take leaps—foresighted leaps. You can look before you leap, and hence be *somewhat* guided from the outset by what you already know. *You* need not be guessing at random, but don't look down your nose at random guesses; among its wonderful products is . . . you!

Evolution is one of the central themes of this book, as of all my books, for the simple reason that it is the central, enabling process not only of life but also of knowledge and learning and understanding. If you attempt to make sense of the world of ideas and meanings, free will and morality, art and science and even philosophy itself without a sound and quite detailed knowledge of evolution, you have one hand tied behind your back. Later, we will look at some tools designed to help you think about some of the more challenging questions of evolution, but here we need to lay a foundation. For evolution, which knows nothing, the steps into novelty are blindly taken by mutations, which

Copyrighted material

are random copying "errors" in DNA. Most of these typographical errors are of no consequence, since nothing reads them! They are as inconsequential as the rough drafts you didn't, or don't, hand in to the teacher for grading. The DNA of a species is rather like a recipe for building a new body, and most of the DNA is never actually consulted in the building process. (It is often called "junk DNA" for just that reason.) In the DNA sequences that do get read and acted upon during development, the vast majority of mutations are harmful; many, in fact, are swiftly fatal. Since the majority of "expressed" mutations are deleterious, the process of natural selection actually works to keep the mutation rate very low. Each of you has very, very good copying machinery in your cells. For instance, you have roughly a trillion cells in your body, and each cell has either a perfect or an *almost* perfect copy of your genome, over three billion symbols long, the recipe for you that first came into existence when your parents' egg and sperm joined forces. Fortunately, the copying machinery does not achieve perfect success, for if it did, evolution would eventually grind to a halt, its sources of novelty dried up. Those tiny blemishes, those "imperfections" in the process, are the source of all the wonderful design and complexity in the living world. (I can't resist adding: if anything deserves to be called Original Sin, these copying mistakes do.)

The chief trick to making good mistakes is not to hide them—especially not from yourself. Instead of turning away in denial when you make a mistake, you should become a connoisseur of your own mistakes, turning them over in your mind as if they were works of art, which in a way they are. The fundamental reaction to any mistake ought to be this: "Well, I won't do *that* again!" Natural selection doesn't actually think the thought; it just wipes out the goofers before they can reproduce; natural selection won't do *that* again, at least not as often. Animals that can learn—learn not to make that noise, touch that wire, eat that food—have *something* with a similar selective force in their brains. (B. F. Skinner and the behaviorists understood the need for this and called it "reinforcement" learning; *that* response is not reinforced and suffers "extinction.") We human

Copyrighted material

beings carry matters to a much more swift and efficient level. We can actually *think the thought*, reflecting on what we have just done: "Well, I won't do *that* again!" And when we reflect, we confront directly the problem that must be solved by any mistake-maker: what, exactly, is *that*? What was it about what I just did that got me into all this trouble? The trick is to take advantage of the particular details of the mess you've made, so that your next attempt will be informed by it and not just another blind stab in the dark.

We have all heard the forlorn refrain "Well, it seemed like a good idea at the time!" This phrase has come to stand for the rueful reflection of an idiot, a sign of stupidity, but in fact we should appreciate it as a pillar of wisdom. Any being, any agent, who can truly say, "Well, it seemed like a good idea at the time!" is standing on the threshold of brilliance. We human beings pride ourselves on our intelligence, and one of its hallmarks is that we can *remember our previous thinking*, and reflect on it—on how it seemed, on why it was tempting in the first place, and then about what went wrong. I know of no evidence to suggest that any other species on the planet can actually think this thought. If they could, they would be almost as smart as we are.

So when you make a mistake, you should learn to take a deep breath, grit your teeth, and then *examine* your own recollections of the mistake as ruthlessly and as dispassionately as you can manage. It's not easy. The natural human reaction to making a mistake is embarrassment and anger (we are never angrier than when we are angry at ourselves), and you have to work hard to overcome these emotional reactions. Try to acquire the weird practice of savoring your mistakes, delighting in uncovering the strange quirks that led you astray. Then, once you have sucked out all the goodness to be gained from having made them, you can cheerfully set them behind you, and go on to the next big opportunity. But that is not enough: you should actively seek out opportunities to make grand mistakes, just so you can then recover from them.

At its simplest, this is a technique we all learned in grade school. Recall how strange and forbidding long division seemed at first: You

Copyrighted material

were confronted by two imponderably large numbers, and you had to figure out how to start. Does the divisor go into the dividend six or seven or eight times? Who knew? You didn't have to know; you just had to take a stab at it, whichever number you liked, and check out the result. I remember being almost shocked when I was told I should start by just "making a guess." Wasn't this *mathematics*? You weren't supposed to play guessing games in such a serious business, were you? But eventually I appreciated, as we all do, the beauty of the tactic. If the chosen number turned out to be too small, you increased it and started over; if too large, you decreased it. The good thing about long division was that it always worked, even if you were maximally stupid in making your first choice, in which case it just took a little longer.

This general technique of making a more-or-less educated guess, working out its implications, and using the result to make a correction for the next phase has found many applications. A key element of this tactic is making a mistake that is clear and precise enough to have definite implications. Before GPS came along, navigators used to determine their position at sea by first making a guess about where they were (they made a guess about *exactly* what their latitude and longitude were), and then calculating exactly how high in the sky the sun would appear to be if that were—by an incredible coincidence—their actual position. When they used this method, they didn't expect to hit the nail on the head. They didn't have to. Instead they then measured the actual elevation angle of the sun (exactly) and compared the two values. With a little more trivial calculation, this told them how big a correction, and in what direction, to make to their initial guess.* In

* This doesn't give navigators their actual position, a point on the globe, but it does give them a *line*. They are somewhere on that *line of position* (LOP). Wait a few hours until the sun has moved on quite a bit. Then choose a point on your LOP, any point, and calculate how high the sun would be now if that point were exactly the right choice. Make the observation, compare the results, apply the correction, and get another LOP. Where it crosses your first LOP is the point where you are. The sun will have changed not only its height but also its compass bearing during those hours so the lines will cross at a pretty good angle. In practice, you are usually moving during those few hours, so you advance your first LOP in the direction you are moving by calculating your speed and drawing an advanced LOP parallel to the original LOP. In real life everything has a bit of slop in it, so you try to get three different

Copyrighted material

such a method it is useful to make a pretty good guess the first time, but it doesn't matter that it is bound to be mistaken; the important thing is to make the mistake, in glorious detail, so there is something serious to correct. (A GPS device uses the same guess-and-fix-it strategy to locate itself relative to the overhead satellites.)

The more complex a problem you're facing, of course, the more difficult the analysis is. This is known to researchers in artificial intelligence (AI) as the problem of "credit assignment" (it could as well be called blame assignment). Figuring out what to credit and what to blame is one of the knottiest problems in AI, and it is also a problem faced by natural selection. Every organism on the earth dies sooner or later after one complicated life story or another. How on earth could natural selection *see through* the fog of all these details in order to figure out what positive factors to "reward" with offspring and what negative factors to "punish" with childless death? Can it really be that some of our ancestors' siblings died childless *because their eyelids were the wrong shape*? If not, how could the process of natural selection explain why our eyelids came to have the excellent shapes they have? Part of the answer is familiar: following the old adage "If it ain't broke, don't fix it," leave almost all your old, conservative design solutions in place and take your risks with a safety net in place. Natural selection automatically conserves *whatever has worked up to now*, and fearlessly explores innovations large and small; the large ones almost always lead immediately to death. A terrible waste, but nobody's counting. Our eyelids were mostly designed by natural selection long before there were human beings or even primates or even mammals. They've had more than a hundred million years to reach the shape they are today, with only a few minor touch-ups in the last six million years, since we shared a common ancestor with the chimpanzees and the bonobos. Another part of the answer is that natural selection works

LOPs. If they all intersect in exactly the same point, you're either incredibly good or incredibly lucky, but more commonly they form a small triangle, called a *cocked hat*. You consider yourself in the middle of the cocked hat, and that's your new calculated position.

Copyrighted material

with large numbers of cases, where even minuscule advantages show up statistically and can be automatically accumulated. (Other parts of the answer are technicalities beyond this elementary discussion.)

Here is a technique that card magicians—at least the best of them—exploit with amazing results. (I don't expect to incur the wrath of the magicians for revealing this trick to you, since this is not a *particular* trick but a deep general principle.) A good card magician knows many tricks that depend on luck—they don't always work, or even often work. There are some effects—they can hardly be called tricks—that might work only once in a thousand times! Here is what you do: You start by telling the audience you are going to perform *a* trick, and without telling them what trick you are doing, you go for the one-in-a-thousand effect. It almost never works, of course, so you glide seamlessly into a second try—for an effect that works about one time in a hundred, perhaps—and when it too fails (as it almost always will), you slide gracefully into effect number 3, which works only about one time in ten, so you'd better be ready with effect number 4, which works half the time (let's say). If all else fails (and by this time, *usually* one of the earlier safety nets will have kept you out of this worst case), you have a failsafe effect, which won't impress the crowd very much but at least it's a surefire trick. In the course of a whole performance, you will be very unlucky indeed if you always have to rely on your final safety net, and whenever you achieve one of the higher-flying effects, the audience will be stupefied. "Impossible! How on earth could you have known which was my card?" *Aha!* You didn't know, but you had a cute way of taking a hopeful stab in the dark that paid off. By hiding all the "mistake" cases from view—the trials that didn't pan out—you create a "miracle."

Evolution works the same way: all the dumb mistakes tend to be invisible, so all we see is a stupendous string of triumphs. For instance, the vast majority—way over 90 percent—of all the creatures that have ever lived died childless, but *not a single one of your ancestors* suffered that fate. Talk about a line of charmed lives!

One big difference between the discipline of science and the disci-

Copyrighted material

pline of stage magic is that while magicians conceal their false starts from the audience as best they can, in science you make your mistakes in public. You show them off so that everybody can learn from them. This way, you get the benefit of everybody else's experience, and not just your own idiosyncratic path through the space of mistakes. (The physicist Wolfgang Pauli famously expressed his contempt for the work of a colleague as "not even wrong." A clear falsehood shared with critics is better than vague mush.) This, by the way, is another reason why we humans are so much smarter than every other species. It is not so much that our brains are bigger or more powerful, or even that we have the knack of reflecting on our own past errors, but that we *share the benefits* that our individual brains have won by their individual histories of trial and error.*

I am amazed at how many really smart people don't understand that you can make big mistakes in public and emerge none the worse for it. I know distinguished researchers who will go to preposterous lengths to avoid having to acknowledge that they were wrong about something. They have never noticed, apparently, that the earth does not swallow people up when they say, "Oops, you're right. I guess I made a mistake." Actually, people love it when somebody admits to making a mistake. All kinds of people love pointing out mistakes. Generous-spirited people appreciate your giving them the opportunity to help, and acknowledging it when they succeed in helping you; mean-spirited people enjoy showing you up. Let them! Either way we all win.

Of course, in general, people do not enjoy correcting the *stupid* mistakes of others. You have to have something *worth* correcting, something original to be right or wrong about, something that

* That is the ideal, but we don't always live up to it, human nature being what it is. One of the recognized but unsolved problems with current scientific practice is that negative results—experiments that didn't uncover what they were designed to uncover—are not published often enough. This flaw in the system is famously explored and deplored in Feynman's "Cargo Cult Lecture," a commencement address he gave at Caltech in 1974, reprinted in Feynman, 1985.

Copyrighted material

requires constructing the sort of pyramid of risky thinking we saw in the card magician's tricks. Carefully building on the works of others, you can get yourself cantilevered out on a limb of your own. And then there's a surprise bonus: if you are one of the big risk-takers, people will get a kick out of correcting your occasional stupid mistakes, which show that you're not so special, you're a regular bungler like the rest of us. I know extremely careful philosophers who have never—apparently—made a mistake in their work. They tend not to get a whole lot accomplished, but what little they produce is pristine, if not venturesome. Their specialty is pointing out the mistakes of others, and this can be a valuable service, but nobody excuses *their* minor errors with a friendly chuckle. It is fair to say, unfortunately, that their best work often gets overshadowed and neglected, drowned out by the passing bandwagons driven by bolder thinkers. In chapter 76 we'll see that the generally good practice of making bold mistakes has other unfortunate side effects as well. Meta-advice: don't take *any* advice too seriously!

Copyrighted material

2. "BY PARODY OF REASONING": USING *REDUCTIO AD ABSURDUM*

The crowbar of rational inquiry, the great lever that enforces consistency, is *reductio ad absurdum*—literally, reduction (of the argument) to absurdity. You take the assertion or conjecture at issue and see if you can pry any contradictions (or just preposterous implications) out of it. If you can, that proposition has to be discarded or sent back to the shop for retooling. We do this all the time without bothering to display the underlying logic: "If that's a bear, then bears have antlers!" or "He won't get here in time for supper unless he can fly like Superman." When the issue is a tricky theoretical controversy, the crowbar gets energetically wielded, but here the distinction between fair criticism and refutation by caricature is hard to draw. Can your opponent really be so stupid as to believe the proposition you have just reduced to absurdity with a few deft moves? I once graded a student paper that had a serendipitous misspelling, replacing "parity" with "parody," creating the delicious phrase "by parody of reasoning," a handy name, I think, for misbegotten *reductio ad absurdum* arguments, which are all too common in the rough-and-tumble of scientific and philosophical controversy.

I recall attending a seminar on cognitive science at MIT some years ago, conducted by the linguist Noam Chomsky and the philosopher Jerry Fodor, in which the audience was regularly regaled with hilarious refutations of cognitive scientists from elsewhere who did not meet with their approval. On this day, Roger Schank, the director of Yale University's artificial intelligence laboratory, was the *bête noir*, and if you went by Chomsky's version, Schank had to be some kind of flaming idiot. I knew Roger and his work pretty well, and though I had disagreements of my own with it, I thought that Noam's version was hardly recognizable, so I raised my hand and suggested that

Copyrighted material

perhaps he didn't appreciate some of the subtleties of Roger's position. "Oh no," Noam insisted, chuckling. "This is what he holds!" And he went back to his demolition job, to the great amusement of those in the room. After a few more minutes of this I intervened again. "I have to admit," I said, "that the views you are criticizing are simply preposterous," and Noam grinned affirmatively, "but then what I want to know is why you're wasting your time and ours criticizing such junk." It was a pretty effective pail of cold water.

What about my own *reductios* of the views of others? Have they been any fairer? Here are a few to consider. You decide. The French neuroscientist Jean-Pierre Changeux and I once debated neuroscientist Sir John Eccles and philosopher Sir Karl Popper about consciousness and the brain at a conference in Venice. Changeux and I were the materialists (who maintain that the mind *is* the brain), and Popper and Eccles the dualists (who claim that a mind is not a material thing like a brain, but some other, second kind of entity that interacts with the brain). Eccles had won the Nobel Prize many years earlier for the discovery of the synapse, the microscopic gap between neurons that glutamate molecules and other neurotransmitters and neuromodulators cross trillions of times a day. According to Eccles, the brain was like a mighty pipe organ and the trillions of synapses composed the keyboards. The immaterial mind—the immortal soul, according to Eccles, a devout Catholic—played the synapses by somehow encouraging quantum-level nudges of the glutamate molecules. "Forget all that theoretical discussion of neural networks and the like; it's irrelevant rubbish," he said. "The mind is in the glutamate!" When it was my turn to speak, I said I wanted to be sure I had understood his position. If the mind was in the glutamate and I poured a bowl of glutamate down the drain, would that not be murder? "Well," he replied, somewhat taken aback, "it would be very hard to tell, wouldn't it?"*

* My other indelible memory of that conference was of Popper's dip in the Grand Canal. He slipped getting out of the motorboat at the boathouse of the Isola di San Giorgio and fell

Copyrighted material

You would think that Sir John Eccles, the Catholic dualist, and Francis Crick, the atheist materialist, would have very little in common, aside from their Nobel Prizes. But at least for a while their respective views of consciousness shared a dubious oversimplification. Many nonscientists don't appreciate how wonderful oversimplifications can be in science; they can cut through the hideous complexity with a working model that is *almost* right, postponing the messy details until later. Arguably the best use of "over"-simplification in the history of science was the end run by Crick and James Watson to find the structure of DNA while Linus Pauling and others were trudging along trying to make sense of all the details. Crick was all for trying the bold stroke just in case it solved the problem in one fell swoop, but of course that doesn't always work. I was once given the opportunity to demonstrate this at one of Crick's famous teas at La Jolla. These afternoon sessions were informal lab meetings where visitors could raise issues and participate in the general discussion. On this particular occasion Crick made a bold pronouncement: it had recently been shown that neurons in cortical area V_4 "cared about" (responded differentially to) color. And then he proposed a strikingly simple hypothesis: the conscious experience of red, for instance, *was* activity in the relevant red-sensitive neurons of that retinal area. *Hmm*, I wondered. "Are you saying, then, that if we were to remove some of those red-sensitive neurons and keep them alive in a petri dish, and stimulate them with a microelectrode, there would be *consciousness of red* in the petri dish?" One way of responding to a

feet first into the canal, submerged up to his knees before being plucked out and set on the pier by two nimble boatmen. The hosts were mortified and ready to rush back to the hotel to get nonagenarian Sir Karl a dry pair of trousers, but the pants he was wearing was the only pair he'd brought—and he was scheduled to lead off the conference in less than half an hour! Italian ingenuity took over, and within about five minutes I enjoyed an unforgettable sight: Sir Karl, sitting regally on a small chair in the exact middle of a marble-floored, domed room (Palladio designed it) surrounded by at least half a dozen young women in miniskirts, on their knees, plying his trouser legs with their hairdryers. The extension cords stretched radially to the walls, making of the tableau a sort of multicolored human daisy, with Sir Karl, unperturbed but unsmiling, in the center. Fifteen minutes later he was dry and pounding his fist on the podium to add emphasis to his dualistic vision.

Copyrighted material

proffered *reductio* is to grasp the nettle and endorse the conclusion, a move I once dubbed *outsmarting*, since the Australian philosopher J. J. C. Smart was famous for saying that yes, according to his theory of ethics, it *was* sometimes right to frame and hang an innocent man! Crick decided to outsmart me. "Yes! It would be an isolated instance of consciousness of red!" Whose consciousness of red? He didn't say. He later refined his thinking on this score, but still, he and neuroscientist Christof Koch, in their quest for what they called the NCC (the neural correlates of consciousness), never quite abandoned their allegiance to this idea.

Perhaps yet another encounter will bring out better what is problematic about the idea of a smidgen of consciousness in a dish. The physicist and mathematician Roger Penrose and the anesthesiologist Stuart Hameroff teamed up to produce a theory of consciousness that depended, not on glutamate, but on quantum effects in the microtubules of neurons. (Microtubules are tubular protein chains that serve as girders and highways inside the cytoplasm of all cells, not just neurons.) At Tucson II, the second international conference on the science of consciousness, after Hameroff's exposition of this view, I asked from the audience, "Stuart, you're an anesthesiologist; have you ever assisted in one of those dramatic surgeries that replaces a severed hand or arm?" No, he had not, but he knew about them. "Tell me if I'm missing something, Stuart, but given your theory, if you were the anesthesiologist in such an operation you would feel morally obliged to anesthetize the severed hand as it lay on its bed of ice, right? After all, the microtubules in the nerves of the hand would be doing their thing, just like the microtubules in the rest of the nervous system, and that hand would be in great pain, would it not?" The look on Stuart's face suggested that this had never occurred to him. The idea that consciousness (of red, of pain, of anything) is some sort of network property, something that involves coordinated activities in myriads of neurons, initially may not be very attractive, but these attempts at *reductios* may help people see why it should be taken seriously.

Copyrighted material

3. RAPOPORT'S RULES

Just how charitable are you supposed to be when criticizing the views of an opponent? If there are *obvious* contradictions in the opponent's case, then of course you should point them out, forcefully. If there are somewhat hidden contradictions, you should carefully expose them to view—and then dump on them. But the search for hidden contradictions often crosses the line into nitpicking, sea-lawyering,* and—as we have seen—outright parody. The thrill of the chase and the conviction that your opponent *has* to be harboring a confusion somewhere encourages uncharitable interpretation, which gives you an easy target to attack. But such easy targets are typically irrelevant to the real issues at stake and simply waste everybody's time and patience, even if they give amusement to your supporters. The best antidote I know for this tendency to caricature one's opponent is a list of rules promulgated many years ago by the social psychologist and game theorist Anatol Rapoport (creator of the winning Tit-for-Tat strategy in Robert Axelrod's legendary prisoner's dilemma tournament).†

How to compose a successful critical commentary:

1. You should attempt to re-express your target's position so clearly, vividly, and fairly that your target says, "Thanks, I wish I'd thought of putting it that way."

---

* Maritime law is notoriously complicated, strewn with hidden traps and escape clauses that only an expert, a *sea lawyer*, can keep track of, so *sea-lawyering* is using technicalities to evade responsibility or assign blame to others.

† The Axelrod tournament (Axelrod and Hamilton, 1981; Axelrod, 1984) opened up the blossoming field of theoretical research on the evolution of altruism. I give an introductory account in *Darwin's Dangerous Idea* (Dennett, 1995, pp. 479–480), and in more recent times there has been an explosion of variations, both simulations and experiments, in laboratories around the world. Rapoport's wonderfully simple implementation of the idea "I won't hit you if you don't hit me" is the seed from which all the later studies and models have grown.

Copyrighted material

2. You should list any points of agreement (especially if they are not matters of general or widespread agreement).
3. You should mention anything you have learned from your target.
4. Only then are you permitted to say so much as a word of rebuttal or criticism.

One immediate effect of following these rules is that your targets will be a receptive audience for your criticism: you have already shown that you understand their positions as well as they do, and have demonstrated good judgment (you agree with them on some important matters and have even been persuaded by something they said).*

Following Rapoport's Rules is always, for me at least, something of a struggle. Some targets, quite frankly, don't deserve such respectful attention, and—I admit—it can be sheer joy to skewer and roast them. But when it is called for, and it works, the results are gratifying. I was particularly diligent in my attempt to do justice to Robert Kane's (1996) brand of incompatibilism (a view about free will with which I profoundly disagree) in my book *Freedom Evolves* (2003), and I treasure the response he wrote to me after I had sent him the draft chapter:

> . . . In fact, I like it a lot, our differences notwithstanding. The treatment of my view is extensive and generally fair, far more so than one usually gets from critics. You convey the complexity of my view and the seriousness of my efforts to address difficult questions rather than merely sweeping them

* The formulation of Rapoport's Rules here is my own, done from memory of correspondence with Rapoport many years ago, now apparently lost. Samuel Ruth recently pointed out to me that the original source of Rapoport's Rules is in his book *Fights, Games, and Debates* (1960) and his paper "Three Modes of Conflict" (1961), which articulates rule 1, attributing it to Carl Rogers, and variations on the rest of the rules. My version is somewhat more portable and versatile.

Copyrighted material

> under the rug. And for this, as well as the extended treatment, I am grateful.

Other recipients of my Rapoport-driven attention have been less cordial. The fairer the criticism seems, the harder to bear in some cases. It is worth reminding yourself that a heroic attempt to find a defensible interpretation of an author, if it comes up empty, can be even more devastating than an angry hatchet job. I recommend it.

Copyrighted material

4. STURGEON'S LAW

The science-fiction author Ted Sturgeon, speaking at the World Science Fiction Convention in Philadelphia in September 1953, said,

> When people talk about the mystery novel, they mention *The Maltese Falcon* and *The Big Sleep*. When they talk about the western, they say there's *The Way West* and *Shane*. But when they talk about science fiction, they call it "that Buck Rogers stuff," and they say "ninety percent of science fiction is crud." Well, they're right. Ninety percent of science fiction is crud. But then ninety percent of everything is crud, and it's the ten percent that isn't crud that is important, and the ten percent of science fiction that isn't crud is as good as or better than anything being written anywhere.

Sturgeon's Law is usually put a little less decorously: *Ninety percent of everything is crap*. Ninety percent of experiments in molecular biology, 90 percent of poetry, 90 percent of philosophy books, 90 percent of peer-reviewed articles in mathematics—and so forth—is crap. Is that true? Well, maybe it's an exaggeration, but let's agree that there is a lot of mediocre work done in every field. (Some curmudgeons say it's more like 99 percent, but let's not get into that game.) A good moral to draw from this observation is that when you want to criticize a field, a genre, a discipline, an art form, . . . *don't waste your time and ours hooting at the crap!* Go after the good stuff, or leave it alone. This advice is often ignored by ideologues intent on destroying the reputation of analytic philosophy, evolutionary psychology, sociology, cultural anthropology, macroeconomics, plastic surgery, improvisational theater, television sitcoms, philosophical theology, massage therapy, you name it. Let's stipulate at the outset that there

Copyrighted material

is a great deal of deplorable, stupid, second-rate stuff out there, of all sorts. Now, in order not to waste your time and try our patience, make sure you concentrate on the best stuff you can find, the flagship examples extolled by the leaders of the field, the prize-winning entries, not the dregs. Notice that this is closely related to Rapoport's Rules: unless you are a comedian whose main purpose is to make people laugh at ludicrous buffoonery, spare us the caricature. This is particularly true, I find, when the target is philosophers. The very best theories and analyses of *any* philosopher, from the greatest, most perceptive sages of ancient Greece to the intellectual heroes of the recent past (Bertrand Russell, Ludwig Wittgenstein, John Dewey, Jean Paul Sartre—to name four very different thinkers), can be made to look like utter idiocy—or tedious nitpicking—with a few deft tweaks. Yuck, yuck. Don't do it. The only one you'll discredit is yourself.

Copyrighted material

5. OCCAM'S RAZOR

Attributed to William of Ockham (or Occam), the fourteenth-century logician and philosopher, this thinking tool is actually a much older rule of thumb. A Latin name for it is *lex parsimoniae*, the law of parsimony. It is usually put into English as the maxim "Do not multiply entities beyond necessity." The idea is straightforward: don't concoct a complicated, extravagant theory if you've got a simpler one (containing fewer ingredients, fewer entities) that handles the phenomenon just as well. If exposure to extremely cold air can account for all the symptoms of frostbite, don't postulate unobserved "snow germs" or "arctic microbes." Kepler's laws explain the orbits of the planets; we have no need to hypothesize pilots guiding the planets from control panels hidden under the surface. This much is uncontroversial, but extensions of the principle have not always met with agreement.

Conwy Lloyd Morgan, a nineteenth-century British psychologist, extended the idea to cover attributions of mentality to animals. Lloyd Morgan's Canon of Parsimony advises us not to attribute fancy minds to insects, fish, and even dolphins, dogs, and cats if their behavior can be explained in simpler terms:

> In no case is an animal activity to be interpreted in terms of higher psychological processes, if it can be fairly interpreted in terms of processes which stand lower in the scale of psychological evolution and development. [1894, p. 128]

Overused, this can be seen as enjoining us to treat all animals and even human beings as having brains but no minds. As we shall see, the tensions that arise when minds are the topic are not well settled by absolute prohibitions.

Copyrighted material

One of the least impressive attempts to apply Occam's Razor to a gnarly problem is the claim (and provoked counterclaims) that postulating a God as creator of the universe is simpler, more parsimonious, than the alternatives. How could postulating something supernatural and incomprehensible be parsimonious? It strikes me as the height of extravagance, but perhaps there are clever ways of rebutting that suggestion. I don't want to argue about it; Occam's Razor is, after all, just a rule of thumb, a frequently useful suggestion. The prospect of turning it into a Metaphysical Principle or Fundamental Requirement of Rationality that could bear the weight of proving or disproving the existence of God in one fell swoop is simply ludicrous. It would be like trying to disprove a theorem of quantum mechanics by showing that it contradicted the axiom "Don't put all your eggs in one basket."

Some thinkers have carried Occam's Razor to drastic extremes, using it to deny the existence of time, matter, numbers, holes, dollars, software, and so on. One of the earliest ultra-stingy thinkers was the ancient Greek philosopher Parmenides, whose catalogue of existing things was minimal indeed. As a student of mine memorably wrote on an exam, "Parmenides is the one who said, 'There's just one thing—and I'm not it.'" I hate to say it, but that does seem to be what Parmenides was trying to tell us. No doubt it loses something in translation. We philosophers get used to taking such ideas seriously if only because we never can tell when a "crazy" idea is going to turn out to be unfairly and unwisely judged, a victim of failure of imagination.

Copyrighted material

6. OCCAM'S BROOM

The molecular biologist Sidney Brenner recently invented a delicious play on Occam's Razor, introducing the new term *Occam's Broom*, to describe the process in which inconvenient facts are whisked under the rug by intellectually dishonest champions of one theory or another. This is our first boom crutch, an *anti-thinking* tool, and you should keep your eyes peeled for it. The practice is particularly insidious when used by propagandists who direct their efforts at the lay public, because like Sherlock Holmes's famous clue about the dog that *didn't* bark in the night, the *absence* of a fact that has been swept off the scene by Occam's Broom is unnoticeable except by experts. For instance, creationists invariably leave out the wealth of embarrassing evidence that their "theories" can't handle, and to a nonbiologist their carefully crafted accounts can be quite convincing simply because the lay reader can't see what *isn't there*.

How on earth can you keep on the lookout for something invisible? Get some help from the experts. Stephen C. Meyer's *Signature in the Cell* (2009) purports to expose the systematic impossibility of life having a natural (nonsupernatural) origin, and gives what seems—even to a relatively well-informed reader—to be a fair and exhaustive survey of the theories and models being worked on around the world, showing how irredeemably hopeless they all are. So persuasive is Meyer's case that in November 2009 the eminent philosopher Thomas Nagel declared it his Best Book of the Year in London's *Times Literary Supplement*, one of the world's most influential publications of book reviews! In a spirited correspondence I had with him after his rave appeared, he demonstrated that he knew quite a lot about the history of work on the origin of life, enough to think he could trust his own judgment. And as he noted in a letter to the *Times Literary Supplement* (January 1, 2010), "Meyer's book seems

Copyrighted material

to me to be written in good faith." Had Nagel consulted with scientists working in the field, he would have been able to see Meyer's exploitation of Occam's Broom, whisking inconvenient facts out of view, and he might also have been dismayed to learn that the experts hadn't been sent an early copy of Meyer's book, as he had, or been asked to referee it before publication. Learning that the book he admired was a stealth operation might have shaken his confidence in his judgment, or it might not have. The scientific establishment has been known to squelch renegade critics unjustly on occasion, and perhaps—perhaps—Meyer had no choice but to launch a sneak attack. But Nagel would have been wise to explore this prospect warily before committing himself. It is fair to say that the scientists working on the origin of life do not yet have a secure and agreed-upon theory, but there is no dearth of candidates, an embarrassment of riches rather than an almost empty arena.

Conspiracy theorists are masters of Occam's Broom, and an instructive exercise on the Internet is to look up a new conspiracy theory, to see if you (a nonexpert on the topic) can find the flaws, before looking elsewhere on the web for the expert rebuttals. When Brenner coined the term, he wasn't talking about creationism and conspiracy theories; he was pointing out that in the heat of battle, even serious scientists sometimes cannot resist "overlooking" some data that seriously undermine their pet theory. It's a temptation to be resisted, no matter what.

Copyrighted material

7. USING LAY AUDIENCES AS DECOYS

One good way of preventing people from *inadvertently* wielding Occam's Broom is a technique that I have been recommending for years, and have several times put to the test—but never as ambitiously as I would like to do. Unlike the other practices I have been describing, this one takes time and money to do properly. I hope others will pursue this technique vigorously and report the results. I have decided to put it here because it addresses some of the same problems of communication that the other general tools confront.

In many fields—not just philosophy—there are controversies that seem never-ending and partly artifactual: people are talking past one another and not making the necessary effort to communicate effectively. Tempers flare, and disrespect and derision start creeping in. People on the sidelines take sides, even when they don't fully understand the issues.

It can get ugly, and it *can* have a very straightforward cause. When experts talk to experts, whether they are in the same discipline or not, they always err on the side of *under-explaining*. The reason is not far to seek: to overexplain something to a fellow expert is a very serious insult—"Do I have to spell it out for you?"—and nobody wants to insult a fellow expert. So just to be safe, people err on the side of under-explaining. It is not done deliberately, for the most part, and it is almost impossible to keep from doing—which is actually a good thing, since being polite in an unstudied way is a nice character trait in anyone. But this gracious disposition to assume more understanding than is apt to be present in one's distinguished audience has an unfortunate by-product: experts often talk past each other.

There is no *direct* cure: entreating all the experts present at a work-

Copyrighted material

shop or conference not to under-explain their positions may be met by earnest promises, but it won't work. If anything it will make matters worse since now people will be particularly sensitive to the issue of inadvertently insulting somebody. But there is an indirect and quite effective cure: have all experts present their views to a small audience of curious nonexperts (here at Tufts I have the advantage of bright undergraduates) while the other experts listen in from the sidelines. They don't have to eavesdrop; this isn't a *devious* suggestion. On the contrary, everybody can and should be fully informed that the point of the exercise is to make it comfortable for participants to speak in terms that *everybody* will understand. By addressing their remarks to the undergraduates (the decoy audience), speakers need not worry *at all* about insulting the experts because they are not addressing the experts. (I suppose they might worry about insulting the undergraduates, but that's another matter.) When all goes well, expert A explains the issues of the controversy to the undergraduates while expert B listens. At some point B's face may light up. "So *that's* what you've been trying to say! Now I get it." Or maybe the good effects will have to wait until it is B's turn to explain to the same undergraduates what the issues are, and provoking just such a welcome reaction in A. It may not go perfectly, but it usually goes well and everybody benefits. The experts dissolve some of the artifactual misunderstandings between their positions, and the undergraduates get a first-rate educational experience.

Several times I have set up such exercises at Tufts, thanks to generous support from the administration. I handpick a small group of undergraduates (less than a dozen) and brief them on their role: they are not to accept anything they don't understand. They will be expected to raise their hands, to interrupt, to alert the experts to anything they find confusing or vague. (They do get required reading to pore over beforehand so that they are not utter novices on the topic; they are interested amateurs.) They love the role, and so they should; they are being given made-to-order tutorials from some big guns. The experts, meanwhile, often find that being set the task (well in advance)

Copyrighted material

to explain their position under these conditions helps them find better ways of making their points than they had ever found before. Sometimes these experts have been "protected" for years by layers of fellow experts, postdocs, and advanced graduate students, and they really need the challenge.

Copyrighted material

8. JOOTSING

It is hard to *find* an application of Occam's Broom, since it operates by whisking inconvenient facts out of sight, and it is even harder to achieve what Doug Hofstadter (1979, 1985) calls *jootsing*, which stands for "*j*umping *out of t*he *s*ystem." This is an important tactic not just in science and philosophy, but also in the arts. Creativity, that ardently sought but only rarely found virtue, often is a heretofore unimagined violation of the rules of the system from which it springs. It might be the system of classical harmony in music, the rules for meter and rhyme in sonnets (or limericks, even), or the "canons" of taste or good form in some genre of art. Or it might be the assumptions and principles of some theory or research program. Being creative is not just a matter of casting about for something novel—anybody can do that, since novelty can be found in any random juxtaposition of stuff—but of making the novelty *jump* out of some *system*, a system that has become somewhat established, for good reasons. When an artistic tradition reaches the point where literally "anything goes," those who want to be creative have a problem: there are no fixed rules to rebel against, no complacent expectations to shatter, nothing to subvert, no background against which to create something that is both surprising and yet meaningful. It helps to *know* the tradition if you want to subvert it. That's why so few dabblers or novices succeed in coming up with anything truly creative.

Sit down at a piano and try to come up with a good new melody and you soon discover how hard it is. All the keys are available, in any combination you choose, but until you can find something to lean on, some style or genre or pattern to lay down and exploit a bit, or allude to, before you twist it, you will come up with nothing but noise. And not just any violation of the rules will do the trick. I know there are at least two flourishing—well, surviving—jazz harpists, but

Copyrighted material