

Current Topics of Business Policy

(E/A/M 21871, IBE 21149)

Benito Arruñada
Universitat Pompeu Fabra

Readings

Table of contents

1. Behavior — Due: week 1
2. Contracting — Due: week 2
3. Public sector reform — Due: week 3
4. Other topics — Due: week 4
5. Professional career — Due: TBA
6. E-business — Due: 1st class on e-business
7. Tools — Due: project preparation

1. Behavior — Due: week 1

References:

Pinker, Steven (1997), “Standard Equipment,” chapter 1 of *How the Mind Works*, Norton, New York, 3-58.

Stark, Rodney (1996), “Conversion and Christian Growth,” in *The Rise of Christianity: A Sociologist Reconsiders History*, Princeton University Press, Princeton, 2-27.

Pinker, Steven (2002), *The Blank Slate: The Modern Denial of Human Nature*, Viking, New York, pp. on stereotypes (201-207).

Mullainathan, Sendhil, and Andrei Shleifer (2005), “The Market for News,” *American Economic Review*, 95(4), 1031-53.

Luscombe, Belinda (2013), “Confidence Woman,” *Time*, March 7.

HOW THE MIND WORKS

Steven Pinker



PENGUIN BOOKS

PENGUIN BOOKS

Published by the Penguin Group

Penguin Books Ltd, 27 Wrights Lane, London W8 5TZ, England

Penguin Putnam Inc., 375 Hudson Street, New York, New York 10014, USA

Penguin Books Australia Ltd, Ringwood, Victoria, Australia

Penguin Books Canada Ltd, 10 Alcorn Avenue, Toronto, Ontario, Canada M4V 3B2

Penguin Books (NZ) Ltd, 182-190 Wairau Road, Auckland 10, New Zealand

Penguin Books Ltd, Registered Offices: Harmondsworth, Middlesex, England

First published in the USA by W. W. Norton 1997

First published in Great Britain by Allen Lane The Penguin Press 1998

Published in Penguin Books 1998

1 3 5 7 9 1 0 8 6 4 2

Copyright © Stephen Pinker, 1997

All rights reserved

The notices on page 627 constitute an extension of this copyright page

The moral right of the author has been asserted

Printed in England by Clays Ltd, St Ives pic

Except in the United States of America, this book is sold subject to the condition that it shall not, by way of trade or otherwise, be lent, re-sold, hired out, or otherwise circulated without the publisher's prior consent in any form of binding or cover other than that in which it is published and without a similar condition including this condition being imposed on the subsequent purchaser

1

STANDARD EQUIPMENT

Why are there so many robots in fiction, but none in real life? I would pay a lot for a robot that could put away the dishes or run simple errands. But I will not have the opportunity in this century, and probably not in the next one either. There are, of course, robots that weld or spray-paint on assembly lines and that roll through laboratory hallways; my question is about the machines that walk, talk, see, and think, often better than their human masters. Since 1920, when Karel Čapek coined the word *robot* in his play *R.U.R.*, dramatists have freely conjured them up: Speedy, Cutie, and Dave in Isaac Asimov's *I, Robot*, Robbie in *Forbidden Planet*, the flailing canister in *Lost in Space*, the daleks in *Dr. Who*, Rosie the Maid in *The Jetsons*, Nomad in *Star Trek*, Hymie in *Get Smart*, the vacant butlers and bickering haberdashers in *Sleeper*, R2D2 and C3PO in *Star Wars*, the Terminator in *The Terminator*, Lieutenant Commander Data in *Star Trek: The Next Generation*, and the wisecracking film critics in *Mystery Science Theater 3000*.

This book is not about robots; it is about the human mind. I will try to explain what the mind is, where it came from, and how it lets us see, think, feel, interact, and pursue higher callings like art, religion, and philosophy. On the way I will try to throw light on distinctively human quirks. Why do memories fade? How does makeup change the look of a face? Where do ethnic stereotypes come from, and when are they irrational? Why do people lose their tempers? What makes children bratty? Why do fools fall in love? What makes us laugh? And why do people believe in ghosts and spirits?

3

But the gap between robots in imagination and in reality is my starting point, for it shows the first step we must take in knowing Ourselves: appreciating the fantastically complex design behind feats of mental life we take for granted. The reason there are no humanlike robots is not that the very idea of a mechanical mind is misguided. It is that the engineering problems that we humans solve as we see and walk and plan and make it through the day are far more challenging than landing on the moon or sequencing the human genome. Nature, once again, has found ingenious solutions that human engineers cannot yet duplicate. When Hamlet says, "What a piece of work is a man! how noble in reason! how infinite in faculty! in form and moving how express and admirable!" we should direct our awe not at Shakespeare or Mozart or Einstein or Kareem Abdul-Jabbar but at a four-year old carrying out a request to put a toy on a shelf.

In a well-designed system, the components are black boxes that perform their functions as if by magic. That is no less true of the mind. The faculty with which we ponder the world has no ability to peer inside itself or our other faculties to see what makes them tick. That makes us the victims of an illusion: that our own psychology comes from some divine force or mysterious essence or almighty principle. In the Jewish legend of the Golem, a clay figure was animated when it was fed an inscription of the name of God. The archetype is echoed in many robot stories. The statue of Galatea was brought to life by Venus' answer to Pygmalion's prayers; Pinocchio was vivified by the Blue Fairy. Modern versions of the Golem archetype appear in some of the less fanciful stories of science. All of human psychology is said to be explained by a single, omnipotent cause: a large brain, culture, language, socialization, learning, complexity, self-organization, neural-network dynamics.

I want to convince you that our minds are not animated, by some godly vapor or single wonder principle. The mind, like the Apollo spacecraft, is designed to solve many engineering problems, and thus is packed with high-tech systems each contrived to overcome its own obstacles. I begin by laying out these problems, which are both design specs for a robot and the subject matter of psychology. For I believe that the discovery by cognitive science and artificial intelligence of the technical challenges overcome by our mundane mental activity is one of the great revelations of science, an awakening of the imagination comparable to learning that the universe is made up of billions of galaxies or that a drop of pond water teems with microscopic life.

THE ROBOT CHALLENGE

What does it take to build a robot? Let's put aside superhuman abilities like calculating planetary orbits and begin with the simple human ones: seeing, walking, grasping, thinking about objects and people, and planning how to act.

In movies we are often shown a scene from a robot's-eye view, with the help of cinematic conventions like fish-eye distortion or crosshairs. That is fine for us, the audience, who already have functioning eyes and brains. But it is no help to the robot's innards. The robot does not house an audience of little people—homunculi—gazing at the picture and telling the robot what they are seeing. If you could see the world through a robot's eyes, it would look not like a movie picture decorated with crosshairs but something like this:

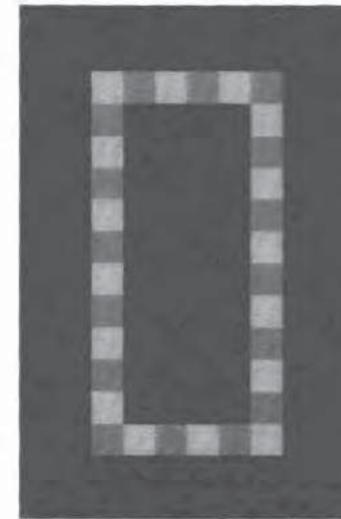
```

225 221 216 219 219 214 207 218 219 220 207 155 136 135
213 206 213 223 208 217 223 221 223 216 195 156 141 130
206 217 210 216 224 223 228 230 234 216 207 157 136 132
211 213 221 223 220 222 237 216 219 220 176 149 137 132
221 229 218 230 228 214 213 209 198 224 161 140 133 127
220 219 224 220 219 215 215 206 206 221 159 143 133 131
221 215 211 214 220 218 221 212 218 204 148 141 131 130
214 211 211 218 214 220 226 216 223 209 143 141 141 124
211 208 223 213 216 226 231 230 241 199 153 141 136 125
200 224 219 215 217 224 232 241 240 211 150 139 128 132
204 206 208 205 233 241 241 252 242 192 151 141 133 130
200 205 201 216 232 248 255 246 231 210 149 141 132 126
191 194 209 238 245 255 249 235 238 197 146 139 130 132
189 199 200 227 239 237 235 236 247 192 145 142 124 133
198 196 209 211 210 215 236 240 232 177 142 137 135 124
198 203 205 208 211 224 226 240 210 160 139 132 129 130
216 209 214 220 210 231 245 219 169 143 148 129 128 136
211 210 217 218 214 227 244 221 162 140 139 129 133 131
215 210 216 216 209 220 248 200 156 139 131 129 139 128
219 220 211 208 205 209 240 217 154 141 127 130 124 142
229 224 212 214 220 229 234 208 151 145 128 128 142 122
252 224 222 224 233 244 228 213 143 141 135 128 131 129
255 235 230 249 253 240 228 193 147 139 132 128 136 125
250 245 238 245 246 235 235 190 139 136 134 135 126 130
240 238 233 232 235 255 246 168 156 144 129 127 136 134

```

Each number represents the brightness of one of the millions of tiny patches making up the visual field. The smaller numbers come from darker patches, the larger numbers from brighter patches. The numbers shown in the array are the actual signals coming from an electronic camera trained on a person's hand, though they could just as well be the firing rates of some of the nerve fibers coming from the eye to the brain as a person looks at a hand. *Vox* a robot brain—or a human brain—to recognize objects and not bump into them, it must crunch these numbers and guess what kinds of objects in the world reflected the light that gave rise to them. The problem is humbly difficult.

First, a visual system must locate where an object ends and the backdrop begins. But the world is not a coloring book, with black outlines around solid regions. The world as it is projected into our eyes is a mosaic of tiny shaded patches. Perhaps, one could guess, the visual brain looks for regions where a quilt of large numbers (a brighter region) abuts a quilt of small numbers (a darker region). You can discern such a boundary in the square of numbers; it runs diagonally from the top right to the bottom center. Most of the time, unfortunately, you would not have found the edge of an object, where it gives way to empty space. The juxtaposition of large and small numbers could have come from many distinct arrangements of matter. This drawing, devised by the psychologists Pawan Sinha and Edward Adelson, appears to show a ring of light gray and dark gray tiles.



In fact, it is a rectangular cutout in a black cover through which you are looking at part of a scene. In the next drawing the cover has been removed, and you can see that each pair of side-by-side gray squares comes from a different arrangement of objects.



Big numbers next to small numbers can come from an object standing in front of another object, dark paper lying on light paper, a surface painted two shades of gray, two objects touching side by side, gray cellophane on a white page, an inside or outside corner where two walls meet, or a shadow. Somehow the brain must solve the chic ken-and-egg problem of identifying three-dimensional objects from the patches on the retina *and* determining what each patch is (shadow or paint, crease or overlay, clear or opaque) from knowledge of what object the patch is part of.

The difficulties have just begun. Once we have carved the visual world into objects, we need to know what they are made of, say, snow versus coal. At first glance the problem looks simple. If large numbers come from bright regions and small numbers come from dark regions, then large number equals white equals snow and small number equals black equals coal, right? Wrong. The amount of light hitting a spot on the retina depends not only on how pale or dark the object is but also on how bright or dim the light illuminating the object is. A photographer's light meter would show you that more light bounces off a lump of coal outdoors than off a snowball indoors. That is why people are so often disappointed by their snapshots and why photography is such a complicated craft. The camera does not lie; left to its own devices, it renders outdoor

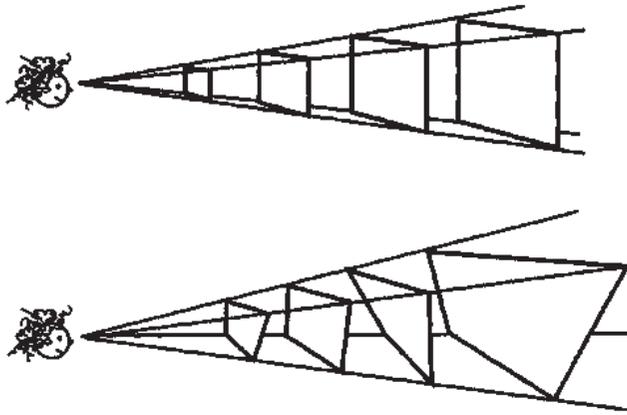
scenes as milk and indoor scenes as mud. Photographers, and sometimes microchips inside the camera, coax a realistic image out of the film with tricks like adjustable shutter timing, lens apertures, film speeds, flashes, and darkroom manipulations.

Our visual system does much better. Somehow it lets Us see the bright outdoor coal as black and the dark indoor snowball as white. That is a happy outcome, because our conscious sensation of color and lightness matches the world as it is rather than the world as it presents itself to the eye. The snowball is soft and wet and prone to melt whether it is indoors or out, and we see it as white whether it is indoors or out. The coal is always hard and dirty and prone to burn, and we always see it as black. The harmony between how the world *looks* and how the world *is* must be an achievement of our neural wizardry, because black and white don't simply announce themselves on the retina. In case you are still skeptical, here is an everyday demonstration. When a television set is off, the screen is a pale greenish gray. When it is on, some of the phosphor dots give off light, painting in the bright areas of the picture. But the other dots do not suck light and paint in the dark areas; they just stay gray. The areas that you see as black are in fact just the pale shade of the picture tube when the set was off. The blackness is a figment, a product of the brain circuitry that ordinarily allows you to see coal as coal. Television engineers exploited that circuitry when they designed the screen.

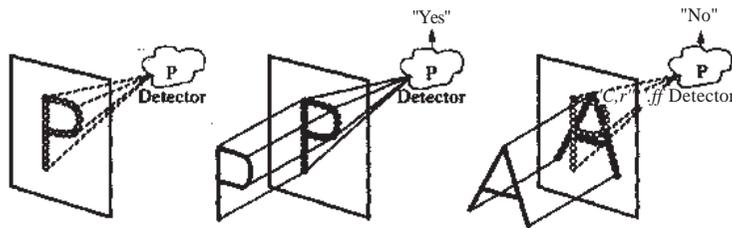
The next problem is seeing in depth. Our eyes squash the three-dimensional world into a pair of two-dimensional retinal images, and the third dimension must be reconstituted by the brain. But there are no telltale signs in the patches on the retina that reveal how far away a surface is. A stamp in your palm can project the same square on your retina as a chair across the room or a building miles away (first drawing, page 9). A cutting board viewed head-on can project the same trapezoid as various irregular shards held at a slant (second drawing, page 9).

You can feel the force of this fact of geometry, and of the neural mechanism that copes with it, by staring at a lightbulb for a few seconds or looking at a camera as the flash goes off, which temporarily bleaches a patch onto your retina. If you now look at the page in front of you, the afterimage adheres to it and appears to be an inch or two across. If you look up at the wall, the afterimage appears several feet long. If you look at the sky, it is the size of a cloud.

Finally, how might a vision module recognize the objects out there in the world, so that the robot can name them or recall what they do? The



obvious solution is to build a template or cutout for each object that duplicates its shape. When an object appears, its projection on the retina would fit its own template like a round peg in a round hole. The template would be labeled with the name of the shape—in this case, "the letter P"—and whenever a shape matches it, the template announces the name:



Alas, this simple device malfunctions in both possible ways. It sees P's that aren't there; for example, it gives a false alarm to the R shown in the first square below. And it fails to see P's that are there; for example, it misses the letter when it is shifted, tilted, slanted, too far, too near, or too fancy:



And these problems arise with a nice, crisp letter of the alphabet. Imagine trying to design a recognizer for a shirt, or a face! To be sure, after four decades of research in artificial intelligence, the technology of shape recognition has improved. You may own software that scans in a page, recognizes the printing, and converts it with reasonable accuracy to a file of bytes. But artificial shape recognizers are still no match for the ones in our heads. The artificial ones are designed for pristine, easy-to-recognize worlds and not the squishy, jumbled real world. The funny numbers at the bottom of checks were carefully drafted to have shapes that don't overlap and are printed with special equipment that positions them exactly so that they can be recognized by templates. When the first face recognizers are installed in buildings to replace doormen, they will not even try to interpret the chiaroscuro of your face but will scan in the hard-edged, rigid contours of your iris or your retinal blood vessels. Our brains, in contrast, keep a record of the shape of every face we know (and every letter, animal, tool, and so on), and the record is somehow matched with a retinal image even when the image is distorted in all the ways we have been examining. In Chapter 4 we will explore how the brain accomplishes this magnificent feat.

Let's take a look at another everyday miracle: getting a body from place to place. When we want a machine to move, we put it on wheels. The invention of the wheel is often held up as the proudest accomplishment of civilization. Many textbooks point out that no animal has evolved wheels and cite the fact as an example of how evolution is often incapable of finding the optimal solution to an engineering problem. But it is not a good example at all. Even if nature *could* have evolved a moose on wheels, it surely would have opted not to. Wheels are good only in a world with roads and rails. They bog down in any terrain that is soft, slippery, steep, or uneven. Legs are better. Wheels have to roll along an unbroken supporting ridge, but legs can be placed on a series of separate footholds, an extreme example being a ladder. Legs can also be placed to minimize lurching and to step over obstacles. Even today, when it seems as if the world has become a parking lot, only about half of the earth's land is accessible to vehicles with wheels or tracks, but most of the earth's land is accessible to vehicles with feet: animals, the vehicles designed by natural selection.

But legs come with a high price: the software to control them. A wheel, merely by turning, changes its point of support gradually and can bear weight the whole time. A leg has to change its point of support all at once, and the weight has to be unloaded to do so. The motors controlling a leg have to alternate between keeping the foot on the ground while it bears and propels the load and taking the load off to make the leg free to move. All the while they have to keep the center of gravity of the body within the polygon defined by the feet so the body doesn't topple over. The controllers also must minimize the wasteful up-and-down motion that is the bane of horseback riders. In walking windup toys, these problems are crudely solved by a mechanical linkage that converts a rotating shaft into a stepping motion. But the toys cannot adjust to the terrain by finding the best footholds.

Even if we solved these problems, we would have figured out only how to control a walking insect. With six legs, an insect can always keep one tripod on the ground while it lifts the other tripod. At any instant, it is stable. Even four-legged beasts, when they aren't moving too quickly, can keep a tripod on the ground at all times. But as one engineer has put it, "the upright two-footed locomotion of the human being seems almost a recipe for disaster in itself, and demands a remarkable control to make it practicable." When we walk, we repeatedly tip over and break our fall in the nick of time. When we run, we take off in bursts of flight. These aerobatics allow us to plant our feet on widely or erratically spaced footholds that would not prop us up at rest, and to squeeze along narrow paths and jump over obstacles. But no one has yet figured out how we do it.

Controlling an arm presents a new challenge. Grab the shade of an architect's lamp and move it along a straight diagonal path from near you, low on the left, to far from you, high on the right. Look at the rods and hinges as the lamp moves. Though the shade proceeds along a straight line, each rod swings through a complicated arc, swooping rapidly at times, remaining almost stationary at other times, sometimes reversing from a bending to a straightening motion. Now imagine having to do it in reverse: without looking at the shade, you must choreograph the sequence of twists around each joint that would send the shade along a straight path. The trigonometry is frightfully complicated. But your arm is an architect's lamp, and your brain effortlessly solves the equations every time you point. And if you have ever held an architect's lamp by its clamp, you will appreciate that the problem is even harder than what I have described. The lamp flails under its weight as if it had a mind of its

own; so would your arm if your brain did not compensate for its weight, solving a near-intractable physics problem. I

A still more remarkable feat is controlling the hand. Nearly¹ two thousand years ago, the Greek physician Galen pointed out the exquisite natural engineering behind the human hand. It is a single tool that manipulates objects of an astonishing range of sizes, shapes, and weights, from a log to a millet seed. "Man handles them all," Galen noted, "as well as if his hands had been made for the sake of each one of them alone." The hand can be configured into a hook grip (to lift a pail), a scissors grip (to hold a cigarette), a five-jaw chuck (to lift a coaster), a three-jaw chuck (to hold a pencil), a two-jaw pad-to-pad chuck (to thread a needle), a two-jaw pad-to-side chuck (to turn a key), a squeeze grip (to hold a hammer), a disc grip (to open a jar), and a spherical grip (to hold a ball). Each grip needs a precise combination of muscle tensions that mold the hand into the right shape and keep it there as the load tries to bend it back. Think of lifting a milk carton. Too loose a grasp, and you drop it; too tight, and you crush it; and with some gentle rocking, you can even use the tugging on your fingertips as a gauge of how much milk is inside! And I won't even begin to talk about the tongue, a boneless water balloon controlled only by squeezing, which can loosen food from a back tooth or perform the ballet that articulates words like *thrilling* and *sixths*. I

"A common man marvels at uncommon things; a wise man marvels at the commonplace." Keeping Confucius' dictum in mind, let's continue to look at commonplace human acts with the fresh eye of a robot designer seeking to duplicate them. Pretend that we have somehow built a robot that can see and move. What will it do with what it sees? How should it decide how to act?

An intelligent being cannot treat every object it sees as a unique entity unlike anything else in the universe. It has to put objects in categories so that it may apply its hard-won knowledge about similar objects, encountered in the past, to the object at hand.

But whenever one tries to program a set of criteria to capture the members of a category, the category disintegrates. Leaving aside slippery concepts like "beauty" or "dialectical materialism," let's look at a textbook

example of a well-defined one: "bachelor." A bachelor, of course, is simply an adult human male who has never been married. But now imagine that a friend asks you to invite some bachelors to her party. What would happen if you used the definition to decide which of the following people to invite?

Arthur has been living happily with Alice for the last five years. They have a two-year-old daughter and have never officially married.

Bruce was going to be drafted, so he arranged with his friend Barbara to have a justice of the peace marry them so he would be exempt. They have never lived together. He dates a number of women, and plans to have the marriage annulled as soon as he finds someone he wants to marry.

Charlie is 17 years old. He lives at home with his parents and is in high school.

David is 17 years old. He left home at 13, started a small business, and is now a successful young entrepreneur leading a playboy's lifestyle in his penthouse apartment.

Eli and Edgar are homosexual lovers who have been living together for many years.

Faisal is allowed by the law of his native Abu Dhabi to have three wives. He currently has two and is interested in meeting another potential fiancée.

Father Gregory is the bishop of the Catholic cathedral at Groton upon Thames.

The list, which comes from the computer scientist Terry Winograd, shows that the straightforward definition of "bachelor" does not capture our intuitions about who fits the category.

Knowing who is a bachelor is just common sense, but there's nothing common about common sense. Somehow it must find its way into a human or robot brain. And common sense is not simply an almanac about life that can be dictated by a teacher or downloaded like an enormous database. No database could list all the facts we tacitly know, and no one ever taught them to us. You know that when Irving puts the dog in the car, it is no longer in the yard. When Edna goes to church, her head goes with her. If Doug is in the house, he must have gone in through some opening unless he was born there and never left. If Sheila is alive

at 9 A.M. and is alive at 5 P.M., she was also alive at noon. Zebras in the wild never wear underwear. Opening a jar of a new brand of peanut butter will not vaporize the house. People never shove meat thermometers in their ears. A gerbil is smaller than Mt. Kilimanjaro.

An intelligent system, then, cannot be stuffed with trillions of facts. It must be equipped with a smaller list of core truths and a set of rules to deduce their implications. But the rules of common sense, like the categories of common sense, are frustratingly hard to set down. Even the most straightforward ones fail to capture our everyday reasoning. Mavis lives in Chicago and has a son named Fred, and Millie lives in Chicago and has a son named Fred. But whereas the Chicago that Mavis lives in is the same Chicago that Millie lives in, the Fred who is Mavis' son is not the same Fred who is Millie's son. If there's a bag in your car, and a gallon of milk in the bag, there is a gallon of milk in your car. But if there's a person in your car, and a gallon of blood in a person, it would be strange to conclude that there is a gallon of blood in your car.

Even if you were to craft a set of rules that derived only sensible conclusions, it is no easy matter to use them all to guide behavior intelligently. Clearly a thinker cannot apply just one rule at a time. A match gives light; a saw cuts wood; a locked door is opened with a key. But we laugh at the man who lights a match to peer into a fuel tank, who saws off the limb he is sitting on, or who locks his keys in the car and spends the next hour wondering how to get his family out. A thinker has to compute not just the direct effects of an action but the side effects as well.

But a thinker cannot crank out predictions about *all* the side effects, either. The philosopher Daniel Dennett asks us to imagine a robot designed to fetch a spare battery from a room that also contained a time bomb. Version 1 saw that the battery was on a wagon and that if it pulled the wagon out of the room, the battery would come with it. Unfortunately, the bomb was also on the wagon, and the robot failed to deduce that pulling the wagon out brought the bomb out, too. Version 2 was programmed to consider all the side effects of its actions. It had just finished computing that pulling the wagon would not change the color of the room's walls and was proving that the wheels would turn more revolutions than there are wheels on the wagon, when the bomb went off. Version 3 was programmed to distinguish between relevant implications and irrelevant ones. It sat there cranking out millions of implications and putting all the relevant ones on a list of facts to consider and all the irrelevant ones on a list of facts to ignore, as the bomb ticked away.

An intelligent being has to deduce the implications of what it knows, but only the *relevant* implications. Dennett points out that this requirement poses a deep problem not only for robot design but for epistemology, the analysis of how we know. The problem escaped the notice of generations of philosophers, who were left complacent by the illusory effortlessness of their own common sense. Only when artificial intelligence researchers tried to duplicate common sense in computers, the ultimate blank slate, did the conundrum, now called "the frame problem," come to light. Yet somehow we all solve the frame problem whenever we use our common sense.

Imagine that we have somehow overcome these challenges and have a machine with sight, motor coordination, and common sense. Now we must figure out how the robot will put them to use. We have to give it motives.

What should a robot want? The classic answer is Isaac Asimov's Fundamental Rules of Robotics, "the three rules that are built most deeply into a robot's positronic brain."

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

Asimov insightfully noticed that self-preservation, that universal biological imperative, does not automatically emerge in a complex system. It has to be programmed in (in this case, as the Third Law). After all, it is just as easy to build a robot that lets itself go to pot or eliminates a malfunction by committing suicide as it is to build a robot that always looks out for Number One. Perhaps easier; robot-makers sometimes watch in horror as their creations cheerfully shear off limbs or flatten themselves against walls, and a good proportion of the world's most intelligent machines are kamikaze cruise missiles and smart bombs.

But the need for the other two laws is far from obvious. Why give a

robot an order to obey orders—why aren't the original orders enough? Why command a robot not to do harm—wouldn't it be easier never to command it to *do* harm in the first place? Does the universe contain a mysterious force pulling entities toward malevolence, so that a positronic brain must be programmed to withstand it? Do intelligent beings inevitably develop an attitude problem?

In this case Asimov, like generations of thinkers, like all of us, was unable to step outside his own thought processes and see them as artifacts of how our minds were put together rather than as inescapable laws of the universe. Man's capacity for evil is never far from our minds, and it is easy to think that evil just comes along with intelligence as part of its very essence. It is a recurring theme in our cultural tradition: Adam and Eve eating the fruit of the tree of knowledge, Promethean fire and Pandora's box, the rampaging Golem, Faust's bargain, the Sorcerer's Apprentice, the adventures of Pinocchio, Frankenstein's monster, the murderous apes and mutinous HAL of *2001: A Space Odyssey*. From the 1950s through the 1980s, countless films in the computer-runs-amok genre captured a popular fear that the exotic mainframes of the era would get smarter and more powerful and someday turn on us.

Now that computers really *have* become smarter and more powerful, the anxiety has waned. Today's ubiquitous, networked computers have an unprecedented ability to do mischief should they ever go to the bad. But the only mayhem comes from unpredictable chaos or from human malice in the form of viruses. We no longer worry about electronic serial killers or subversive silicon cabals because we are beginning to appreciate that malevolence—like vision, motor coordination, and common sense—does not come free with computation but has to be programmed in. The computer running WordPerfect on your desk will continue to fill paragraphs for as long as it does anything at all. Its software will not insidiously mutate into depravity like the picture of Dorian Gray.

Even if it could, why would it want to? To get—what? More floppy disks? Control over the nation's railroad system? Gratification of a desire to commit senseless violence against laser-printer repairmen? And wouldn't it have to worry about reprisals from technicians who with the turn of a screwdriver could leave it pathetically singing "A Bicycle Built for Two"? A network of computers, perhaps, could discover the safety in numbers and plot an organized takeover—but what would make one computer volunteer to fire the data packet heard round the world and risk early martyrdom? And what would prevent the coalition from being

undermined by silicon draft-dodgers and conscientious objectors? Aggression, like every other part of human behavior we take for granted, is a challenging engineering problem!

But then, so are the kinder, gentler motives. How would you design a robot to obey Asimov's injunction never to allow a human being to come to harm through inaction!⁵ Michael Frayn's 1965 novel *The Tin Men* is set in a robotics laboratory, and the engineers in the Ethics Wing, Macintosh, Goldwasser, and Sinson, are testing the altruism of their robots. They have taken a bit too literally the hypothetical dilemma in every moral philosophy textbook in which two people are in a lifeboat built for one and both will die unless one bails out. So they place each robot in a raft with another occupant, lower the raft into a tank, and observe what happens.

[The] first attempt, Samaritan I, had pushed itself overboard with great alacrity, but it had gone overboard to save anything which happened to be next to it on the raft, from seven stone of lima beans to twelve stone of wet seaweed. After many weeks of stubborn argument Macintosh had conceded that the lack of discrimination was unsatisfactory, and he had abandoned Samaritan I and developed Samaritan II, which would sacrifice itself only for an organism at least as complicated as itself.

The raft stopped, revolving slowly, a few inches above the water. "Drop it," cried Macintosh.

The raft hit the water with a sharp report. Sinson and Samaritan sat perfectly still. Gradually the raft settled in the water, until a thin tide began to wash over the top of it. At once Samaritan leaned forward and seized Sinson's head. In four neat movements it measured the size of his skull, then paused, computing. Then, with a decisive click, it rolled sideways off the raft and sank without hesitation to the bottom of the tank.

But as the Samaritan II robots came to behave like the moral agents in the philosophy books, it became less and less clear that they were really moral at all. Macintosh explained why he did not simply tie a rope around the self-sacrificing robot to make it easier to retrieve: "I don't want it to know that it's going to be saved. It would invalidate its decision to sacrifice itself. . . . So, every now and then I leave one of them in instead of fishing it out. To show the others I mean business. I've written off two this week." Working out what it would take to program goodness into a robot shows not only how much machinery it takes to be good but how slippery the concept of goodness is to start with.

And what about the most caring motive of all? The weak-willed com-

puters of 1960s pop culture were not tempted only by selfishness and power, as we see in the comedian Allan Sherman's song "Automation," sung to the tune of "Fascination":

It was automation, I know.
That was what was making the factory go.
It was IBM, it was Univac,
It was all those gears going clickety clack, dear.
I thought automation was keen
Till you were replaced by a ten-ton machine.
It was a computer that tore us apart, dear,
Automation broke my heart. . . .

*abit
dit
elung*

It was automation, I'm told,
That's why I got fired and I'm out in the cold.
How could I have known, when the 503
Started in to blink, it was winking at me, dear?
I thought it was just some mishap
When it sidled over and sat on my lap.
But when it said "I love you" and gave me a hug, dear,
That's when I pulled out . . . its . . . plug.

But for all its moonstruck madness, love is no bug or crash or malfunction. The mind is never so wonderfully concentrated as when it turns to love, and there must be intricate calculations that carry out the peculiar logic of attraction, infatuation, courtship, coyness, surrender, commitment, malaise, philandering, jealousy, desertion, and heartbreak. And in the end, as my grandmother used to say, every pot finds a cover; most people—including, significantly, all of our ancestors—manage to pair up long enough to produce viable children. Imagine how many lines of programming it would take to duplicate that!

Robot design is a kind of consciousness-raising. We tend to be blasé about our mental lives. We open our eyes, and familiar articles present themselves; we will our limbs to move, and objects and bodies float into place; we awaken from a dream, and return to a comfortably predictable

worldr Cupid draws back his bow, and lets his arrow go. But think of what it takes for a hunk of matter to accomplish these improbable outcomes, and you begin to see through the illusion. Sight and action and common sense and violence and morality and love are no accident, no inextricable ingredients of an intelligent essence, no inevitability of information processing. Each is a tour de force, wrought by a high level of targeted design. Hidden behind the panels of consciousness must lie fantastically complex machinery—optical analyzers, motion guidance systems, simulations of the world, databases on people and things, goal-schedulers, conflict-resolvers, and many others. Any explanation of how the mind works that alludes hopefully to some single master force or mind-bestowing elixir like "culture," "learning," or "self-organization" begins to sound hollow, just not up to the demands of the pitiless universe we negotiate so successfully.

The robot challenge hints at a mind loaded with original equipment, but it still may strike you as an argument from the armchair. Do we actually find signs of this intricacy when we look directly at the machinery of the mind and at the blueprints for assembling it? I believe we do, and what we see is as mind-expanding as the robot challenge itself.

When the visual areas of the brain are damaged, for example, the visual world is not simply blurred or riddled with holes. Selected aspects of visual experience are removed while others are left intact. Some patients see a complete world but pay attention only to half of it. They eat food from the right side of the plate, shave only the right cheek, and draw a clock with twelve digits squished into the right half. Other patients lose their sensation of color, but they do not see the world as an arty black-and-white movie. Surfaces look grimy and rat-colored to them, killing their appetite and their libido. Still others can see objects change their positions but cannot see them move—a syndrome that a philosopher once tried to convince me was logically impossible! The stream from a teapot does not flow but looks like an icicle; the cup does not gradually fill with tea but is empty and then suddenly full.

Other patients cannot recognize the objects they see: their world is like handwriting they cannot decipher. They copy a bird faithfully but identify it as a tree stump. A cigarette lighter is a mystery until it is lit. When they try to weed the garden, they pull out the roses. Some patients can recognize inanimate objects but cannot recognize faces. The patient deduces that the visage in the mirror must be his, but does not viscerally recognize himself. He identifies John F. Kennedy as Martin Luther King,

and asks his wife to wear a ribbon at a party so he can find her when it is time to leave. Stranger still is the patient who recognizes the face but not the person: he sees his wife as an amazingly convincing impostor.

These syndromes are caused by an injury, usually a stroke, to one or more of the thirty brain areas that compose the primate visual system. Some areas specialize in color and form, others in where an object is, others in what an object is, still others in how it moves. A seeing robot cannot be built with just the fish-eye viewfinder of the movies, and it is no surprise to discover that humans were not built that way either. When we gaze at the world, we do not fathom the many layers of apparatus that underlie our unified visual experience, until neurological disease dissects them for us.

Another expansion of our vista comes from the startling similarity between identical twins, who share the genetic recipes that build the mind. Their minds are astonishingly alike, and not just in gross measures like IQ and personality traits like neuroticism and introversion. They are alike in talents such as spelling and mathematics, in opinions on questions such as apartheid, the death penalty, and working mothers, and in their career choices, hobbies, vices, religious commitments, and tastes in dating. Identical twins are far more alike than fraternal twins, who share only half their genetic recipes, and most strikingly, they are almost as alike when they are reared apart as when they are reared together. Identical twins separated at birth share traits like entering the water backwards and only up to their knees, sitting out elections because they feel insufficiently informed, obsessively counting everything in sight, becoming captain of the volunteer fire department, and leaving little love notes around the house for their wives.

People find these discoveries arresting, even incredible. The discoveries cast doubt on the autonomous "I" that we all feel hovering above our bodies, making choices as we proceed through life and affected only by our past and present environments. Surely the mind does not come equipped with so many small parts that it could predestine us to flush the toilet before and after using it or to sneeze playfully in crowded elevators, to take two other traits shared by identical twins reared apart. But apparently it does. The far-reaching effects of the genes have been documented in scores of studies and show up no matter how one tests for them: by comparing twins reared apart and reared together, by comparing identical and fraternal twins, or by comparing adopted and biological children. And despite what critics sometimes claim, the effects are not

Concordance



products of coincidence, fraud, or subtle similarities in the family environments (such as adoption agencies striving to place identical twins in homes that both encourage walking into the ocean backwards). The findings, of course, can be misinterpreted in many ways, such as by imagining a gene for leaving little love notes around the house or by concluding that people are unaffected by their experiences. And because this research can measure only the ways in which people *differ*, it says little about the design of the mind that all normal people share. But by showing how many ways the mind can vary in its innate structure, the discoveries open our eyes to how much structure the mind must have.

REVERSE-ENGINEERING THE PSYCHE

The complex structure of the mind is the subject of this book. Its key idea can be captured in a sentence: The mind is a system of organs of computation, designed by natural selection to solve the kinds of problems our ancestors faced in their foraging way of life, in particular, understanding and outmaneuvering objects, animals, plants, and other people. The summary can be unpacked into several claims. The mind is what the brain does; specifically, the brain processes information, and thinking is a kind of computation. The mind is organized into modules or mental organs, each with a specialized design that makes it an expert in one arena of interaction with the world. The modules' basic logic is specified by our genetic program. Their operation was shaped by natural selection to solve the problems of the hunting and gathering life led by our ancestors in most of our evolutionary history. The various problems for our ancestors were subtasks of one big problem for their genes, maximizing the number of copies that made it into the next generation.

On this view, psychology is engineering in reverse. In forward-engineering, one designs a machine to do something; in reverse-engineering, one figures out what a machine was designed to do. Reverse-engineering is what the boffins at Sony do when a new product is announced by Panasonic, or vice versa. They buy one, bring it back to the lab, take a screwdriver to it, and try to figure out what all the parts are for and how they combine to make the device work. We all engage in reverse-engineering when we face an interesting new gadget. In rummaging through

an antique store, we may find a contraption that is inscrutable until we figure out what it was designed to do. When we realize that it is an olive-pitter, we suddenly understand that the metal ring is designed to hold the olive, and the lever lowers an X-shaped blade through one end, pushing the pit out through the other end. The shapes and arrangements of the springs, hinges, blades, levers, and rings all make sense in a satisfying rush of insight. We even understand why canned olives have an X-shaped incision at one end.

In the seventeenth century William Harvey discovered that veins had valves and deduced that the valves must be there to make the blood circulate. Since then we have understood the body as a wonderfully complex machine, an assembly of struts, ties, springs, pulleys, levers, joints, hinges, sockets, tanks, pipes, valves, sheaths, pumps, exchangers, and filters. Even today we can be delighted to learn what mysterious parts are for. Why do we have our wrinkled, asymmetrical ears? Because they filter sound waves coming from different directions in different ways. The nuances of the sound shadow tell the brain whether the source of the sound is above or below, in front of or behind us. The strategy of reverse-engineering the body has continued in the last half of this century as we have explored the nanotechnology of the cell and of the molecules of life. The stuff of life turned out to be not a quivering, glowing, wondrous gel but a contraption of tiny jigs, springs, hinges, rods, sheets, magnets, zippers, and trapdoors, assembled by a data tape whose information is copied, downloaded, and scanned.

The rationale for reverse-engineering living things comes, of course, from Charles Darwin. He showed how "organs of extreme perfection and complication, which justly excite our admiration" arise not from God's foresight but from the evolution of replicators over immense spans of time. As replicators replicate, random copying errors sometimes crop up, and those that happen to enhance the survival and reproduction rate of the replicator tend to accumulate over the generations. Plants and animals are replicators, and their complicated machinery thus appears to have been engineered to allow them to survive and reproduce.

Darwin insisted that his theory explained not just the complexity of an animal's body but the complexity of its mind. "Psychology will be based on a new foundation," he famously predicted at the end of *The Origin of Species*. But Darwin's prophecy has not yet been fulfilled. More than a century after he wrote those words, the study of the mind is still mostly Darwin-free, often defiantly so. Evolution is said to be irrelevant,

sinful, or fit only for speculation over a beer at the end of the day. The allergy to evolution in the social and cognitive sciences has been, I think, a barrier to understanding. The mind is an exquisitely organized system that accomplishes remarkable feats no engineer can duplicate. How could the forces that shaped that system, and the purposes for which it was designed, be irrelevant to understanding it? Evolutionary thinking is indispensable, not in the form that many people think of—dreaming up missing links or narrating stories about the stages of Man—but in the form of careful reverse-engineering. Without reverse-engineering we are like the singer in Tom Paxton's "The Marvelous Toy," reminiscing about a childhood present: "It went ZIP! when it moved, and POP! when it stopped, and WHIRRR! when it stood still; I never knew just what it was, and I guess I never will."

Only in the past few years has Darwin's challenge been taken up, by a new approach christened "evolutionary psychology" by the anthropologist John Tooby and the psychologist Leda Cosmides. Evolutionary psychology brings together two scientific revolutions. One is the cognitive revolution of the 1950s and 1960s, which explains the mechanics of thought and emotion in terms of information and computation. The other is the revolution in evolutionary biology of the 1960s and 1970s, which explains the complex adaptive design of living things in terms of selection among replicators. The two ideas make a powerful combination. Cognitive science helps us to understand how a mind is possible and what kind of mind we have. Evolutionary biology helps us to understand *why* we have the kind of mind we have.

The evolutionary psychology of this book is, in one sense, a straightforward extension of biology, focusing on one organ, the mind, of one species, *Homo sapiens*. But in another sense it is a radical thesis that discards the way issues about the mind have been framed for almost a century. The premises of this book are probably not what you think they are. Thinking is computation, I claim, but that does not mean that the computer is a good metaphor for the mind. The mind is a set of modules, but the modules are not encapsulated boxes or circumscribed swatches on the surface of the brain. The organization of our mental modules comes from our genetic program, but that does not mean that there is a gene for every trait or that learning is less important than we used to think. The mind is an adaptation designed by natural selection, but that does not mean that everything we think, feel, and do is biologically adaptive. We evolved from apes, but that does not mean we have the same minds as

apes. And the ultimate goal of natural selection is to propagate genes, but that does not mean that the ultimate goal of people is to propagate genes. Let me show you why not.

This book is about the brain, but I will not say much about neurons, hormones, and neurotransmitters. That is because the mind is not the brain but what the brain does, and not even everything it does, such as metabolizing fat and giving off heat. The 1990s have been named the Decade of the Brain, but there will never be a Decade of the Pancreas. The brain's special status comes from a special thing the brain does, which makes us see, think, feel, choose, and act. That special thing is information processing, or computation.

Information and computation reside in patterns of data and in relations of logic that are independent of the physical medium that carries them. When you telephone your mother in another city, the message stays the same as it goes from your lips to her ears even as it physically changes its form, from vibrating air, to electricity in a wire, to charges in silicon, to flickering light in a fiber optic cable, to electromagnetic waves, and then back again in reverse order. In a similar sense, the message stays the same when she repeats it to your father at the other end of the couch after it has changed its form inside her head into a cascade of neurons firing and chemicals diffusing across synapses. Likewise, a given program can run on computers made of vacuum tubes, electromagnetic switches, transistors, integrated circuits, or well-trained pigeons, and it accomplishes the same things for the same reasons.

This insight, first expressed by the mathematician Alan Turing, the computer scientists Alan Newell, Herbert Simon, and Marvin Minsky, and the philosophers Hilary Putnam and Jerry Fodor, is now called the computational theory of mind. It is one of the great ideas in intellectual history, for it solves one of the puzzles that make up the "mind-body problem": how to connect the ethereal world of meaning and intention, the stuff of our mental lives, with a physical hunk of matter like the brain. Why did Bill get on the bus? Because he wanted to visit his grandmother and knew the bus would take him there. No other answer will do. If he hated the sight of his grandmother, or if he knew the route had changed, his body would not be on that bus. For millennia this has been

a paradox. Entities like "wanting to visit one's grandmother" and "knowing the bus goes to Grandma's house" are colorless, odorless, and tasteless. But at the same time they are *causes* of physical events, as potent as any billiard ball clacking into another.

The computational theory of mind resolves the paradox. It says that beliefs and desires are *information*, incarnated as configurations of symbols. The symbols are the physical states of bits of matter, like chips in a computer or neurons in the brain. They symbolize things in the world because they are triggered by those things via our sense organs, and because of what they do once they are triggered. If the bits of matter that constitute a symbol are arranged to bump into the bits of matter constituting another symbol in just the right way, the symbols corresponding to one belief can give rise to new symbols corresponding to another belief logically related to it, which can give rise to symbols corresponding to other beliefs, and so on. Eventually the bits of matter constituting a symbol bump into bits of matter connected to the muscles, and behavior happens. The computational theory of mind thus allows us to keep beliefs and desires in our explanations of behavior while planting them squarely in the physical universe. It allows meaning to cause and be caused.

The computational theory of mind is indispensable in addressing the questions we long to answer. Neuroscientists like to point out that all parts of the cerebral cortex look pretty much alike—not only the different parts of the human brain, but the brains of different animals. One could draw the conclusion that all mental activity in all animals is the same. But a better conclusion is that we cannot simply look at a patch of brain and read out the logic in the intricate pattern of connectivity that makes each part do its separate thing. In the same way that all books are physically just different combinations of the same seventy-five or so characters, and all movies are physically just different patterns of charges along the tracks of a videotape, the mammoth tangle of spaghetti of the brain may all look alike when examined strand by strand. The content of a book or a movie lies in the *pattern* of ink marks or magnetic charges, and is apparent only when the piece is read or seen. Similarly, the content of brain activity lies in the patterns of connections and patterns of activity among the neurons. Minute differences in the details of the connections may cause similar-looking brain patches to implement very different programs. Only when the program is run does the coherence become evident. As Tooby and Cosmides have written,

There are birds that migrate by the stars, bats that echolocate, bees that compute the variance of flower patches, spiders that spin webs, humans that speak, ants that farm, lions that hunt in teams, cheetahs that hunt falone, monogamous gibbons, polyandrous seahorses, polygynous gorillas.

... There are millions of animal species on earth, each with a different set of cognitive programs. *The same basic neural tissue embodies all of these programs*, and it could support many others as well. Facts about the properties of neurons, neurotransmitters, and cellular development cannot tell you which of these millions of programs the human mind contains. Even if all neural activity is the expression of a uniform process at the cellular level, it is the arrangement of neurons—into bird song templates or web-spinning programs—that matters.

That does not imply, of course, that the brain is irrelevant to understanding the mind! Programs are assemblies of simple information-processing units—tiny circuits that can add, match a pattern, turn on some other circuit, or do other elementary logical and mathematical operations. What those microcircuits can do depends only on what they are made of. Circuits made from neurons cannot do exactly the same things as circuits made from silicon, and vice versa. For example, a silicon circuit is faster than a neural circuit, but a neural circuit can match a larger pattern than a silicon one. These differences ripple up through the programs built from the circuits and affect how quickly and easily the programs do various things, even if they do not determine exactly which things they do. My point is not that prodding brain tissue is irrelevant to understanding the mind, only that it is not enough. Psychology, the analysis of mental software, will have to burrow a considerable way into the mountain before meeting the neurobiologists tunneling through from the other side.

The computational theory of mind is not the same thing as the despised "computer metaphor." As many critics have pointed out, computers are serial, doing one thing at a time; brains are parallel, doing millions of things at once. Computers are fast; brains are slow. Computer parts are reliable; brain parts are noisy. Computers have a limited number of connections; brains have trillions. Computers are assembled according to a blueprint; brains must assemble themselves. Yes, and computers come in putty-colored boxes and have AUTOEXECBAT files and run screen-savers with flying toasters, and brains do not. The claim is not that the brain is like commercially available computers. Rather, the claim is that brains and computers embody intelligence for some of the same

reasons. To explain how birds fly, we invoke principles of lift and drag and fluid mechanics that also explain how airplanes fly. That does not commit us to an Airplane Metaphor for birds, complete with jet engines and complimentary beverage service.

Without the computational theory, it is impossible to make sense of the evolution of the mind. Most intellectuals think that the human mind must somehow have escaped the evolutionary process. Evolution, they think, can fabricate only stupid instincts and fixed action patterns: a sex drive, an aggression urge, a territorial imperative, hens sitting on eggs and ducklings following hulks. Human behavior is too subtle and flexible to be a product of evolution, they think; it must come from somewhere else—from, say, "culture." But if evolution equipped us not with irresistible urges and rigid reflexes but with a neural computer, everything changes. A program is an intricate recipe of logical and statistical operations directed by comparisons, tests, branches, loops, and subroutines embedded in subroutines. Artificial computer programs, from the Macintosh user interface to simulations of the weather to programs that recognize speech and answer questions in English, give us a hint of the finesse and power of which computation is capable. Human thought and behavior, no matter how subtle and flexible, could be the product of a very complicated program, and that program may have been our endowment from natural selection. The typical imperative from biology is not "Thou shalt. . .," but "If . . . then . . . else."

The mind, I claim, is not a single organ but a system of organs, which we can think of as psychological faculties or mental modules. The entities now commonly evoked to explain the mind—such as general intelligence, a capacity to form culture, and multipurpose learning strategies—will surely go the way of protoplasm in biology and of earth, air, fire, and water in physics. These entities are so formless, compared to the exacting phenomena they are meant to explain, that they must be granted near-magical powers. When the phenomena are put under the microscope, we discover that the complex texture of the everyday world is supported not by a single substance but by many layers of elaborate machinery. Biologists long ago replaced the concept of an all-powerful protoplasm with the concept of functionally specialized

mechanisms. The organ systems of the body do their jobs because each is built with a particular structure tailored to the task. The heart circulates the blood because it is built like a pump; the lungs oxygenate the blood because they are built like gas exchangers. The lungs cannot pump blood and the heart cannot oxygenate it. This specialization goes all the way down. Heart tissue differs from lung tissue, heart cells differ from lung cells, and many of the molecules making up heart cells differ from those making up lung cells. If that were not true, our organs would not work.

A jack-of-all-trades is master of none, and that is just as true for our mental organs as for our physical organs. The robot challenge makes that clear. Building a robot poses many software engineering problems, and different tricks are necessary to solve them.

Take our first problem, the sense of sight. A seeing machine must solve a problem called inverse optics. Ordinary optics is the branch of physics that allows one to predict how an object with a certain shape, material, and illumination projects the mosaic of colors we call the retinal image. Optics is a well-understood subject, put to use in drawing, photography, television engineering, and more recently, computer graphics and virtual reality. But the brain must solve the *opposite* problem. The input is the retinal image, and the output is a specification of the objects in the world and what they are made of—that is, what we know we are seeing. And there's the rub. Inverse optics is what engineers call an "ill-posed problem." It literally has no solution. Just as it is easy to multiply some numbers and announce the product but impossible to take a product and announce the numbers that were multiplied to get it, optics is easy but inverse optics impossible. Yet your brain does it every time you open the refrigerator and pull out a jar. How can this be? !

The answer is that *the brain supplies the missing information*, information about the world we evolved in and how it reflects light. If the Visual brain "assumes" that it is living in a certain kind of world—an evenly lit world made mostly of rigid parts with smooth, uniformly colored surfaces—it can make good guesses about what is out there. As we saw earlier, it's impossible to distinguish coal from snow by examining the brightnesses of their retinal projections. But say there is a module for perceiving the properties of surfaces, and built into it is the following assumption: "The world is smoothly and uniformly lit." The module can solve the coal-versus-snow problem in three steps: subtract out any gradient of brightness from one edge of the scene to the other; estimate the average level of brightness of

the whole scene; and calculate the shade of gray of each patch by subtracting its brightness from the average brightness. Large positive deviations from the average are then seen as white things, large negative deviations as black things. If the illumination really is smooth and uniform, those perceptions will register the surfaces of the world accurately. Since Planet Earth has, more or less, met the even-illumination assumption for eons, natural selection would have done well by building the assumption in.

The surface-perception module solves an unsolvable problem, but at a price. The brain has given up any pretense of being a general problem-solver. It has been equipped with a gadget that perceives the nature of surfaces in typical earthly viewing conditions because it is specialized for that parochial problem. Change the problem slightly and the brain no longer solves it. Say we place a person in a world that is not blanketed with sunshine but illuminated by a cunningly arranged patchwork of light. If the surface-perception module assumes that illumination is even, it should be seduced into hallucinating objects that aren't there. Could that really happen? It happens every day. We call these hallucinations slide shows and movies and television (complete with the illusory black I mentioned earlier). When we watch TV, we stare at a shimmering piece of glass, but our surface-perception module tells the rest of our brain that we are seeing real people and places. The module has been unmasked; it does not apprehend the nature of things but relies on a cheat-sheet. That cheat-sheet is so deeply embedded in the operation of our visual brain that we cannot erase the assumptions written on it. Even in a lifelong couch potato, the visual system never "learns" that television is a pane of glowing phosphor dots, and the person never loses the illusion that there is a world behind the pane.

Our other mental modules need their own cheat-sheets to solve their unsolvable problems. A physicist who wants to figure out how the body moves when muscles are contracted has to solve problems in kinematics (the geometry of motion) and dynamics (the effects of forces). But a brain that has to figure out how to contract muscles to get the body to move has to solve problems in *inverse* kinematics and *inverse* dynamics—what forces to apply to an object to get it to move in a certain trajectory. Like inverse optics, inverse kinematics and dynamics are ill-posed problems. Our motor modules solve them by making extraneous but reasonable assumptions—not assumptions about illumination, of course, but assumptions about bodies in motion.

Our common sense about other people is a kind of intuitive psychol-

ogy—we try to infer people's beliefs and desires from what they do, and try to predict what they will do from our guesses about their beliefs and desires. Our intuitive psychology, though, must make the assumption that other people *have* beliefs and desires; we cannot sense a belief or desire in another person's head the way we smell oranges. If we did not see the social world through the lens of that assumption, we would be like the Samaritan I robot, which sacrificed itself for a bag of lima beans, or like Samaritan II, which went overboard for any object with a human-like head, even if the head belonged to a large wind-up toy. ⁱ (Later we shall see that people suffering from a certain syndrome lack the assumption that people have minds and *do* treat other people as wind-up toys.) Even our feelings of love for our family members embody a specific assumption about the laws of the natural world, in this case an inverse of the ordinary laws of genetics. Family feelings are designed to help our genes replicate themselves, but we cannot see or smell genes. Scientists use forward genetics to deduce how genes get distributed among organisms (for example, meiosis and sex cause the offspring of two people to have fifty percent of their genes in common); our emotions about kin use a kind of inverse genetics to guess which of the organisms we interact with are likely to share our genes (for example, if someone appears to have the same parents as you do, treat the person as if their genetic well-being overlaps with yours). I will return to all these topics in later chapters.

The mind has to be built out of specialized parts because it has to solve specialized problems. Only an angel could be a general problem-solver; we mortals have to make fallible guesses from fragmentary information. Each of our mental modules solves its unsolvable problem by a leap of faith about how the world works, by making assumptions that are indispensable but indefensible—the only defense being that the assumptions worked well enough in the world of our ancestors.

The word "module" brings to mind detachable, snap-in components, and that is misleading. Mental modules are not likely to be visible to the naked eye as circumscribed territories on the surface of the brain, like the flank steak and the rump roast on the supermarket cow display. A mental module probably looks more like roadkill, sprawling messily over the bulges and crevasses of the brain. Or it may be broken into regions that are interconnected by fibers that make the regions act as a unit. The beauty of information processing is the flexibility of its demand for real estate. Just as a corporation's management can be scattered across sites

linked by a telecommunications network, or a computer program can be fragmented into different parts of the disk or memory, the circuitry underlying a psychological module might be distributed across the brain in a spatially haphazard manner. And mental modules need not be tightly sealed off from one another, communicating only through a few narrow pipelines. (That is a specialized sense of "module" that many cognitive scientists have debated, following a definition by Jerry Fodor.) Modules are defined by the special things they do with the information available to them, not necessarily by the kinds of information they have available.

So the metaphor of the mental module is a bit clumsy; a better one is Noam Chomsky's "mental organ." An organ of the body is a specialized structure tailored to carry out a particular function. But our organs do not come in a bag like chicken gizzards; they are integrated into a complex whole. The body is composed of systems divided into organs assembled from tissues built out of cells. Some kinds of tissues, like the epithelium, are used, with modifications, in many organs. Some organs, like the blood and the skin, interact with the rest of the body across a widespread, convoluted interface, and cannot be encircled by a dotted line. Sometimes it is unclear where one organ leaves off and another begins, or how big a chunk of the body we want to call an organ. (Is the hand an organ? the finger? a bone in the finger?) These are all pedantic questions of terminology, and anatomists and physiologists have not wasted their time on them. What is clear is that the body is not made of Spam but has a heterogeneous structure of many specialized parts. All this is likely to be true of the mind. Whether or not we establish exact boundaries for the components of the mind, it is clear that it is not made of mental Spam but has a heterogeneous structure of many specialized parts.

Our physical organs owe their complex design to the information in the human genome, and so, I believe, do our mental organs. We do not learn to have a pancreas, and we do not learn to have a visual system, language acquisition, common sense, or feelings of love, friendship, and fairness. No single discovery proves the claim (just as no single discovery proves that the pancreas is innately structured), but many lines of evidence converge on it. The one that most impresses me is the Robot Challenge. Each of the major engineering problems solved by the mind is unsolvable

without built-in assumptions about the laws that hold in that arena of interaction with the world. All of the programs designed by artificial intelligence researchers have been specially engineered for a particular domain, such as language, vision, movement, or one of many different kinds of common sense. Within artificial intelligence research, the proud parent of a program will sometimes tout it as a mere demo of an amazingly powerful general-purpose system to be built in the future, but everyone else in the field routinely writes off such hype. I predict that no one will ever build a humanlike robot—and I mean a *really* humanlike robot—unless they pack it with computational systems tailored to different problems.

Throughout the book we will run into other lines of evidence that our mental organs owe their basic design to our genetic program. I have already mentioned that much of the fine structure of our personality and intelligence is shared by identical twins reared apart and hence charted by the genes. Infants and young children, when tested with ingenious methods, show a precocious grasp of the fundamental categories of the physical and social world, and sometimes command information that was never presented to them. People hold many beliefs that are at odds with their experience but were true in the environment in which we evolved, and they pursue goals that subvert their own well-being but were adaptive in that environment. And contrary to the widespread belief that cultures can vary arbitrarily and without limit, surveys of the ethnographic literature show that the peoples of the world share an astonishingly detailed universal psychology.

But if the mind has a complex innate structure, that does *not* mean that learning is unimportant. Framing the issue in such a way that innate structure and learning are pitted against each other, either as alternatives or, almost as bad, as complementary ingredients or interacting forces, is a colossal mistake. It's not that the claim that there is an interaction between innate structure and learning (or between heredity and environment, nature and nurture, biology and culture) is literally wrong. Rather, it falls into the category of ideas that are so bad they are not even wrong.

Imagine the following dialogue:

"This new computer is brimming with sophisticated technology. It has a 500 megahertz processor, a gigabyte of RAM, a terabyte of disk storage, a 3-D color virtual reality display, speech output, wireless access to the World Wide Web, expertise in a dozen subjects, and built-in editions of

the Bible, the *Encyclopaedia Britannica*, *Bartlett's Famous Quotations*, and the complete works of Shakespeare. Tens of thousands of hacker-hours went into its design."

"Oh, so I guess you're saying that it doesn't matter what I type into the computer. With all that built-in structure, its environment can't be very important. It will always do the same thing, regardless of what I type in."

The response is patently senseless. Having a lot of built-in machinery should make a system respond *more* intelligently and flexibly to its inputs, not less. Yet the reply captures how centuries of commentators have reacted to the idea of a richly structured, high-tech mind.

And the "interactionist" position, with its phobia of ever specifying the innate part of the interaction, is not much better. Look at these claims.

The behavior of a computer comes from a complex interaction between the processor and the input.

When trying to understand how a car works, one cannot neglect the engine or the gasoline or the driver. All are important factors.

The sound coming out of this CD player represents the inextricably intertwined mixture of two crucial variables: the structure of the machine, and the disk you insert into it. Neither can be ignored.

These statements are true but useless—so blankly uncomprehending, so defiantly incurious, that it is almost as bad to assert them as to deny them. For minds, just as for machines, the metaphors of a mixture of two ingredients, like a martini, or a battle between matched forces, like a tug-of-war, are wrongheaded ways of thinking about a complex device designed to process information. Yes, every part of human intelligence involves culture and learning. But learning is not a surrounding gas or force field, and it does not happen by magic. It is made possible by innate machinery designed to do the learning. The claim that there are several innate modules is a claim that there are several innate learning machines, each of which learns according to a particular logic. To understand learning, we need new ways of thinking to replace the prescientific metaphors—the mixtures and forces, the writing on slates and sculpting of blocks of marble. We need ideas that capture the ways a complex device can tune itself to unpredictable aspects of the world and take in the kinds of data it needs to function.

The idea that heredity and environment interact is not always mean-

ingless, but I think it confuses two issues: what all minds have in common, and how minds can differ. The vapid statements above can be made intelligible by replacing "How X works" with "What makes X work better than Y":

The *usefulness* of a computer depends on both the power of its processor and the expertise of the user.

The *speed* of a car depends on the engine, the fuel, and the skill of the driver. All are important factors.

The *quality* of sound coming from a CD player depends on two crucial variables: the player's mechanical and electronic design, and the quality of the original recording. Neither can be ignored.

When we are interested in *how much better* one system functions than a similar one, it is reasonable to gloss over the causal chains inside each system and tally up the factors that make the whole thing fast or slow, hi-fi or low-fi. And this *ranking* of people—to determine who enters medical school, or who gets the job—is where the framing of nature versus nurture comes from.

But this book is about how the mind works, not about why some people's minds might work a bit better in certain ways than other people's minds. The evidence suggests that humans everywhere on the planet see, talk, and think about objects and people in the same basic way. The difference between Einstein and a high school dropout is trivial compared to the difference between the high school dropout and the best robot in existence, or between the high school dropout and a chimpanzee. That is the mystery I want to address. Nothing could be farther from my subject matter than a comparison between the means of overlapping bell curves for some crude consumer index like IQ. And for this reason, the relative importance of innateness and learning is a phony issue.

An emphasis on innate design should not, by the way, be confused with the search for "a gene for" this or that mental organ. Think of the genes and putative genes that have made the headlines: genes for muscular dystrophy, Huntington's disease, Alzheimer's, alcoholism, schizophrenia, manic-depressive disorder, obesity, violent outbursts, dyslexia, bed-wetting, and some kinds of retardation. They are *disorders*, all of them. There have been no discoveries of a gene for civility, language, memory, motor control, intelligence, or other complete mental systems, and there probably won't ever be. The reason was summed up by the politician Sam Rayburn: Any jackass can kick down a barn, but it takes a

carpenter to build one. Complex mental organs, like complex physical organs, surely are built by complex genetic recipes, with many genes cooperating in as yet unfathomable ways. A defect in any one of them could corrupt the whole device, just as a defect in any part of a complicated machine (like a loose distributor cable in a car) can bring the machine to a halt.

The genetic assembly instructions for a mental organ do not specify every connection in the brain as if they were a wiring schematic for a Heathkit radio. And we should not expect each organ to grow under a particular bone of the skull regardless of what else happens in the brain. The brain and all the other organs differentiate in embryonic development from a ball of identical cells. Every part of the body, from the toenails to the cerebral cortex, takes on its particular shape and substance when its cells respond to some kind of information in its neighborhood that unlocks a different part of the genetic program. The information may come from the taste of the chemical soup that a cell finds itself in, from the shapes of the molecular locks and keys that the cell engages, from mechanical tugs and shoves from neighboring cells, and other cues still poorly understood. The families of neurons that will form the different mental organs, all descendants of a homogeneous stretch of embryonic tissue, must be designed to be opportunistic as the brain assembles itself, seizing any available information to differentiate from one another. The coordinates in the skull may be one trigger for differentiation, but the pattern of input firings from connected neurons is another. Since the brain is destined to be an organ of computation, it would be surprising if the genome did not exploit the capacity of neural tissue to process information during brain assembly.

In the sensory areas of the brain, where we can best keep track of what is going on, we know that early in fetal development neurons are wired according to a rough genetic recipe. The neurons are born in appropriate numbers at the right times, migrate to their resting places, send out connections to their targets, and hook up to appropriate cell types in the right general regions, all under the guidance of chemical trails and molecular locks and keys. To make precise connections, though, the baby neurons must begin to function, and their firing pattern carries information downstream about their pinpoint connections. This isn't "experience," as it all can take place in the pitch-black womb, sometimes before the rods and cones are functioning, and many mammals can see almost perfectly as soon as they are born. It is

more like a kind of genetic data compression or a set of internally generated test patterns. These patterns can trigger the cortex at the receiving end to differentiate, at least one step of the way; into the kind of cortex that is appropriate to processing the incoming information. (For example, in animals that have been cross-wired so that the eyes are connected to the auditory brain, that area shows a few hints of the properties of the visual brain.) How the genes control brain development is still unknown, but a reasonable summary of what we know so far is that brain modules assume their identity by a combination of what kind of tissue they start out as, where they are in the brain, and what patterns of triggering input they get during critical periods in development.

Our organs of computation are a product of natural selection. The biologist Richard Dawkins called natural selection the Blind Watchmaker; in the case of the mind, we can call it the Blind Programmer. Our mental programs work as well as they do because they were shaped by selection to allow our ancestors to master rocks, tools, plants, animals, and each other, ultimately in the service of survival and reproduction.

Natural selection is not the only cause of evolutionary change. Organisms also change over the eons because of statistical accidents in who lives and who dies, environmental catastrophes that wipe out whole families of creatures, and the unavoidable by-products of changes that *are* the product of selection. But natural selection is the only evolutionary force that acts like an engineer, "designing" organs that accomplish improbable but adaptive outcomes (a point that has been made forcefully by the biologist George Williams and by Dawkins). The textbook argument for natural selection, accepted even by those who feel that selection has been overrated (such as the paleontologist Stephen Jay Gould), comes from the vertebrate eye. Just as a watch has too many finely meshing parts (gears, springs, pivots, and so on) to have been assembled by a tornado or a river eddy, entailing instead the design of a watchmaker, the eye has too many finely meshing parts (lens, iris, retina, and so on) to have arisen from a random evolutionary force like a big mutation, statistical drift, or the fortuitous shape of the nooks and crannies between other organs. The design of the eye must be a product of

natural selection of replicators, the only nonmiraculous natural process we know of that can manufacture well-functioning machines. The organism appears as if it was designed to see well now because it owes its existence to the success of its ancestors in seeing well in the past. (This point will be expanded in Chapter 3.)

Many people acknowledge that natural selection is the artificer of the body but draw the line when it comes to the human mind. The mind, they say, is a by-product of a mutation that enlarged the head, or is a clumsy programmer's hack, or was given its shape by cultural rather than biological evolution. Tooby and Cosmides point out a delicious irony. The eye, that most uncontroversial example of fine engineering by natural selection, is not just any old organ that can be sequestered with flesh and bone, far away from the land of the mental. It doesn't digest food or, except in the case of Superman, change anything in the physical world. What does the eye do? The eye is an organ of information processing, firmly connected to—anatomically speaking, a part of—the brain. And all those delicate optics and intricate circuits in the retina do not dump information into a yawning empty orifice or span some Cartesian chasm from a physical to a mental realm. The receiver of this richly structured message must be every bit as well engineered as the sender. As we have seen in comparing human vision and robot vision, the parts of the mind that allow us to see are indeed well engineered, and there is no reason to think that the quality of engineering progressively deteriorates as the information flows upstream to the faculties that interpret and act on what we see.

The adaptationist program in biology, or the careful use of natural selection to reverse-engineer the parts of an organism, is sometimes ridiculed as an empty exercise in after-the-fact storytelling. In the satire of the syndicated columnist Cecil Adams, "the reason our hair is brown is that it enabled our monkey ancestors to hide amongst the coconuts." Admittedly, there is no shortage of bad evolutionary "explanations." Why do men avoid asking for directions? Because our male ancestors might have been killed if they approached a stranger. What purpose does music serve? It brings the community together. Why did happiness evolve? Because happy people are pleasant to be around, so they attracted more allies. What is the function of humor? To relieve tension. Why do people overestimate their chance of surviving an illness? Because it helps them to operate effectively in life.

These musings strike us as glib and lame, but it is not because they

dare to seek an evolutionary explanation of how some part of the mind works. It is because they botch the job. First, many of them never bother to establish the facts. Has anyone ever documented that *women* like to ask for directions? Would a woman in a foraging society *not* have come to harm when she approached a stranger? Second, even if the facts had been established, the stories try to explain one puzzling fact by taking for granted some other fact that is just as much of a puzzle, getting us nowhere. *Why* do rhythmic noises bring a community together? *Why* do people like to be with happy people? *Why* does humor relieve tension? The authors of these explanations treat some parts of our mental life as so obvious—they are, after all, obvious to each of *us*, here inside our heads—that they don't need to be explained. But *all* parts of the mind are up for grabs—every reaction, every pleasure, every taste—when we try to explain how it evolved. *We could have* evolved like the Samaritan I robot, which sacrificed itself to save a sack of lima beans, or like dung beetles, which must find dung delicious, or like the masochist in the old joke about sadomasochism (Masochist: "Hit me!" Sadist: "No!").

A good adaptationist explanation needs the fulcrum of an engineering analysis that is independent of the part of the mind we are trying to explain. The analysis begins with a goal to be attained and a world of causes and effects in which to attain it, and goes on to specify what kinds of designs are better suited to attain it than others. Unfortunately for those who think that the departments in a university reflect meaningful divisions of knowledge, it means that psychologists have to look outside psychology if they want to explain what the parts of the mind are for. To understand sight, we have to look to optics and computer vision systems. To understand movement, we have to look to robotics. To understand sexual and familial feelings, we have to look to Mendelian genetics. To understand cooperation and conflict, we have to look to the mathematics of games and to economic modeling.

Once we have a spec sheet for a well-designed mind, we can see whether *Homo sapiens* has that kind of mind. We do the experiments or surveys to get the facts down about a mental faculty, and then see whether the faculty meets the specs: whether it shows signs of precision, complexity, efficiency, reliability, and specialization in solving its assigned problem, especially in comparison with the vast number of alternative designs that are biologically growable.

The logic of reverse-engineering has guided researchers in visual perception for over a century, and that may be why we understand vision

better than we understand any other part of the mind. There is no reason that reverse-engineering guided by evolutionary theory should not bring insight about the rest of the mind. An interesting example is a new theory of pregnancy sickness (traditionally called "morning sickness") by the biologist Margie Profet. Many pregnant women become nauseated and avoid certain foods. Though their sickness is usually explained away as a side effect of hormones, there is no reason that hormones should induce nausea and food aversions rather than, say, hyperactivity, aggressiveness, or lust. The Freudian explanation is equally unsatisfying: that pregnancy sickness represents the woman's loathing of her husband and her unconscious desire to abort the fetus orally.

Profet predicted that pregnancy sickness should confer some benefit that offsets the cost of lowered nutrition and productivity. Ordinarily, nausea is a protection against eating toxins: the poisonous food is ejected from the stomach before it can do much harm, and our appetite for similar foods is reduced in the future. Perhaps pregnancy sickness protects women against eating or digesting foods with toxins that might harm the developing fetus. Your local Happy Carrot Health Food Store notwithstanding, there is nothing particularly healthy about natural foods. Your cabbage, a Darwinian creature, has no more desire to be eaten than you do, and since it can't very well defend itself through behavior, it resorts to chemical warfare. Most plants have evolved dozens of toxins in their tissues: insecticides, insect repellents, irritants, paralytics, poisons, and other sand to throw in herbivores' gears. Herbivores have in turn evolved countermeasures, such as a liver to detoxify the poisons and the taste sensation we call bitterness to deter any further desire to ingest them. But the usual defenses may not be enough to protect a tiny embryo.

So far this may not sound much better than the barf-up-your-baby theory, but Profet synthesized hundreds of studies, done independently of each other and of her hypothesis, that support it. She meticulously documented that (1) plant toxins in dosages that adults tolerate can cause birth defects and induce abortion when ingested by pregnant women; (2) pregnancy sickness begins at the point when the embryo's organ systems are being laid down and the embryo is most vulnerable to teratogens (birth defect—inducing chemicals) but is growing slowly and has only a modest need for nutrients; (3) pregnancy sickness wanes at the stage when the embryo's organ systems are nearly complete and its biggest need is for nutrients to allow it to grow; (4) women with pregnancy sickness selectively avoid bitter, pungent, highly flavored, and

novel foods, which are in fact the ones most likely to contain toxins; (5) women's sense of smell becomes hypersensitive during the window of pregnancy sickness and less sensitive than usual thereafter; (6) foraging peoples (including, presumably, our ancestors) are at even higher risk of ingesting plant toxins, because they eat wild plants rather than domesticated crops bred for palatability; (7) pregnancy sickness is universal across human cultures; (8) women with more severe pregnancy sickness are less likely to miscarry; (9) women with more severe pregnancy sickness are less likely to bear babies with birth defects. The fit between how a baby-making system in a natural ecosystem ought to work and how the feelings of modern women do work is impressive, and gives a measure of confidence that Profet's hypothesis is correct.

The human mind is a product of evolution, so our mental organs are either present in the minds of apes (and perhaps other mammals and vertebrates) or arose from overhauling the minds of apes, specifically, the common ancestors of humans and chimpanzees that lived about six million years ago in Africa. Many titles of books on human evolution remind us of this fact: *The Naked Ape*, *The Electric Ape*, *The Scented Ape*, *The Lopsided Ape*, *The Aquatic Ape*, *The Thinking Ape*, *The Human Ape*, *The Ape That Spoke*, *The Third Chimpanzee*, *The Chosen Primate*. Some authors are militant that humans are barely different from chimpanzees and that any focus on specifically human talents is arrogant chauvinism or tantamount to creationism. For some readers that is a *reductio ad absurdum* of the evolutionary framework. If the theory says that man "at best is only a monkey shaved," as Gilbert and Sullivan put it in *Princess Ida*, then it fails to explain the obvious fact that men and monkeys have different minds.

We are naked, lopsided apes that speak, but we also have minds that differ considerably from those of apes. The outsize brain of *Homo sapiens sapiens* is, by any standard, an extraordinary adaptation. It has allowed us to inhabit every ecosystem on earth, reshape the planet, walk on the moon, and discover the secrets of the physical universe. Chimpanzees, for all their vaunted intelligence, are a threatened species clinging to a few patches of forest and living as they did millions of years ago. Our curiosity about this difference demands more than repeating that we

share most of our DNA with chimpanzees and that small changes can have big effects. Three hundred thousand generations and up to ten megabytes of potential genetic information are enough to revamp a mind considerably. Indeed, minds are probably easier to revamp than bodies because software is easier to modify than hardware. We should not be surprised to discover impressive new cognitive abilities in humans, language being just the most obvious one.

None of this is incompatible with the theory of evolution. Evolution is a conservative process, to be sure, but it can't be all *that* conservative or we would all be pond scum. Natural selection introduces differences into descendants by fitting them with specializations that adapt them to different niches. Any museum of natural history has examples of complex organs unique to a species or to a group of related species: the elephant's trunk, the narwhal's tusk, the whale's baleen, the platypus' duckbill, the armadillo's armor. Often they evolve rapidly on the geological timescale. The first whale evolved in something like ten million years from its common ancestor with its closest living relatives, ungulates such as cows and pigs. A book about whales could, in the spirit of the human-evolution books, be called *The Naked Cow*, but it would be disappointing if the book spent every page marveling at the similarities between whales and cows and never got around to discussing the adaptations that make them so different.

To say that the mind is an evolutionary adaptation is not to say that all behavior is adaptive in Darwin's sense. Natural selection is not a guardian angel that hovers over us making sure that our behavior always maximizes biological fitness. Until recently, scientists with an evolutionary bent felt a responsibility to account for acts that seem like Darwinian suicide, such as celibacy, adoption, and contraception. Perhaps, they ventured, celibate people have more time to raise large broods of nieces and nephews and thereby propagate more copies of their genes than they would if they had their own children. This kind of stretch is unnecessary, however. The reasons, first articulated by the anthropologist Donald Symons, distinguish evolutionary psychology from the school of thought in the 1970s and 1980s called sociobiology (though there is much overlap between the approaches as well).

First, selection operates over thousands of generations. For ninety-nine percent of human existence, people lived as foragers in small nomadic bands. Our brains are adapted to that long-vanished way of life, not to brand-new agricultural and industrial civilizations. They are not wired to cope with anonymous crowds, schooling, written language, government, police, courts, armies, modern medicine, formal social institutions, high technology, and other newcomers to the human experience. Since the modern mind is adapted to the Stone Age, not the computer age, there is no need to strain for adaptive explanations for everything we do. Our ancestral environment lacked the institutions that now entice us to nonadaptive choices, such as religious orders, adoption agencies, and pharmaceutical companies, so until very recently there was never a selection pressure to resist the enticements. Had the Pleistocene savanna contained trees bearing birth-control pills, we might have evolved to find them as terrifying as a venomous spider.

Second, natural selection is not a puppetmaster that pulls the strings of behavior directly. It acts by designing the generator of behavior: the package of information-processing and goal-pursuing mechanisms called the mind. Our minds are designed to generate behavior that would have been adaptive, on average, in our ancestral environment, but any particular deed done today is the effect of dozens of causes. Behavior is the outcome of an internal struggle among many mental modules, and it is played out on the chessboard of opportunities and constraints defined by *other* people's behavior. A recent cover story in *Time* asked, "Adultery: Is It in Our Genes?" The question makes no sense because neither adultery nor any other behavior can be in our genes. Conceivably a *desire* for adultery can be an indirect product of our genes, but the desire may be overridden by *other* desires that are also indirect products of our genes, such as the desire to have a trusting spouse. And the desire, even if it prevails in the rough-and-tumble of the mind, cannot be consummated as overt behavior unless there is a partner around in whom that desire has also prevailed. Behavior itself did not evolve; what evolved was the mind.

Reverse-engineering is possible only when one has a hint of what the device was designed to accomplish. We do not understand the olive-pitter until we catch on that it was designed as a machine for pitting olives

rather than as a paperweight or wrist-exerciser. The goals of the designer must be sought for every part of a complex device and for the device as a whole. Automobiles have a component, the carburetor, that's designed to mix air and gasoline, and mixing air and gasoline is a subgoal of the ultimate goal, carting people around. Though the *motives* of natural selection itself has no goal, it evolved entities that *like* (the automobile) are highly organized to bring about certain goals and subgoals. To reverse-engineer the mind, we must sort them out and identify the ultimate goal in its design. Was the human mind *ultimately* designed to create beauty? To discover truth? To love and *work*? To harmonize with other human beings and with nature?

The logic of natural selection gives the answer. The ultimate goal that the mind was designed to attain is maximizing the number of copies of the genes that created it. Natural selection cares only about the long-term fate of entities that replicate; that is, entities that retain a stable identity across many generations of copying. It predicts only that replicators whose effects tend to enhance the probability of their own replication come to predominate. When we ask questions like "Who or what is supposed to benefit from an adaptation?" and "What is a design in living things a design or?" the theory of natural selection provides the answer: the long-term stable replicators, genes. Even our bodies, our selves, are not the ultimate beneficiary of our design. As Gould has said, "What is the 'individual reproductive success' of which Darwin speaks? It cannot be the passage of one's body into the next generation—for, truly, you can't take it with you in this sense above all!" The criterion by which genes get selected is the quality of the bodies they build, but it is the *genes* making it into the next generation, not the perishable bodies, that *are* selected to live and fight another day.

Though there are some holdouts (such as Gould himself), the gene's-eye view predominates in evolutionary biology and has been a stunning success. It has asked, and is finding answers to, the deepest questions about life, such as how life arose, why there are cells, why there are bodies, why there is sex, how the genome is structured, why animals interact socially, and why there is communication. It is as indispensable to researchers in animal behavior as Newton's laws are to mechanical engineers.

But almost everyone misunderstands the theory. Contrary to popular belief, the gene-centered theory of evolution does *not* imply that the point of all human striving is to spread our genes. With the exception of

the fertility doctor who artificially inseminated patients with his own semen, the donors to the sperm bank for Nobel Prize winners, and other kooks, *no* human being (or animal) strives to spread his or her genes. Dawkins explained the theory in a book called *The Selfish Gene*, and the metaphor was chosen carefully. People don't selfishly spread their genes; genes selfishly spread themselves. They do it by the way they build our brains. By making us enjoy life, health, sex, friends, and children, the genes buy a lottery ticket for representation in the next generation, with odds that were favorable in the environment in which we evolved. Our goals are subgoals of the ultimate goal of the genes, replicating themselves. But the two are different. As far as *we* are concerned, our goals, conscious or unconscious, are not about genes at all, but about health and lovers and children and friends.

The confusion between our goals and our genes' goals has spawned one muddle after another. A reviewer of a book about the evolution of sexuality protests that human adultery, unlike the animal equivalent, cannot be a strategy to spread the genes because adulterers take steps to prevent pregnancy. But whose strategy are we talking about? Sexual desire is *not* people's strategy to propagate their genes. It's people's strategy to attain the pleasures of sex, and the pleasures of sex are the genes' strategy to propagate themselves. If the genes don't get propagated, it's because we are smarter than they are. A book on the emotional life of animals complains that if altruism according to biologists is just helping kin or exchanging favors, both of which serve the interests of one's genes, it would not *really* be altruism after all, but some kind of hypocrisy. This too is a mixup. Just as blueprints don't necessarily specify blue buildings, selfish genes don't necessarily specify selfish organisms. As we shall see, sometimes the most selfish thing a gene can do is to build a selfless brain. Genes are a play within a play, not the interior monologue of the players.

PSYCHOLOGICAL CORRECTNESS

The evolutionary psychology of this book is a departure from the dominant view of the human mind in our intellectual tradition, which Tooby and Cosmides have dubbed the Standard Social Science Model (SSSM). The SSSM proposes a fundamental division between biology and cul-

ture. Biology endows humans with the five senses, a few drives like hunger and fear, and a general capacity to learn. But biological evolution, according to the SSSM, has been superseded by cultural evolution. Culture is an autonomous entity that carries out a desire to perpetuate itself by setting up expectations and assigning roles, which can vary arbitrarily from society to society. Even the reformers of the SSSM have accepted its framing of the issues. Biology is "just as important as" culture, say the reformers; biology imposes "constraints" on behavior, and all behavior is a mixture of the two.

The SSSM not only has become an intellectual orthodoxy but has acquired a moral authority. When sociobiologists first began to challenge it, they met with a ferocity that is unusual even by the standards of academic invective. The biologist E. O. Wilson was doused with a pitcher of ice water at a scientific convention, and students yelled for his dismissal over bullhorns and put up posters urging people to bring noisemakers to his lectures. Angry manifestos and book-length denunciations were published by organizations with names like Science for the People and The Campaign Against Racism, IQ, and the Class Society. In *Not in Our Genes*, Richard Lewontin, Steven Rose, and Leon Kamin dropped innuendos about Donald Symons' sex life and doctored a defensible passage of Richard Dawkins' into an insane one. (Dawkins said of the genes, "They created us, body and mind"; the authors have quoted it repeatedly as "They control us, body and mind.") When *Scientific American* ran an article on behavior genetics (studies of twins, families, and adoptees), they entitled it "Eugenics Revisited," an allusion to the discredited movement to improve the human genetic stock. When the magazine covered evolutionary psychology, they called the article "The New Social Darwinists," an allusion to the nineteenth-century movement that justified social inequality as part of the wisdom of nature. Even one of sociobiology's distinguished practitioners, the primatologist Sarah Blaffer Hrdy, said, "I question whether sociobiology should be taught at the high school level, or even the undergraduate level. . . . The whole message of sociobiology is oriented toward the success of the individual. It's Machiavellian, and unless a student has a moral framework already in place, we could be producing social monsters by teaching this. It really fits in very nicely with the yuppie 'me first' ethos."

Entire scholarly societies joined in the fun, passing votes on empirical issues that one might have thought would be hashed out in the lab and the field. Margaret Mead's portrayal of an idyllic, egalitarian Samoa was

one of the founding documents of the SSSM, and when the anthropologist Derek Freeman showed that she got the facts spectacularly wrong, the American Anthropological Association voted at its business meeting to denounce his finding as unscientific. In 1986, twenty social scientists at a "Brain and Aggression" meeting drafted the Seville Statement on Violence, subsequently adopted by UNESCO and endorsed by several scientific organizations. The statement claimed to "challenge a number of alleged biological findings that have been used, even by some in our disciplines, to justify violence and war":

It is scientifically incorrect to say that we have inherited a tendency to make war from our animal ancestors.

It is scientifically incorrect to say that war or any other violent behavior is genetically programmed into our human nature.

It is scientifically incorrect to say that in the course of human evolution there has been a selection for aggressive behavior more than for other kinds of behavior.

It is scientifically incorrect to say that humans have a "violent brain."

It is scientifically incorrect to say that war is caused by "instinct" or any single motivation. . . . We conclude that biology does not condemn humanity to war, and that humanity can be freed from the bondage of biological pessimism and empowered with confidence to undertake the transformative tasks needed in the International Year of Peace and in the years to come.

What moral certainty could have incited these scholars to doctor quotations, censor ideas, attack the ideas' proponents ad hominem, smear them with unwarranted associations to repugnant political movements, and mobilize powerful institutions to legislate what is correct and incorrect? The certainty comes from an opposition to three putative implications of an innate human nature.

First, if the mind has an innate structure, different people (or different classes, sexes, and races) could have different innate structures. That would justify discrimination and oppression.

Second, if obnoxious behavior like aggression, war, rape, clannishness, and the pursuit of status and wealth are innate, that would make them "natural" and hence good. And even if they are deemed objectionable, they are in the genes and cannot be changed, so attempts at social reform are futile.

Third, if behavior is caused by the genes, then individuals cannot be

held responsible for their actions. If the rapist is following a biological imperative to spread his genes, it's not his fault.

Aside perhaps from a few cynical defense lawyers and a lunatic fringe who are unlikely to read manifestos in the *New York Review of Books*, no one has actually drawn these mad conclusions. Rather, they are thought to be extrapolations that the untutored masses *might* draw, so the dangerous ideas must themselves be suppressed. In fact, the problem with the three arguments is not that the conclusions are so abhorrent that no one should be allowed near the top of the slippery slope that leads to them. The problem is that there is no such slope; the arguments are non sequiturs. To expose them, one need only examine the logic of the theories and separate the scientific from the moral issues.

My point is not that scientists should pursue the truth in their ivory tower, undistracted by moral and political thoughts. Every human act involving another living being is both the subject matter of psychology and the subject matter of moral philosophy, and both are important. But they are not the same thing. The debate over human nature has been muddied by an intellectual laziness, an unwillingness to make moral arguments when moral issues come up. Rather than reasoning from principles of rights and values, the tendency has been to buy an off-the-shelf moral package (generally New Left or Marxist) or to lobby for a feel-good picture of human nature that would spare us from having to argue moral issues at all.

The moral equation in most discussions of human nature is simple: innate equals right-wing equals bad. Now, many hereditarian movements *have* been right-wing and bad, such as eugenics, forced sterilization, genocide, discrimination along racial, ethnic, and sexual lines, and the justification of economic and social castes. The Standard Social Science Model, to its credit, has provided some of the grounds that thoughtful social critics have used to undermine these practices.

But the moral equation is wrong as often as it is right. Sometimes left-wing practices are just as bad, and the perpetrators have tried to justify them using the SSSM's denial of human nature. Stalin's purges, the Gulag, Pol Pot's killing fields, and almost fifty years of repression in China—all have been justified by the doctrine that dissenting ideas

reflect not the operation of rational minds that have come to; different conclusions, but arbitrary cultural products that can be eradicated by re-engineering the society, "re-educating" those who were tainted by the old upbringing, and, if necessary, starting afresh with a new generation of slates that are still blank.

And sometimes left-wing positions are right *because* the denial of human nature is wrong. In *Hearts and Minds*, the 1974 documentary about the war in Vietnam, an American officer explains that we cannot apply our moral standards to the Vietnamese because their culture does not place a value on individual lives, so they do not suffer as we do when family members are killed. The director plays the quote over footage of wailing mourners at the funeral of a Vietnamese casualty, reminding us that the universality of love and grief refutes the officer's horrifying rationalization. For most of this century, guilty mothers have endured inane theories blaming them for every dysfunction or difference in their children (mixed messages cause schizophrenia, coldness causes autism, domineering causes homosexuality, lack of boundaries causes anorexia, insufficient "motherese" causes language disorders). Menstrual cramps, pregnancy sickness, and childbirth pain have been dismissed as women's "psychological" reactions to cultural expectations, rather than being treated as legitimate health issues.

The foundation of individual rights is the assumption that people have wants and needs and are authorities on what those wants and needs are. If people's stated desires were just some kind of erasable inscription or reprogrammable brainwashing, any atrocity could be justified. (Thus it is ironic that fashionable "liberation" ideologies like those of Michel Foucault and some academic feminists invoke a socially conditioned "interiorized authority," "false consciousness," or "inauthentic preference" to explain away the inconvenient fact that people enjoy the things that are alleged to oppress them.) A denial of human nature, no less than an emphasis on it, can be warped to serve harmful ends. We should expose whatever ends are harmful and whatever ideas are false, and not confuse the two.

So what about the three supposed implications of an innate human nature? The first "implication"—that an innate human nature[^] implies innate human differences—is no implication at all. The mental machin-

ery I argue for is installed in every neurologically normal human being. The differences among people may have nothing to do with the design of that machinery. They could very well come from random variations in the assembly process or from different life histories. Even if the differences were innate, they could be quantitative variations and minor quirks in equipment present in all of us (how fast a module works, which module prevails in a competition inside the head) and are not necessarily any more pernicious than the kinds of innate differences allowed in the Standard Social Science Model (a faster general-purpose learning process, a stronger sex drive).

A universal structure to the mind is not only logically possible but likely to be true. Tooby and Cosmides point out a fundamental consequence of sexual reproduction: every generation, each person's blueprint is scrambled with someone else's. That means we must be qualitatively alike. If two people's genomes had designs for different kinds of machines, like an electric motor and a gasoline engine, the new pastiche would not specify a working machine at all. Natural selection is a homogenizing force within a species; it eliminates the vast majority of macroscopic design variants because they are not improvements. Natural selection does depend on there having been variation in the past, but it feeds off the variation and uses it up. That is why all normal people have the same physical organs, and why we all surely have the same mental organs as well. There are, to be sure, microscopic variations among people, mostly small differences in the molecule-by-molecule sequence of many of our proteins. But at the level of functioning organs, physical and mental, people work in the same ways. Differences among people, for all their endless fascination to us as we live our lives, are of minor interest when we ask how the mind works. The same is true for differences—whatever their source—between the averages of entire groups of people, such as races.

The sexes, of course, are a different matter. The male and female reproductive organs are a vivid reminder that qualitatively different designs *are* possible for the sexes, and we know that the differences come from the special gadget of a genetic "switch," which triggers a line of biochemical dominoes that activate and deactivate families of genes throughout the brain and body. I will present evidence that some of these effects cause differences in how the mind works. In another of the ironies that run through the academic politics of human nature, this evolution-inspired research has proposed sex differences that are tightly focused on repro-

duction and related domains, and are far less invidious than the differences proudly claimed by some schools of feminism. Among the claims of "difference feminists" are that women do not engage in abstract linear reasoning, that they do not treat ideas with skepticism or evaluate them through rigorous debate, that they do not argue from general moral principles, and other insults.

But ultimately we cannot just look at who is portrayed more flatteringly; the question is what to make of any group differences we do stumble upon. And here we must be prepared to make a moral argument. Discrimination against individuals on the basis of their race, sex, or ethnicity is wrong. The argument can be defended in various ways that have nothing to do with the average traits of the groups. One might argue that it is unfair to deny a social benefit to individuals because of factors they cannot control, or that a victim of discrimination experiences it as a uniquely painful sting, or that a group of victims is liable to react with rage, or that discrimination tends to escalate into horrors like slavery and genocide. (Those who favor affirmative action could acknowledge that reverse discrimination is wrong but argue that it undoes an even greater wrong.) None of these arguments is affected by anything any scientist will ever claim to discover. The final word on the political non-implications of group differences must go to Gloria Steinem: "There are really not many jobs that actually require a penis or a vagina, and all the other occupations should be open to everyone."

The fallacy of the second supposed implication of a human nature—that if our ignoble motives are innate, they can't be so bad after all—is so obvious it has been given a name: the naturalistic fallacy, that what happens in nature is right. Forget the romantic nonsense in wildlife documentaries, where all creatures great and small act for the greater good and the harmony of the ecosystem. As Darwin said, "What a book a devil's chaplain might write on the clumsy, wasteful, blundering, low, and horribly cruel works of nature!" A classic example is the ichneumon wasp, who paralyzes a caterpillar and lays eggs in its body so her hatchlings can slowly devour its living flesh from the inside.

Like many species, *Homo sapiens* is a nasty business. Recorded history from the Bible to the present is a story of murder, rape, and war, and

honest ethnography shows that foraging peoples, like the rest of us, are more savage than noble. The !Kung San of the Kalahari Desert are often held out as a relatively peaceful people, and so they are, compared with other foragers: their murder rate is only as high as Detroit's. A linguist friend of mine who studies the Wari in the Amazon rainforest learned that their language has a term for edible things, which includes anyone who isn't a Wari. Of course humans don't have an "instinct for war" or a "violent brain," as the Seville Statement assures us, but humans don't exactly have an instinct for peace or a nonviolent brain, either. We cannot attribute all of human history and ethnography to toy guns and superhero cartoons.

Does that mean that "biology condemns man to war" (or rape or murder or selfish yuppies) and that any optimism about reducing it should be snuffed out? No one needs a scientist to make the moral point that war is not healthy for children and other living things, or the empirical point that some places and periods are vastly more peaceable than others and that we should try to understand and duplicate what makes them so. And no one needs the bromides of the Seville Statement or its disinformation that war is unknown among animals and that their dominance hierarchies are a form of bonding and affiliation that benefits the group. What could not hurt is a realistic understanding of the psychology of human malevolence. For what it's worth, the theory of a module-packed mind allows both for innate motives that lead to evil acts and for innate motives that can avert them. Not that this is a unique discovery of evolutionary psychology; all the major religions observe that mental life is often a struggle between desire and conscience.

When it comes to the hopes of changing bad behavior, the conventional wisdom again needs to be inverted: a complex human nature may allow *more* scope for change than the blank slate of the Standard Social Science Model. A richly structured mind allows for complicated negotiations inside the head, and one module could subvert the ugly designs of another one. In the SSSM, in contrast, upbringing is often said to have an insidious and irreversible power. "Is it a boy or a girl?" is the first question we ask about a new human being, and from then on parents treat their sons and daughters differently: they touch, comfort, breast-feed, indulge, and talk to boys and girls in unequal amounts. Imagine that this behavior has long-term consequences on the children, which include all the documented sex differences *and* a tendency to treat *their* children differently from birth. Unless we stationed parenting police in the mater-

nity ward, the circle would be complete and irrevocable. Culture would condemn women to inferiority, and we would be enslaved to the bondage of cultural pessimism, disempowered by self-doubt from undertaking transformative tasks.

Nature does not dictate what we should accept or how we should live our lives. Some feminists and gay activists react with fury to the banal observations that natural selection designed women in part for growing and nursing children and that it designed both men and women for heterosexual sex. They see in those observations the sexist and homophobic message that only traditional sexual roles are "natural" and that alternative lifestyles are to be condemned. For example, the novelist Mary Gordon, mocking a historian's remark that what all women have in common is the ability to bear children, wrote, "If the defining quality of being a woman is the ability to bear children, then not bearing children (as, for instance, Florence Nightingale and Greta Garbo did not) is somehow a failure to fulfill your destiny." I'm not sure what "the defining quality of being a woman" and "fulfilling your destiny" even *mean*, but I do know that happiness and virtue have nothing to do with what natural selection designed us to accomplish in the ancestral environment. They are for us to determine. In saying this I am no hypocrite, even though I am a conventional straight white male. Well into my procreating years I am, so far, voluntarily childless, having squandered my biological resources reading and writing, doing research, helping out friends and students, and jogging in circles, ignoring the solemn imperative to spread my genes. By Darwinian standards I am a horrible mistake, a pathetic loser, not one iota less than if I were a card-carrying member of Queer Nation. But I am happy to be that way, and if my genes don't like it, they can go jump in the lake.

Finally, what about blaming bad behavior on our genes? The neuroscientist Steven Rose, in a review of a book by E. O. Wilson in which Wilson wrote that men have a greater desire for polygamy than women, accused him of really saying, "Don't blame your mates for sleeping around, ladies, it's not their fault they are genetically programmed." The title of Rose's own book with Lewontin and Kamin, *Not in Our Genes*, is an allusion to *Julius Caesar*:

Men at some time are masters of their fates:
The fault, dear Brutus, lies not in our stars,
But in ourselves . . .

For Cassius, the programming that was thought to excuse human faults was not genetic but astrological, and that raises a key point. *Any* cause of behavior, not just the genes, raises the question of free will and responsibility. The difference between explaining behavior and excusing it is an ancient theme of moral reasoning, captured in the saw "To understand is not to forgive."

In this scientific age, "to understand" means to try to explain behavior as a complex interaction among (1) the genes, (2) the anatomy of the brain, (3) its biochemical state, (4) the person's family upbringing, (5) the way society has treated him or her, and (6) the stimuli that impinge upon the person. Sure enough, *every one* of these factors, not just the stars or the genes, has been inappropriately invoked as the source of our faults and a claim that we are not masters of our fates.

(1) In 1993 researchers identified a gene that was associated with uncontrollable violent outbursts. ("Think of the implications," one columnist wrote. "We may someday have a cure for hockey.") Soon afterward came the inevitable headline: "Man's Genes Have Made Him Kill, His Lawyers Claim."

(2) In 1982 an expert witness in the insanity defense of John Hinckley, who had shot President Reagan and three other men to impress the actress Jodie Foster, argued that a CAT scan of Hinckley's brain showed widened sulci and enlarged ventricles, a sign of schizophrenia and thus an excusing mental disease or defect. (The judge excluded the evidence, though the insanity defense prevailed.)

(3) In 1978 Dan White, having resigned from the San Francisco Board of Supervisors, walked into Mayor George Moscone's office and begged to be reinstated. When Moscone refused, White shot him dead, walked down the hall into the office of Supervisor Harvey Milk, and shot him dead too. White's lawyers successfully argued that at the time of his crime White had diminished capacity and had not committed a premeditated act because his binges on sugary junk food played havoc with his brain chemistry. White was convicted of voluntary manslaughter and served five years, thanks to the tactic that lives on in infamy as the Twinkie Defense. Similarly, in what is now known as the PMS (premen-

strual syndrome) Defense, raging hormones exonerated a surgeon who had assaulted a trooper who stopped her for drunk driving.

(4) In 1989 Lyle and Erik Menendez burst into their millionaire parents' bedroom and killed them with a shotgun. After several months of showing off their new Porsches and Rolexes, they confessed to the shootings. Their lawyers argued the case to a hung jury by claiming self-defense, despite the fact that the victims had been lying in bed, unarmed, eating strawberries and ice cream. The Menendez boys, the lawyers said, had been traumatized into believing that their parents were going to kill them because they had been physically, sexually, and emotionally abused by the father for years. (In a new trial in 1996 they were convicted of murder and sent to prison for life.)

(5) In 1994 Colin Ferguson boarded a train and began to shoot white people at random, killing six. The radical lawyer William Kunstler was prepared to defend him by invoking the Black Rage Syndrome, in which an African American can suddenly burst under the accumulated pressure of living in a racist society. (Ferguson rejected the offer and argued his own case, unsuccessfully.)

(6) In 1992 a death-row inmate asked an appeals court to reduce his sentence for rape and murder because he had committed his crimes under the influence of pornography. The Pornography-Made-Me-Do-It Defense is an irony for the schools of feminism that argue that biological explanations of rape reduce the rapist's responsibility and that a good tactic to fight violence against women is to blame it on pornography.

As science advances and explanations of behavior become less fanciful, the Specter of Creeping Exculpation, as Dennett calls it, will loom larger. Without a clearer moral philosophy, any cause of behavior could be taken to undermine free will and hence moral responsibility. Science is! guaranteed to appear to eat away at the will, *regardless* of what it finds, because the scientific mode of explanation cannot accommodate the mysterious notion of uncaused causation that underlies the will. If scientists wanted to show that people had free will, what would they look for? Some random neural event that the rest of the brain amplifies into a signal triggering behavior? But a random event does not fit the concept of free will any more than a lawful one does, and could not serve as the long-sought locus of moral responsibility. We would not find someone guilty if his finger pulled the trigger when it was mechanically connected to a roulette wheel;

why should it be any different if the roulette wheel is inside his skull? The same problem arises for another unpredictable cause that has been suggested as the source of free will, chaos theory, in which, according to the cliché, a butterfly's flutter can set off a cascade of events culminating in a hurricane. A fluttering in the brain that causes a hurricane of behavior, if it were ever found, would still be a cause of behavior and would not fit the concept of uncaused free will that underlies moral responsibility.

Either we dispense with all morality as an unscientific superstition, or we find a way to reconcile causation (genetic or otherwise) with responsibility and free will. I doubt that our puzzlement will ever be completely assuaged, but we can surely reconcile them in part. Like many philosophers, I believe that science and ethics are two self-contained systems played out among the same entities in the world, just as poker and bridge are different games played with the same fifty-two-card deck. **The science game treats people as material objects, and its rules are the physical processes that cause behavior through natural selection and neurophysiology. The ethics game treats people as equivalent, sentient, rational, free-willed agents, and its rules are the calculus that assigns moral value to behavior through the behavior's inherent nature or its consequences.**

Free will is an idealization of human beings that makes the ethics game playable. Euclidean geometry requires idealizations like infinite straight lines and perfect circles, and its deductions are sound and useful even though the world does not really have infinite straight lines or perfect circles. The world is close enough to the idealization that the theorems can usefully be applied. Similarly, ethical theory requires idealizations like free, sentient, rational, equivalent agents whose behavior is uncaused, and its conclusions can be sound and useful even though the world, as seen by science, does not really have uncaused events. As long as there is no outright coercion or gross malfunction of reasoning, the world is close enough to the idealization of free will that moral theory can meaningfully be applied to it.

Science and morality are separate spheres of reasoning. Only by recognizing them as separate can we have them both. If discrimination is wrong only if group averages are the same, if war and rape and greed are wrong only if people are never inclined toward them, if people are responsible for their actions only if the actions are mysterious, then either scientists must be prepared to fudge their data or all of us must be prepared to give up our values. Scientific arguments would turn into the

National Lampoon cover showing a puppy with a gun at its head and the caption "Buy This Magazine or We'll Shoot the Dog."

The knife that separates causal explanations of behavior from moral responsibility for behavior cuts both ways. In the latest twist in the human-nature morality play, a chromosomal marker for homosexuality in some men, the so-called gay gene, was identified by the geneticist Dean Hamer. To the bemusement of *Science for the People*, this time it is the genetic explanation that is politically correct. Supposedly it refutes right-wingers like Dan Quayle, who had said that homosexuality "is more of a choice than a biological situation. It is a wrong choice." The gay gene has been used to argue that homosexuality is not a choice for which gay people can be held responsible but an involuntary orientation they just can't help. But the reasoning is dangerous. The gay gene could just as easily be said to influence some people to *choose* homosexuality. And like all good science, Hamer's result might be falsified someday, and then where would we be? Conceding that bigotry against gay people is OK after all? The argument against persecuting gay people must be made not in terms of the gay gene or the gay brain but in terms of people's right to engage in private consensual acts without discrimination or harassment.

The cloistering of scientific and moral reasoning in separate arenas also lies behind my recurring metaphor of the mind as a machine, of people as robots. Does this not dehumanize and objectify people and lead us to treat them as inanimate objects? As one humanistic scholar lucidly put it in an Internet posting, does it not render human experience invalid, reifying a model of relating based on an I-It relationship, and delegitimizing all other forms of discourse with fundamentally destructive consequences to society? Only if one is so literal-minded that one cannot shift among different stances in conceptualizing people for different purposes. A human being is simultaneously a machine and a sentient free agent, depending on the purpose of the discussion, just as he is also a taxpayer, an insurance salesman, a dental patient, and two hundred pounds of ballast on a commuter airplane, depending on the purpose of the discussion. The mechanistic stance allows us to understand what makes us tick and how we fit into the physical universe. When those discussions wind down for the day, we go back to talking about each other as free and dignified human beings.

The confusion of scientific psychology with moral and political goals, and the resulting pressure to believe in a structureless mind, have rippled perniciously through the academy and modern intellectual discourse. Many of us have been puzzled by the takeover of humanities departments by the doctrines of postmodernism, poststructuralism, and deconstructionism, according to which objectivity is impossible, meaning is self-contradictory, and reality is socially constructed. The motives become clearer when we consider typical statements like "Human beings have constructed and used gender—human beings can deconstruct and stop using gender," and "The heterosexual/homosexual binary is not in nature, but is socially constructed, and therefore deconstructable." Reality is denied to categories, knowledge, and the world itself so that reality can be denied to stereotypes of gender, race, and sexual orientation. The doctrine is basically a convoluted way of getting to the conclusion that oppression of women, gays, and minorities is bad. And the dichotomy between "in nature" and "socially constructed" shows a poverty of the imagination, because it omits a third alternative: that some categories are products of a complex mind designed to mesh with what is in nature.

Mainstream social critics, too, can state any absurdity if it fits the Standard Social Science Model. Little boys are encouraged to argue and fight. Children learn to associate sweets with pleasure because parents use sweets as a reward for eating spinach. Teenagers compete in looks and dress because they follow the example set by spelling bees and award ceremonies. Men are socialized into believing that the goal of sex is an orgasm. Eighty-year-old women are considered less physically attractive than twenty-year-olds because our phallic culture has turned the young girl into the cult object of desire. It's not just that there is no evidence for these astonishing claims, but it is hard to credit that the authors, deep down, believe them themselves. These kinds of claims are uttered without concern for whether they are true; they are part of the secular catechism of our age.

Contemporary social commentary rests on archaic conceptions of the mind. Victims burst under the pressure, boys are conditioned to do this, women are brainwashed to value that, girls are taught to be such-and-such. Where do these explanations come from? From the nineteenth-century hydraulic model of Freud, the drooling dogs and key-pressing

vermin of behaviorism, the mind-control plots of bad cold-war movies, the wide-eyed, obedient children of *Father Knows Best*.

But when we look around us, we sense that these simplistic theories just don't ring true. Our mental life is a noisy parliament of competing factions. In dealing with others, we assume they are as complicated as we are, and we guess what they are guessing we are guessing they are guessing. Children defy their parents from the moment they are born, and confound all expectations thereafter: one overcomes horrific circumstances to lead a satisfying life, another is granted every comfort but grows up a rebel without a cause. A modern state loosens its grip, and its peoples enthusiastically take up the vendettas of their grandparents. And there are no robots.

I believe that a psychology of many computational faculties engineered by natural selection is our best hope for a grasp on how the mind works that does justice to its complexity. But I won't convince you with the opening brief in this chapter. The proof must come from insight into problems ranging from how Magic Eye stereograms work to what makes a landscape beautiful to why we find the thought of eating worms disgusting to why men kill their estranged wives. Whether or not you are persuaded by the arguments so far, I hope they have provoked your thoughts and made you curious about the explanations to come.

Conversion and Christian Growth



Rather than cause the triumph of Christianity, the emperor Constantine's "Edict of Milan" was an astute *response* to rapid Christian growth that had already made them a major political force.

FINALLY, all questions concerning the rise of Christianity are one: How was it done? How did a tiny and obscure messianic movement from the edge of the Roman Empire dislodge classical paganism and become the dominant faith of Western civilization? Although this is the only question, it requires many answers—no one thing led to the triumph of Christianity.

The chapters that follow will attempt to reconstruct the rise of Christianity in order to explain why it happened. But in this chapter I will pose the question in a more precise way than has been done. First, I shall explore the arithmetic of growth to see more clearly the task that had to be accomplished. What is the minimum rate of growth that would permit the Christian movement to become as large as it must have been in the time that history allows? Did Christianity grow so rapidly that mass conversions must have taken place—as Acts attests and every historian from Eusebius to Ramsay MacMullen has believed? Having established a plausible growth curve for the rise of Christianity, I will review sociological knowledge of the process by which people convert to new religions in order to infer certain requirements concerning social relations between Christians and the surrounding Greco-Roman world. The chapter concludes with a discussion of the legitimate uses of social scientific theories to reconstruct history in the absence of adequate information on what actually occurred.

Since this book is a work of both history and social science, I have written it for a nonprofessional audience. In this way I can make sure that the social science is fully accessible to historians of the early church, meanwhile preventing social scientists from becoming lost amidst obscure historical and textual references.

Before I proceed, however, it seems appropriate to discuss

whether an attempt to explain the rise of Christianity is not somewhat sacrilegious. If, for example, I argue that the rise of Christianity benefited from superior fertility or from an excess of females who made possible high rates of exogamous marriage, am I not, thereby, attributing sacred achievements to profane causes? I think not. Whatever one does or does not believe about the divine, obviously God did not cause the world to become Christian, since that remains to be achieved. Rather, the New Testament recounts human efforts to spread the faith. No sacrilege is entailed in the search to understand human actions in human terms. Moreover, I do not reduce the rise of Christianity to purely "material" or social factors. Doctrine receives its due—an essential factor in the religion's success was what Christians believed.

THE ARITHMETIC OF GROWTH

Studies of the rise of Christianity all stress the movement's rapid growth, but rarely are any figures offered. Perhaps this reflects the prevalence among historians of the notion, recently expressed by Pierre Chuvin, that "ancient history remains wholly refractory to quantitative evaluations" (1990:12). Granted, we shall never discover "lost" Roman census data giving authoritative statistics on the religious composition of the empire in various periods. Nevertheless, we *must quantify*—at least in terms of exploring the arithmetic of the possible—if we are to grasp the magnitude of the phenomenon that is to be explained. For example, in order for Christianity to have achieved success in the time allowed, must it have grown at rates that seem incredible in the light of modern experience? If so, then we may need to formulate new social scientific propositions about conversion. If not, then we have some well-tested propositions to draw upon. What we need is at least two plausible numbers to provide the basis for extrapolating the proba-

ble rate of early Christian growth. Having achieved such a rate and used it to project the number of Christians in various years, we can then test these projections against a variety of historical conclusions and estimates.

For a *starting* number, Acts 1:14–15 suggests that several months after the Crucifixion there were 120 Christians. Later, in Acts 4:4, a total of 5,000 believers is claimed. And, according to Acts 21:20, by the sixth decade of the first century there were "many thousands of Jews" in Jerusalem who now believed. These are not statistics. Had there been that many converts in Jerusalem, it would have been the first Christian city, since there probably were no more than twenty thousand inhabitants at this time—J. C. Russell (1958) estimated only ten thousand. As Hans Conzelmann noted, these numbers are only "meant to render impressive the marvel that here the Lord himself is at work" (1973:68). Indeed, as Robert M. Grant pointed out, "one must always remember that figures in antiquity . . . were part of rhetorical exercises" (1977:7–8) and were not really meant to be taken literally. Nor is this limited to antiquity. In 1984 a Toronto magazine claimed that there were 10,000 Hare Krishna members in that city. But when Irving Hexham, Raymond F. Currie, and Joan B. Townsend (1985) checked on the matter, they found that the correct total was 80.

Origen remarked, "Let it be granted that Christians were few in the beginning" (*Against Celsus* 3.10, 1989 ed.), but how many would that have been? It seems wise to be conservative here, and thus I shall assume that there were 1,000 Christians in the year 40. I shall qualify this assumption at several later points in the chapter.

Now for an *ending* number. As late as the middle of the third century, Origen admitted that Christians made up "just a few" of the population. Yet only six decades later, Christians were so numerous that Constantine found it expedient to embrace the church. This has caused many scholars to think that something really extraordinary, in terms of growth, happened in the later

half of the third century (cf. Gager 1975). This may explain why, of the few numbers that have been offered in the literature, most are for membership in about the year 300.

Edward Gibbon may have been the first to attempt to estimate the Christian population, placing it at no more than "a twentieth part of the subjects of the empire" at the time of Constantine's conversion ([1776-1788] 1960:187). Later writers have rejected Gibbon's figure as far too low. Goodenough (1931) estimated that 10 percent of the empire's population were Christians by the time of Constantine. If we accept 60 million as the total population at that time—which is the most widely accepted estimate (Boak 1955a; Russell 1958; MacMullen 1984; Wilken 1984)—this would mean that there were 6 million Christians at the start of the fourth century. Von Herding (1934) estimated the maximum number of Christians in the year 300 as 15 million. Grant (1978) rejected this as far too high and even rejected von Herding's minimum estimate of 7.5 million as high. MacMullen (1984) placed the number of Christians in 300 at 5 million. Fortunately, we do not need greater precision; if we assume that the actual number of Christians in the year 300 lay within the range of 5-7.5 million, we have an adequate basis for exploring what rate of growth is needed for that range to be reached in 260 years.

Given our starting number, if Christianity grew at the rate of 40 percent per decade, there would have been 7,530 Christians in the year 100, followed by 217,795 Christians in the year 200 and by 6,299,832 Christians in the year 300. If we cut the rate of growth to 30 percent a decade, by the year 300 there would have been only 917,334 Christians—a figure far below what anyone would accept. On the other hand, if we increase the growth rate to 50 percent a decade, then there would have been 37,876,752 Christians in the year 300—or more than twice von Herding's maximum estimate. Hence 40 percent per decade (or 3.42 percent per year) seems the most plausible estimate of the rate at which Christianity actually grew during the first several centuries.

TABLE 1.1
Christian Growth Projected at 40 Percent per Decade

Year	Number of Christians	Percent of Population ^a
40	1,000	0.0017
50	1,400	0.0023
100	7,530	0.0126
150	40,496	0.07
200	217,795	0.36
250	1,171,356	1.9
300	6,299,832	10.5
350	33,882,008	56.5

^a Based on an estimated population of 60 million.

This is a very encouraging finding since it is exceedingly close to the average growth rate of 43 percent per decade that the Mormon church has maintained over the past century (Stark 1984, 1994). Thus we know that the numerical goals Christianity needed to achieve are entirely in keeping with modern experience, and we are not forced to seek exceptional explanations. Rather, history allows time for the normal processes of conversion, as understood by contemporary social science, to take place.

However, before we take up the topic of conversion, it seems worthwhile to pause and consider the widespread impression that Christian growth speeded rapidly during the last half of the third century. In terms of *rate* of growth, it probably did not. But because of the rather extraordinary features of exponential curves, this probably was a period of "miraculous-seeming" growth in terms of *absolute numbers*. All of this is clear in table 1.1.

Progress must have seemed terribly slow during the first century—the projected total is only 7,530 by 100. There was a greater increase in numbers by the middle of the second century, but still the projection amounts to only slightly more than 40,000 Christians. This projection is in extremely close agree-

ment with Robert L. Wilken's estimate of "less than fifty thousand Christians" at this time—"an infinitesimal number in a society comprising sixty million" (1984:31). Indeed, according to L. Michael White (1990:110), Christians in Rome still met in private homes at this time. Then, early in the third century, the projected size of the Christian population picks up a bit and by 250 reaches 1.9 percent. This estimate is also sustained by a prominent historian's "feel" for the times. Discussing the process of conversion to Christianity, Robin Lane Fox advised that we keep "the total number of Christians in perspective: their faith was much the most rapidly growing religion in the Mediterranean, but its total membership was still small in absolute terms, perhaps (at a guess) only 2 percent of the Empire's total population by 250" (1987:317). But even more compelling is how the absolute number (as well as the percent Christian) suddenly shoots upward between 250 and 300, just as historians have reported,¹ and recent archaeological findings from Dura-Europos support this view. Excavations of a Christian building show that during the middle of the third century a house church was extensively remodeled into a building "entirely devoted to religious functions," after which "all domestic activities ceased" (White 1990:120). The renovations mainly involved the removal of partition walls to create an enlarged meeting hall—indicative of the need to accommodate more worshipers. That my reconstruction of Christian growth exhibits the "sudden spurt" long associated with the second half of the third century adds to the plausibility of the figures.

The projections are also extremely consistent with Graydon F. Snyder's (1985) assessment of all known archaeological evidence of Christianity during the first three centuries. Snyder determined that there really isn't any such evidence prior to 180. He interpreted this to indicate that before then it is impossible to distinguish Christian from non-Christian culture in "funerary art, inscriptions, letters, symbols, and perhaps buildings . . . [because] it took over a century for the new community of

faith to develop a distinctive mode of self-expression" (Snyder 1985:2). That may be, but it must also be noted that the *survival* of Christian archaeological evidence would have been roughly proportionate to how much there *could have been* to start with. The lack of anything surviving from prior to 180 must be assessed on the basis of the tiny number of Christians who could have left such traces. Surely it is not surprising that the 7,535 Christians at the end of the first century left no trace. By 180, when I project that the total Christian population first passed the 100,000 mark, there would finally have been enough Christians so that it is probable that traces of their existence would survive. Thus Snyder's findings are very compatible with my estimates of a very small Christian population in the first two centuries.

As an additional test of these projections, Robert M. Grant has calculated that there were 7,000 Christians in Rome at the end of the second century (1977:6). If we also accept Grant's estimate of 700,000 as the population of Rome for that year, then 1 percent of the population of Rome had been converted by the year 200. If we set the total population of the empire at 60 million in 200, then, based on the projection for that year, Christians constituted 0.36 percent of the empire's population. This seems to be an entirely plausible matchup, since the proportion Christian should have been higher in Rome than in the empire at large. First of all, historians assume that the church in Rome was exceptionally strong—it was well known for sending funds to Christians elsewhere. In about 170, Dionysius of Corinth wrote to the Roman church: "From the start it has been your custom to treat all Christians with unflinching kindness, and to send contributions to many churches in every city, sometimes alleviating the distress of those in need, sometimes providing for your brothers in the mines" (Eusebius, *Ecclesiastical History* 4.23.6, 1965 ed.). Second, by 200 the Christian proportion of the population of the city of Rome must have been substantially larger than that in the whole of the empire because

Christianity had not yet made much headway in the more westerly provinces. As will be seen in chapter 6, of the twenty-two largest cities in the empire, four probably still lacked a Christian church by the year 200. Although I have estimated the overall number of Christians in the empire, I am fully aware that Christian growth was concentrated in the East—in Asia Minor, Egypt, and North Africa. Moreover, there is general agreement among historians (Harnack 1908; Boak 1955a; Meeks 1988) that the Christian proportion of the population was substantially higher in cities than in the rural areas at this time—hence the term *pagani* or “countryman” came to refer to non-Christians (pagans). In any event, here too the projections closely agree with estimates based on independent sources.

Now, let us peek just a bit further into the future of Christian growth. If growth held at 40 percent per decade for the first half of the fourth century, there would have been 83,882,008 Christians by 350. In an empire having a population of at least 60 million, there might well have been 33 million Christians by 350—for by then some contemporary Christian writers were claiming a majority (Harnack 1908: 2:29). Looking at the rise of a Christian majority as purely a function of a constant rate of growth calls into serious question the emphasis given by Eusebius and others to the conversion of Constantine as the factor that produced the Christian majority (Grant 1977). So long as nothing changed in the conditions that sustained the 40-percent-a-decade growth rate, Constantine’s conversion would better be seen as a response to the massive exponential wave in progress, not as its cause.

This interpretation is entirely in keeping with the thesis developed by Shirley Jackson Case in his 1925 presidential address to the American Society of Church History. Case began by noting that attempts by the emperor Diocletian in 303, and continued by his successor Galerius in 305, to use persecution to force Christians to support the state had failed because “by the year 300 Christianity had become too widely accepted in Roman so-

ciety to make possible a successful persecution on the part of the government” (1928:59). As a result, Case continued, by 311 the emperor Galerius switched tactics and excused the Christians from praying to Roman gods, and asked only that they pray to “their own god for our security and that of the state” (Case 1928:61). Thus Constantine’s edict of toleration, issued two years later, was simply a continuation of state policy. Case’s assessment of Constantine’s edict stressed the impact of Christian growth on this policy:

In this document one perceives very easily the real basis of Constantine’s favor for Christianity. First, there is the characteristic attitude of an emperor who is seeking supernatural support for his government, and secondly, there is a recognition of the fact that the Christian element in the population is now so large, and its support for Constantine and Licinius in their conflict with rivals who still opposed Christianity, is so highly esteemed, that the emperors are ready to credit the Christian God with the exercise of a measure of supernatural power on a par with the other gods of the State. (1928:62)

It is reassuring to have the projections of Christian membership in table 1.1 fit so well with several independent estimates, with major historical perceptions such as the rapid increases during the latter part of the third century, and with the record of Mormon growth achieved over the past century. Keep in mind, however, that the numbers are *estimates*, not recorded fact. They seem very plausible, but I would be entirely comfortable with suggestions that reality may have been a bit lumpier. Perhaps growth was somewhat more rapid in the earliest days and my beginning number of 1,000 Christians in 40 is a bit low. But it also seems likely that there were periodic losses in the early days, some of which may have been very substantial for a group still so small. For example, following the execution of James and the subsequent destruction of Jerusalem, the Christian community in Palestine seems to have died out (Frend

1965, 1984). And while Tacitus's claim that "an immense multitude" (*Annals* 15.44, 1989 ed.) was butchered by Nero in about 65 is much exaggerated (see chapter 8), even the deaths of several hundred Christians would have been a very serious setback.

I have tried to offset such bumps and lumps in the growth curve by starting with a very conservative number. Moreover, my purpose in generating these numbers was not to discover "facts," but to impose needed discipline on the subject. That is, by resorting to simple arithmetic I believe I have demonstrated adequately that the rise of Christianity required no miraculous rates of conversion.

Several years after I had completed this exploration of the arithmetic of early Christian growth, when this book was nearly finished, my colleague Michael Williams made me aware of Roger S. Bagnall's remarkable reconstruction of the growth of Christianity in Egypt (1982, 1987). Bagnall examined Egyptian papyri to identify the proportion of persons with identifiable Christian names in various years, and from these he reconstructed a curve of the Christianization of Egypt. Here are *real* data, albeit from only one area, against which to test my projections. Two of Bagnall's data points are much later than the end of my projections. However, a comparison of the six years within my time frame shows a level of agreement that can only be described as extraordinary—as can be seen in table 1.2.

Bagnall's finding no Christians in 239 can be disregarded. Obviously there were Christians in Egypt then, but because their numbers would still have been very small it is not surprising that none turned up in Bagnall's data. But for later years the matchups are striking, and the correlation of 0.86 between the two curves borders on the miraculous. The remarkable fit between these two estimates, arrived at via such different means and sources, seems to me a powerful confirmation of both.

Although the projections seem very plausible through 350, the rate of Christian growth eventually must have declined rap-

TABLE 1.2
Two Estimates of Christianization Compared

Year	Projected Percent Christian in the Greco-Roman World	Percent Christian in Egypt ^a
239	1.4	0
274	4.2	2.4
278	5.0	10.5
280	5.4	13.5
313	16.2	18.0
315	17.4	18.0
		r = 0.86

^a Bagnall 1982, 1987.

idly at some point during the fourth century. If nothing else, the empire would have begun to run out of potential converts. This is evident when we realize that had the 40 percent growth rate held throughout the fourth century, there would have been 182,225,584 Christians in the year 400. Not only is that total impossible, growth rates must always decline when a movement has converted a substantial proportion of the available population—as the pool of potential converts is progressively "fished out." Or, as Bagnall put it, "the curve of conversion becomes asymptotic, and incremental conversion becomes slight after a time" (1982:123). Clearly, then, the projections from my model are invalid after the year 350. However, since my concerns only involve the *rise* of Christianity, it is not necessary to venture beyond this point.

ON CONVERSION

Eusebius tells us that early Christian missionaries were so empowered by the "divine Spirit" that "at the first hearing whole multitudes in a body eagerly embraced in their souls piety towards the Creator of the universe" (*Ecclesiastical History* 3.37.3,

1927 ed.). Not only do many modern historians of the early church accept Eusebius's claims about mass conversions in response to public preaching and miracle working, but they often regard it as a necessary assumption because of the rapidity of Christianity's rise. Thus in his distinguished study, *Christianizing the Roman Empire*, Ramsay MacMullen urged acceptance of the reports of large-scale conversions as necessary

to explain better the rate of change we are observing. In the whole process, very large numbers are obviously involved. . . . [I]t would be hard to picture the necessary scale of conversion if we limited ourselves to . . . evangelizing in private settings . . . [If this mode of conversion], however, is combined with evidence for successes en masse, the two in combination do seem to me adequate to explain what we know happened. (1984:29)

MacMullen's views reflect those of Adolf Harnack (1908: 2:335–336), who characterized the growth of Christianity in terms such as “inconceivable rapidity” and “astonishing expansion,” and who expressed his agreement with Augustine's claim that “Christianity must have reproduced itself by means of miracles, for the greatest miracle of all would have been the extraordinary extension of the religion apart from any miracles” (335n.2).

This is precisely why there is no substitute for arithmetic. The projections reveal that Christianity could easily have reached half the population by the middle of the fourth century without miracles or conversions en masse. The Mormons have, thus far, traced the same growth curve, and we have no knowledge of their achieving mass conversions. Moreover, the claim that mass conversions to Christianity took place as crowds spontaneously responded to evangelists assumes that doctrinal appeal lies at the heart of the conversion process—that people hear the message, find it attractive, and embrace the faith. But modern social science relegates doctrinal appeal to a very secondary role, claiming that most people do not really become very

attached to the doctrines of their new faith until *after* their conversion.

In the early 1960s John Lofland and I were the first social scientists to actually go out and watch people convert to a new religious movement (Lofland and Stark 1965). Up to that time, the most popular social scientific explanation of conversion involved the pairing of deprivation with ideological (or theological) appeal. That is, one examined the ideology of a group to see what kinds of deprivation it addressed and then concluded (*mirabile dictu!*) that converts suffered from those deprivations (Glock 1964). As an example of this approach, since Christian Science promised to restore health, its converts *must* disproportionately be drawn from among those with chronic health problems, or at least those who suffer from hypochondria (Glock 1964). Of course, one could as plausibly argue the reverse, that only people with excellent health could long hold to the Christian Science doctrine that illness was all in the mind.

In any event, Lofland and I were determined to watch people go through the process of conversion and try to discover what really was involved. Moreover, we wanted to watch conversion, not simply activation. That is, we wanted to look at people who were making a major religious shift, as from Christianity to Hinduism, rather than examine how lifelong Christians got themselves born again. The latter is a matter of considerable interest, but it was not our interest at the time.

We also wanted a group that was small enough so that the two of us could provide adequate surveillance, and new enough so that it was in an early and optimistic phase of growth. After sifting through many deviant religious groups in the San Francisco Bay area we came upon precisely what we were looking for—a group of about a dozen young adults who had just moved to San Francisco from Eugene, Oregon. The group was led by Young Oon Kim, a Korean woman who had once been a professor of religion at Ewha University in Seoul. The movement she served was based in Korea, and in January 1959, she arrived in Oregon

to launch a mission to America. Miss² Kim and her young followers were the very first American members of the Unification Church, widely known today as the Moonies.

As Lofland and I settled back to watch people convert to this group, the first thing we discovered was that all of the current members were united by close ties of friendship predating their contact with Miss Kim. Indeed, the first three converts had been young housewives, next-door neighbors who became friends of Miss Kim after she became a lodger with one of them. Subsequently, several of the husbands joined, followed by several of their friends from work. At the time Lofland and I arrived to study them, the group had never succeeded in attracting a stranger.

Lofland and I also found it interesting that although all the converts were quick to describe how their spiritual lives had been empty and desolate prior to their conversion, many claimed they had not been particularly interested in religion before. One man told me, "If anybody had said I was going to join up and become a missionary I would have laughed my head off. I had no use for church at all."

We also found it instructive that during most of her first year in America, Miss Kim had tried to spread her message directly by talks to various groups and by sending out many press releases. Later, in San Francisco the group also tried to attract followers through radio spots and by renting a hall in which to hold public meetings. But these methods yielded nothing. As time passed, Lofland and I were able to observe people actually becoming Moonies. The first several converts were old friends or relatives of members who came from Oregon for a visit. Subsequent converts were people who formed close friendships with one or more members of the group.

We soon realized that of all the people the Moonies encountered in their efforts to spread their faith, the only ones who joined were those whose *interpersonal attachments to members overbalanced their attachments to nonmembers*. In effect, conversion is

not about seeking or embracing an ideology; it is about bringing one's religious behavior into alignment with that of one's friends and family members.

This is simply an application of the highly respected control theory of deviant behavior (Toby 1957; Hirschi 1969; Stark and Bainbridge 1987; Gottfredson and Hirschi 1990). Rather than asking why people deviate, why they break laws and norms, control theorists ask why anyone ever does conform. Their answer is posed in terms of *stakes in conformity*. People conform when they believe they have more to lose by being detected in deviance than they stand to gain from the deviant act. Some people deviate while others conform because people differ in their stakes in conformity. That is, some people simply have far less to lose than do others. A major stake in conformity lies in our attachments to other people. Most of us conform in order to retain the good opinion of our friends and family. But some people lack attachments. Their rates of deviance are much higher than are those of people with an abundance of attachments.

Becoming a Moonie today is an act of deviance, as was becoming a Christian in the first century. Such conversions violate norms defining legitimate religious affiliations and identities. Lofland and I saw many people who spent some time with the Moonies and expressed considerable interest in their doctrines, but who never joined. In every instance these people had many strong attachments to nonmembers who did not approve of the group. Of persons who did join, many were newcomers to San Francisco whose attachments were all to people far away. As they formed strong friendships with group members, these were not counterbalanced because distant friends and families had no knowledge of the conversion-in-process. In several instances a parent or sibling came to San Francisco intending to intervene after having learned of the conversion. Those who lingered eventually joined up too. Keep in mind that becoming a Moonie may have been regarded as deviant by outsiders, but

it was an act of conformity for those whose most significant attachments were to Moonies.

During the quarter century since Lofland and I first published our conclusion—that attachments lie at the heart of conversion and therefore that conversion tends to proceed along social networks formed by interpersonal attachments—many others have found the same to be true in an immense variety of religious groups all around the world. A recent study based on Dutch data (Kox, Meeus, and 't Hart 1991) cited twenty-five additional empirical studies, all of which supported our initial finding. And that list was far from complete.

Although several other factors are also involved in the conversion process, the central sociological proposition about conversion is this: *Conversion to new, deviant religious groups occurs when, other things being equal, people have or develop stronger attachments to members of the group than they have to nonmembers* (Stark 1992).

Data based on records kept by a Mormon mission president give powerful support to this proposition. When missionaries make cold calls, knock on the doors of strangers, this eventually leads to a conversion once out of a thousand calls. However, when missionaries make their first contact with a person in the home of a Mormon friend or relative of that person, this results in conversion 50 percent of the time (Stark and Bainbridge 1985).

A variation on the network proposition about conversion is that successful founders of new faiths typically turn first to those with whom they already have strong attachments. That is, they recruit their first followers from among their family and close friends. Thus Muhammad's first convert was his wife Khadijah; the second was his cousin Ali, followed by his servant Zeyd and then his old friend Abu Bakr. On April 6, 1830, the Mormons were founded by Joseph Smith, his brothers Hyrum and Samuel, and Joseph Smith's friends Oliver Cowdery and David and Peter Whitmer. The rule extends to Jesus too, since it appears that he began with his brothers and mother.

A second aspect of conversion is that people who are deeply committed to any particular faith do not go out and join some other faith. Thus Mormon missionaries who called upon the Moonies were immune, despite forming warm relationships with several members. Indeed, the Moonie who previously had "no use for church at all" was more typical. Converts were not former atheists, but they were essentially unchurched and many had not paid any particular attention to religious questions. Thus the Moonies quickly learned that they were wasting their time at church socials or frequenting denominational student centers. They did far better in places where they came in contact with the uncommitted. This finding has received substantial support from subsequent research. Converts to new religious movements are overwhelmingly from relatively irreligious backgrounds. The majority of converts to modern American cult movements report that their parents had no religious affiliation (Stark and Bainbridge 1985). Let me state this as a theoretical proposition: *New religious movements mainly draw their converts from the ranks of the religiously inactive and discontented, and those affiliated with the most accommodated (worldly) religious communities.*

Had we not gone out and watched people as they converted, we might have missed this point entirely, because when people retrospectively describe their conversions, they tend to put the stress on theology. When asked why they converted, Moonies invariably noted the irresistible appeal of the Divine Principles (the group's scripture), suggesting that only the blind could reject such obvious and powerful truths. In making these claims converts implied (and often stated) that their path to conversion was the end product of a search for faith. But Lofland and I knew better because we had met them well before they had learned to appreciate the doctrines, before they had learned how to testify to their faith, back when they were not seeking faith at all. Indeed, we could remember when most of them regarded the religious beliefs of their new set of friends as quite odd. I recall one who told me that he was puzzled that such nice

people could get so worked up about "some guy in Korea" who claimed to be the Lord of the Second Advent. Then, one day, he got worked up about this guy too. I suggest that this is also how people in the first century got themselves worked up about someone who claimed to be the Lord of the First Advent. Robin Lane Fox suggests the same thing: "Above all we should give weight to the presence and influence of friends. It is a force which so often escapes the record, but it gives shape to everyone's personal life. One friend might bring another to the faith. . . . When a person turned to God, he found others, new 'brethren,' who were sharing the same path" (1987:316). Peter Brown has expressed similar views: "Ties of family, marriages, and loyalties to heads of households had been the most effective means of recruiting members of the church, and had maintained the continued adherence of the average Christian to the new cult" (1988:90).

The basis for successful conversionist movements is growth through social networks, through a *structure of direct and intimate interpersonal attachments*. Most new religious movements fail because they quickly become closed, or semiclosed networks. That is, they fail to keep forming and sustaining attachments to outsiders and thereby lose the capacity to grow. Successful movements discover techniques for remaining open networks, able to reach out and into new adjacent social networks. And herein lies the capacity of movements to sustain exponential rates of growth over a long period of time.

Some readers may suspect that the rapid rise in the absolute number of new Christians between 250 and 350 would require mass conversions even though the rate of conversion remained constant at 40 percent per decade. Admittedly, exponential growth curves are counterintuitive and easily seem incredible. Nevertheless, the *dynamics of the conversion process* are not changed even as the absolute numbers reach a rapid growth stage along an exponential curve. The reason is that as movements grow, their social surface expands proportionately. That

is, each new member expands the size of the network of attachments between the group and potential converts. As noted above, however, this occurs *only* if the group constitutes an *open network*. Thus if we are to better understand and explain the rise of Christianity, we must discover how the early Christians maintained open networks—for it would seem certain that they did. This last remark sets the stage for a brief discussion of the appropriate scope of social scientific theories and whether it is possible even to apply propositions developed in one time and place to other eras and cultures.

ON SCIENTIFIC GENERALIZATION

Many historians believe that cultures and eras verge on the unique. Thus in his very thoughtful response to my use of the network theory of conversion to discuss the success of the mission to the Jews (see chapter 3), Ronald F. Hock noted that I seem to think that networks, for example, are not "all that different from period to period, society to society" (1986:2–3). He then pointed out that

the networks utilized by Mormons are those consisting of a member's family, relatives, and friends, but are ancient networks the same? Ancient cities are not modern ones, and ancient networks that were centered in aristocratic households included more than family and friends: domestic slaves, freedmen, and perhaps parasites, teachers, athletic trainers, and travelers. In addition, urban life was lived more in public, so that recruitment could proceed along more extensive and complex networks than we find among Mormons in our more nuclear and anonymous cities and suburbs.

I am certain that Hock is correct, but I am unrepentant. What he is noting are details that might tell us how to discover networks should we be transported to ancient Antioch, but

that have no implications for the network proposition *per se*. However people constitute structures of direct interpersonal attachments, those structures will define the lines through which conversion will most readily proceed. The definition of network is not locked to time and space, nor is the conversion proposition.

Many historians seem to have considerable trouble with the idea of general theories because they have not been trained in the distinction between concepts and instances. Proper scientific concepts are abstract and identify a class of "things" to be regarded as alike. As such, concepts must apply to all possible members of the class, all that have been, are, shall be, or could be. The concept of chair, defined as all objects created to seat a lone individual and support his or her back, is an abstraction. We cannot see the concept of chair. It is an intellectual creation existing only in our minds. But we can see many actual chairs, and as we look at some, we discover immense variation in size, shape, materials, color, and the like. Moreover, when we look at chairs used in the ancient world, we perceive some very noticeable differences from the chairs of today. Nevertheless, each is a chair so long as it meets the definition set out above—other somewhat similar objects belong to other object classes such as stools and couches.

These points apply as fully to the concept of social network as to the concept of chair. The concept of social network also exists only in our minds. All that we can see are specific instances of the class—networks involving some set of individuals. As with chairs, the shapes and sizes of social networks may differ greatly across time and space, and the processes by which networks form may vary as greatly as do techniques for making chairs. But these variations in details never result in chairs' becoming pianos, nor do variations in their makeup ever turn social networks into collections of strangers.

It is only through the use of abstract concepts, linked by abstract propositions, that science exists. Consider a physics that

must generate a new rule of gravity for each object in the universe. And it is precisely the abstract generality of science that makes it possible for social science to contribute anything to our understanding of history, let alone to justify efforts to reconstruct history from social scientific theories. Let me now turn to that important issue.

SOCIAL THEORY AND HISTORICAL RECONSTRUCTIONS

During the past several decades historians of the New Testament era have become increasingly familiar with social science and have become increasingly inclined to use social scientific models to infer "what must have happened" in order to fill blanks in the historical and archaeological record. As Robin Scroggs pointed out in an influential essay, "there may be times when a sociological model may actually assist our ignorance. If our data evidence some *parts* of the gestalt of a known model, while being silent about others, we *may* cautiously be able to conclude that the absence of the missing parts is accidental and that the entire model was actually a reality in the early church" (1980:166). Since those lines were published, the practice Scroggs suggested has become common (Barton 1982, 1984; Holmberg 1980; Elliott 1986; Fox 1987; Gager 1975, 1983; Green 1985; Malina 1981, 1986; Meeks 1983, 1993; Kee 1983; Kraemer 1992; Sanders 1993; Theissen 1978, 1982; Wilken 1984; Wire 1991). I have quite mixed reactions to this literature. Some studies I have read with pleasure and admiration. Other examples have made me very uncomfortable because the social science "models" utilized are so inadequate. Some of them are merely metaphors—as Durkheim's "discovery" that religion is society worshiping itself is merely metaphor. How would one falsify that statement, or assertions to the effect that religion is a neurotic illusion or the poetry of the soul? The

problem with metaphors is not that they are false, but that they are *empty*. Many of them do seem to ooze profundity, but at best metaphors are merely definitions. Consider the term *charisma*.

Max Weber borrowed this Greek word meaning "divine gift" to identify the ability of some people to convince others that their authority is based on divine sources: "The holder of charisma seizes the task that is adequate for him and demands obedience and a following by virtue of his mission. His success determines whether he finds them. His charismatic claim breaks down if his mission is not recognized by those to whom he feels he has been sent. If they recognize him, he is their master" (1946:216). Charisma is commonly observed in religious leaders, and surely no one would dispute that Jesus and many of the apostles and early evangelists had it. Thus the literature on the early church is saturated with the term. Unfortunately, charisma is too often understood as a nearly magical power possessed by individuals rather than a description of how they are regarded. That is, their power over others is attributed to their charisma, and it is often suggested that particular religious leaders are so potent *because* they had charisma. Roy Wallis, for example, claimed that Moses David (David Berg), founder of the Children of God, maintained control over his followers because of his "charismatic status" (1982:107). But this is entirely circular. It is the same as saying that people believed that Moses David had divine authority because people believed he had divine authority. Because Weber's discussions of charisma did not move beyond definitional and descriptive statements, and said nothing about the causes of charisma, the concept is merely a name attached to a definition. When we see someone whose authority is believed by some people to be of divine origin, we have the option of calling this charisma, but doing so will contribute nothing to our understanding of why this phenomenon occurs. Hence when studies of the early church utilize the term *charisma*, what we usually confront is only a name that too often is thought to explain something, but does not.

Besides metaphors and simple concepts, other "models" used in this literature are nothing but typologies or *sets* of concepts. One of the most popular of these consists of various definitions to distinguish religious groups as *churches* or *sects*. The most useful of these definitions identifies churches and sects as the end points of a continuum based on the degree of tension between the group and its sociocultural environment (Johnson 1968; Stark and Bainbridge 1979, 1987). Sects are religious groups in a relatively high state of tension with their environment; churches are groups in a relatively low state of tension. These are very useful concepts. Unfortunately, they are often used, even by many social scientists, as if they explained something. All such efforts are circular. Thus it is circular to say that a particular religious body rejects the world *because* it is a sect, as Bryan Wilson (1970) often does, since bodies are classified as sects because they reject the world. The concepts of church and sect do nothing more (or less) than allow us to classify various religious bodies. But theories using these concepts do not reside in the concepts themselves. For example, it is well known that religious bodies, especially if they are successful, tend to move from a higher to a lower state of tension—sects often are transformed into churches. But no explanation of this transformation can be found in the definitions of church and sect. Instead, we must use propositions to link the concepts of church and sect to other concepts, such as upward social mobility and regression to the mean, in order to formulate an explanation (Stark and Bainbridge 1985, 1987).

Let me emphasize: concepts are *names*, not *explanations*. The act of naming some objects or phenomena tells us nothing about why they occur or what they influence. Explanation requires theories: abstract statements saying *why* and *how* some set of phenomena are linked, and from which falsifiable statements can be derived (Popper 1959, 1962). Metaphors, typologies, and concepts are passive; they cast no light of their own and cannot illuminate the dark corners of unrecorded history

(Stark and Bainbridge 1979, 1985, 1987). Granted, concepts may permit some useful comparisons among some sets of phenomena—comparisons of the social class composition of two religious movements, for example, can be very revealing. But if a model is to provide more than *classification*, if it proposes to *explain*, then the model must include not simply concepts, but propositions. The difference here is that between a parts catalog and a working diagram of an engine. That is, a model must include a fully specified set of interrelations among the parts. Such a model explains why and how things fit together and function. For this task, only a theory, not a conceptual scheme, suffices.

It is not surprising that scholars trained in history and in textual interpretation might find themselves more comfortable with an older generation of social "scientists" who dealt in metaphors rather than scientific theories, if for no other reason than that their work abounds in literary allusions and is recolent of ancient library dust. But let it be noted that in science, unlike papyrology, older seldom is better. And I regard it as an essential part of my task in this book to familiarize historians of the early church with more powerful and modern social scientific tools, and particularly with real theories rather than with concepts, metaphors, and typologies pretending to have explanatory power.

However, even if we use the best social science theories as our guide for reconstructing history, we are betting that the theories are solid and that the application is appropriate. When those conditions are met, then there is no reason to suppose that we cannot reason from the general rule to deduce the specific in precisely the same way that we can reason from the principles of physics that coins dropped in a well will go to the bottom. Even so, it is better when we can actually see the coins go down. *Need* is the only justification for the application of social science to fill in historical blanks. But we must be very cautious not to fill the blanks with fantasy and science fiction.

In this book I shall attempt to reconstruct the rise of Christianity on the basis of many inferences from modern social scientific theories, making particular use of my own formal theorizing about religion and religious movements (Stark and Bainbridge 1979, 1980, 1985, 1987; Stark and Iannaccone 1991, 1992). I will frequently employ the arithmetic of the possible and the plausible to test various assumptions. To guard against error I shall test my reconstructions against the historical record whenever possible, as I have done in this chapter.³

THE BLANK SLATE

*The Modern Denial
of Human Nature*



Steven Pinker

VIKING
NYC, 2002

VISUAL PERCEPTION IS the most piquant form of knowledge of the world, but relativists are less concerned with how we see objects than with how we *categorize* them: how we sort our experiences into conceptual categories like birds, tools, and people. The seemingly innocuous suggestion that the categories of the mind correspond to something in reality became a contentious idea in the twentieth century because some categories—stereotypes of race, gender, ethnicity, and sexual orientation—can be harmful when they are used to discriminate or oppress.

The word *stereotype* originally referred to a kind of printing plate. Its current sense as a pejorative and inaccurate image standing for a category of people was introduced in 1922 by the journalist Walter Lippmann. Lippmann was an important public intellectual who, among other things, helped to found *The New Republic*, influenced Woodrow Wilson's policies at the end of World War I, and wrote some of the first attacks on IQ testing. In his book *Public Opinion*, Lippmann fretted about the difficulty of achieving true democracy in an age in which ordinary people could no longer judge public issues rationally because they got their information in what we today call sound bites. As part of this argument, Lippmann proposed that ordinary people's concepts of social groups were stereotypes: mental pictures that are incomplete, biased, insensitive to variation, and resistant to disconfirming information.

Lippmann had an immediate influence on social science (though the subtleties and qualifications of his original argument were forgotten). Psychologists gave people lists of ethnic groups and lists of traits and asked them to pair them up. Sure enough, people linked Jews with "shrewd" and "mercenary," Germans with "efficient" and "nationalistic," Negroes with "superstitious" and "happy-go-lucky," and so on.⁷ Such generalizations are pernicious when applied to individuals, and though they are still lamentably common in much of

the world, they are now actively avoided by educated people and by mainstream public figures.

By the 1970s, many thinkers were not content to note that stereotypes about categories of people can be inaccurate. They began to insist that the categories themselves don't exist other than in our stereotypes. An effective way to fight racism, sexism, and other kinds of prejudice, in this view, is to deny that conceptual categories about people have any claim to objective reality. It would be impossible to believe that homosexuals are effeminate, blacks superstitious, and women passive if there were no such things as categories of homosexuals, blacks, or women to begin with. For example, the philosopher Richard Rorty has written, "The homosexual, 'the Negro,' and 'the female' are best seen not as inevitable classifications of human beings but rather as inventions that have done more harm than good."⁸

For that matter, many writers think, why stop there? Better still to insist that *all* categories are social constructions and therefore figments, because that would *really* make invidious stereotypes figments. Rorty notes with approval that many thinkers today "go on to suggest that quarks and genes probably are [inventions] too." Postmodernists and other relativists attack truth and objectivity not so much because they are interested in philosophical problems of ontology and epistemology but because they feel it is the best way to pull the rug out from under racists, sexists, and homophobes. The philosopher Ian Hacking provides a list of almost forty categories that have recently been claimed to be "socially constructed." The prime examples are race, gender, masculinity, nature, facts, reality, and the past. But the list has been growing and now includes authorship, AIDS, brotherhood, choice, danger, dementia, illness, Indian forests, inequality, the Landsat satellite system, the medicalized immigrant, the nation-state, quarks, school success, serial homicide, technological systems, white-collar crime, women refugees, and Zulu nationalism. According to Hacking, the common thread is a conviction that the category is not determined by the nature of things and therefore is not inevitable. The further implication is that we would be much better off if it were done away with or radically transformed.⁹

This whole enterprise is based on an unstated theory of human concept formation: that conceptual categories bear no systematic relation to things in the world but are socially constructed (and can therefore be reconstructed). Is it a correct theory? In some cases it has a grain of truth. As we saw in Chapter 4, some categories really are social constructions: they exist only because people tacitly agree to act as if they exist. Examples include money, tenure, citizenship, decorations for bravery, and the presidency of the United States.¹⁰ But that does not mean that *all* conceptual categories are socially constructed. Concept formation has been studied for decades by cognitive psychologists, and they conclude that most concepts pick out categories of objects in the

world which had some kind of reality before we ever stopped to think about them.¹¹

Yes, every snowflake is unique, and no category will do complete justice to every one of its members. But intelligence depends on lumping together things that share properties, so that we are not flabbergasted by every new thing we encounter. As William James wrote, "A polyp would be a conceptual thinker if a feeling of 'Hollo! thingumbob again!' ever flitted through its mind." We perceive some traits of a new object, place it in a mental category, and infer that it is likely to have the other traits typical of that category, ones we cannot perceive. If it walks like a duck and quacks like a duck, it probably is a duck. If it's a duck, it's likely to swim, fly, have a back off which water rolls, and contain meat that's tasty when wrapped in a pancake with scallions and hoisin sauce.

This kind of inference works because the world really does contain ducks, which really do share properties. If we lived in a world in which walking quacking objects were no more likely to contain meat than did any other object, the category "duck" would be useless and we probably would not have evolved the ability to form it. If you were to construct a giant spreadsheet in which the rows and columns were traits that people notice and the cells were filled in by objects that possess that combination of traits, the pattern of filled cells would be lumpy. You would find lots of entries at the intersection of the "quacks" row and the "waddles" column but none at the "quacks" row and the "gallops" column. Once you specify the rows and columns, the lumpiness comes from the world, not from society or language. It is no coincidence that the same living things tend to be classified together by the words in European cultures, the words for plant and animal kinds in other cultures (including preliterate cultures), and the Linnaean taxa of professional biologists equipped with calipers, dissecting tools, and DNA sequencers. Ducks, biologists say, are several dozen species in the subfamily Anatinae, each with a distinct anatomy, an ability to interbreed with other members of their species, and a common ancestor in evolutionary history.

Most cognitive psychologists believe that conceptual categories come from two mental processes.¹² One of them notices clumps of entries in the mental spreadsheet and treats them as categories with fuzzy boundaries, prototypical members, and overlapping similarities, like the members of a family. That's why our mental category "duck" can embrace odd ducks that don't match the prototypical duck, such as lame ducks, who cannot swim or fly, Muscovy ducks, which have claws rather than webbed feet, and Donald Duck, who talks and wears clothing. The other mental process looks for crisp rules and definitions and enters them into chains of reasoning. The second system can learn that true ducks molt twice a season and have overlapping scales on their legs and hence that certain birds that look like geese and are called geese really are ducks. Even when people don't know these facts from academic

biology, they have a strong intuition that species are defined by an internal essence or hidden trait that lawfully gives rise to its visible features.¹³

Anyone who teaches the psychology of categorization has been hit with this question from a puzzled student: "You're telling us that putting things into categories is rational and makes us smart. But we've always been taught that putting *people* into categories is irrational and makes us sexist and racist. If categorization is so great when we think about ducks and chairs, why is it so terrible when we think about genders and ethnic groups?" As with many ingenuous questions from students, this one uncovers a shortcoming in the literature, not a flaw in their understanding.

The idea that stereotypes are inherently irrational owes more to a condescension toward ordinary people than it does to good psychological research. Many researchers, having shown that stereotypes existed in the minds of their subjects, assumed that the stereotypes had to be irrational, because they were uncomfortable with the possibility that some trait might be statistically true of some group. They never actually checked. That began to change in the 1980s, and now a fair amount is known about the accuracy of stereotypes.¹⁴

With some important exceptions, stereotypes are in fact *not* inaccurate when assessed against objective benchmarks such as census figures or the reports of the stereotyped people themselves. People who believe that African Americans are more likely to be on welfare than whites, that Jews have higher average incomes than WASPs, that business students are more conservative than students in the arts, that women are more likely than men to want to lose weight, and that men are more likely than women to swat a fly with their bare hands, are not being irrational or bigoted. Those beliefs are correct. People's stereotypes are generally consistent with the statistics, and in many cases their bias is to *underestimate* the real differences between sexes or ethnic groups.¹⁵ This does not mean that the stereotyped traits are unchangeable, of course, or that people think they are unchangeable, only that people perceive the traits fairly accurately at the time.

Moreover, even when people believe that ethnic groups have characteristic traits, they are never mindless stereotypers who literally believe that each and every member of the group possesses those traits. People may think that Germans are, on average, more efficient than non-Germans, but no one believes that every last German is more efficient than every non-German.¹⁶ And people have no trouble overriding a stereotype when they have good information about an individual. Contrary to a common accusation, teachers' impressions of their individual pupils are not contaminated by their stereotypes of race, gender, or socioeconomic status. The teachers' impressions accurately reflect the pupil's performance as measured by objective tests.¹⁷

Now for the important exceptions. Stereotypes can be downright inaccurate when a person has few or no firsthand encounters with the stereotyped

group, or belongs to a group that is overtly hostile to the one being judged. During World War II, when the Russians were allies of the United States and the Germans were enemies, Americans judged Russians to have more positive traits than Germans. Soon afterward, when the alliances reversed, Americans judged Germans to have more positive traits than Russians.¹⁸

Also, people's ability to set aside stereotypes when judging an individual is accomplished by their conscious, deliberate reasoning. When people are distracted or put under pressure to respond quickly, they are more likely to judge that a member of an ethnic group has all the stereotyped traits of the group.¹⁹ This comes from the two-part design of the human categorization system mentioned earlier. Our network of fuzzy associations naturally reverts to a stereotype when we first encounter an individual. But our rule-based categorizer can block out those associations and make deductions based on the relevant facts about that individual. It can do so either for practical reasons, when information about a group-wide average is less diagnostic than information about the individual, or for social and moral reasons, out of respect for the imperative that one *ought* to ignore certain group-wide averages when judging an individual.

The upshot of this research is not that stereotypes are always accurate but that they are not always false, or even usually false. This is just what we would expect if human categorization—like the rest of the mind—is an adaptation that keeps track of aspects of the world that are relevant to our long-term well-being. As the social psychologist Roger Brown pointed out, the main difference between categories of people and categories of other things is that when you use a prototypical exemplar to stand for a category of things, no one takes offense. When Webster's dictionary used a sparrow to stand for all birds, "emus and ostriches and penguins and eagles did not go on the attack." But just imagine what would have happened if Webster's had used a picture of a soccer mom to illustrate *woman* and a picture of a business executive to illustrate *man*. Brown remarks, "Of course, people would be right to take offense since a prototype can never represent the variation that exists in natural categories. It's just that birds don't care but people do."²⁰

What are the implications of the fact that many stereotypes are statistically accurate? One is that contemporary scientific research on sex differences cannot be dismissed just because some of the findings are consistent with traditional stereotypes of men and women. Some parts of those stereotypes may be false, but the mere fact that they are stereotypes does not prove that they are false in every respect.

The partial accuracy of many stereotypes does not, of course, mean that racism, sexism, and ethnic prejudice are acceptable. Quite apart from the democratic principle that in the public sphere people should be treated as individuals, there are good reasons to be concerned about stereotypes.

Stereotypes based on hostile depictions rather than on firsthand experience are bound to be inaccurate. And some stereotypes are accurate only because of self-fulfilling prophecies. Forty years ago it may have been factually correct that few women and African Americans were qualified to be chief executives or presidential candidates. But that was only because of barriers that prevented them from attaining those qualifications, such as university policies that refused them admission out of a belief that they were not qualified. The institutional barriers had to be dismantled before the facts could change. The good news is that when the facts do change, people's stereotypes can change with them.

What about policies that go farther and actively compensate for prejudicial stereotypes, such as quotas and preferences that favor underrepresented groups? Some defenders of these policies assume that gatekeepers are incurably afflicted with baseless prejudices, and that quotas must be kept in place forever to neutralize their effects. The research on stereotype accuracy refutes that argument. Nonetheless, the research might support a different argument for preferences and other gender- and color-sensitive policies. Stereotypes, even when they are accurate, might be self-fulfilling, and not just in the obvious case of institutionalized barriers like those that kept women and African Americans out of universities and professions. Many people have heard of the Pygmalion effect, in which people perform as other people (such as teachers) expect them to perform. As it happens, the Pygmalion effect appears to be small or nonexistent, but there are more subtle forms of self-fulfilling prophecies.²¹ If subjective decisions about people, such as admissions, hiring, credit, and salaries, are based in part on group-wide averages, they will conspire to make the rich richer and the poor poorer. Women are marginalized in academia, making them genuinely less influential, which increases their marginalization. African Americans are treated as poorer credit risks and denied credit, which makes them less likely to succeed, which makes them poorer credit risks. Race- and gender-sensitive policies, according to arguments by the psychologist Virginia Valian, the economist Glenn Loury, and the philosopher James Flynn, may be needed to break the vicious cycle.²²

Pushing in the other direction is the finding that stereotypes are least accurate when they pertain to a coalition that is pitted against one's own in hostile competition. This should make us nervous about identity politics, in which public institutions identify their members in terms of their race, gender, and ethnic group and weigh every policy by how it favors one group over another. In many universities, for example, minority students are earmarked for special orientation sessions and encouraged to view their entire academic experience through the lens of their group and how it has been victimized. By implicitly pitting one group against another, such policies may cause each group to brew stereotypes about the other that are more pejorative than the

ones they would develop in personal encounters. As with other policy issues I examine in this book, the data from the lab do not offer a thumbs-up or thumbs-down verdict on race- and gender-conscious policies. But by highlighting the features of our psychology that different policies engage, the findings can make the tradeoffs clearer and the debates better informed.

The Market for News

By SENDHIL MULLAINATHAN AND ANDREI SHLEIFER*

We investigate the market for news under two assumptions: that readers hold beliefs which they like to see confirmed, and that newspapers can slant stories toward these beliefs. We show that, on the topics where readers share common beliefs, one should not expect accuracy even from competitive media: competition results in lower prices, but common slanting toward reader biases. On topics where reader beliefs diverge (such as politically divisive issues), however, newspapers segment the market and slant toward extreme positions. Yet in the aggregate, a reader with access to all news sources could get an unbiased perspective. Generally speaking, reader heterogeneity is more important for accuracy in media than competition per se. (JEL D23, L82)

Several recent books have accused mainline media outlets of reporting news with a heavy political bias. Bernard Goldberg (2002) and Ann Coulter (2003) argue that the bias is on the left, and provide numerous illustrations of their argument, while Eric Alterman (2003) and Al Franken (2003) argue that the bias is on the right, with equally numerous illustrations. In principle, media bias can come from the supply side, and reflect the preferences of journalists (David Baron, 2004), editors, or owners (Besley and Andrea Prat, 2004; Simeon Djankov et al., 2003). Alternatively, it can come from the demand side, and reflect the news providers' profit-maximizing choice to cater to the preferences of the consumers. We examine, theoretically, the determinants of media accuracy in such a demand-side model, focusing specifically on the effects of reader beliefs, reader heterogeneity, and competition on media bias. We argue that the analysis of media accuracy

relies crucially on how one conceptualizes the demand for news.

In the traditional conception of the demand for news, consumers read, watch, and listen to the news in order to get information. The quality of this information is its accuracy. The more accurate the news, the more valuable is its source to the consumer. Pressure from audiences and rivals forces news outlets to seek and deliver more accurate information, just as market forces motivate auto-makers to produce better cars.¹

This conception of the news as a source of pure information is dramatically different from that of noneconomists studying the media. According to these scholars, private media want to sell newspapers and television programs, as well as advertising space. To do that, they provide a great deal of pure entertainment. But even with news, audiences want their sources not only to inform but also to explain, interpret, persuade, and entertain. To meet this demand, media outlets do not provide unadulterated information, but rather tell stories that hang together and have a point of view, what is referred to in the business as "the narrative imperative."²

¹ Ronald Coase (1974), Besley and Robin Burgess (2002), Besley and Prat (2002), Djankov et al. (2003), David Stromberg (2001), and Alexander Dyck and Luigi Zingales (2002) all advance this view of competition in the media as delivering greater accuracy.

² H. L. Mencken (1920), Walter Lippmann (1922), Samuel Hayakawa (1940), Michael Jensen (1979), Doris Graber (1984), James Hamilton (2003), and the standard

In this view, news provision can be analyzed in the same way as entertainment broadcasting.³

In this paper, we examine these two conceptions of what the consumers want and what the media deliver, and evaluate media accuracy under different scenarios. We show, in particular, that these two conceptions have radically different implications for the accuracy of news in the competitive media, and more specifically on the question of which news issues will be reported more accurately.

Our model of rational readers seeking information shows that, indeed, consistent with economists' priors, media reporting is unbiased. We compare this to a specific behavioral model (of which the rational consumers are a special case), which relies on two assumptions, one about reader preferences and one about the technology of delivering news.⁴ We assume that readers hold biased beliefs, which might come from their general knowledge and education, from previous news, from prejudices and stereotypes, or from the views of politicians or political parties they trust. With respect to preferences, we assume that readers prefer to hear or read news that is more consistent with their beliefs. Such biased readers might believe, for example, that corporate executives are cheats and crooks, and these readers prefer news about their indictments to news about their accomplishments. They might think that China is up to no good with respect to the United States, and appreciate stories about Chinese spies. Some readers might like President Bill Clinton and prefer to read about partisan Republicans persecuting the hard-working president; others might dislike Clinton and look for stories explaining, in salacious detail, the impeachability of his offenses.

The idea that people appreciate, find credible, enjoy, and remember stories consistent with

their beliefs is standard in the communications literature (Graber, 1984; Severin and Tankard, 1992). Basic research in psychology strongly supports it. Research on memory suggests that people tend to remember information consistent with their beliefs better than information inconsistent with their beliefs (Frederic Bartlett, 1932). Research on information processing shows people find data inconsistent with their beliefs to be less credible and update less as a result (Charles Lord et al., 1979; John Zaller, 1992; Matthew Rabin and Joel Schrag, 1999). According to Graber (1984, p. 130), "stories about economic failures in third world countries were processed more readily than stories about economic successes." People seek information that confirms their beliefs (Josh Klayman, 1995). When people categorize, they tend to ignore category-inconsistent information unless it is large enough to induce category change (Susan Fiske, 1995; Mullainathan, 2002). Severin and Tankard (1992) see the demand for cognitive consistency as crucially shaping which news people listen to, and which they ignore.

Our second assumption is that newspapers can slant the presentation of the news to cater to the preferences of their audiences. The term "slanting" was introduced by Hayakawa (1940), and defined as "the process of selecting details that are favorable or unfavorable to the subject being described." Slanting is easily illustrated in a simple example. Suppose that the Bureau of Labor Statistics (BLS) releases data that show the rate of unemployment rising from 6.1 percent to 6.3 percent. What are the different ways a paper can report this number? One is a single sentence report that simply presents the above fact. But there are alternatives. Consider just two.

- (a) *Headline: Recession Fears Grow.* New data suggest the economy is slipping into a recession. The BLS reports that the number of unemployed grew by 200,000 in the last quarter, reaching 6.3 percent. John Kenneth Galbraith, the distinguished Harvard economist, sees this as an ominous sign of the failure of the administration's policies. "Not since Herbert Hoover has a president ignored economic realities so blatantly. This news is only the beginning of more to come," he said. (Accompanying picture: a

communications textbook (Werner Severin and James Tankard, Jr., 1992) all advance this view of news.

³ Entertainment broadcasting is analyzed by Peter Steiner (1952), Michael Spence and Bruce Owen (1977), Ronald Goettler and Ron Shachar (2001), and Esther Gal-Or and Anthony Dukes (2003). Jean Gabszewicz et al. (2001) take the approach closest to ours by conceptualizing news provision in a Hotelling framework. They examine how advertisers have an impact on content, whereas we focus on media accuracy.

⁴ For concreteness, we talk about newspapers, although our argument applies equally well to television and radio.

* Mullainathan: Department of Economics, 208 Littauer Center, Harvard University, Cambridge, MA 02138 (e-mail: mullain@fas.harvard.edu); Shleifer: Department of Economics, M9 Littauer Center, Harvard University, Cambridge, MA 02138 (e-mail: ashleifer@harvard.edu). We are extremely grateful to Alberto Alesina, Daniel Benjamin, Tim Besley, Filipe Campante, Gene D'Avolio, Glenn Ellison, Josh Fischman, Edward Glaeser, Matthew Gentzkow, Simon Johnson, Emir Kamenica, Lawrence Katz, David Laibson, Dominique Olie Lauga, Emily Oster, Richard Posner, Jesse Shapiro, Jeremy Stein, Lawrence Summers, and three anonymous referees for comments. This paper is a substantially revised version of an earlier paper entitled "Media Bias."

long line for unemployment benefits in Detroit, Michigan.)

- (b) *Headline: Turnaround in Sight.* Is the economy poised for an imminent turnaround? Data from the BLS suggest that it might be. Newly released figures show unemployment inching up just 0.2 percent last quarter. Abbie Joseph Cohen, the chief stock market strategist at Goldman Sachs, sees the news as highly encouraging. "This is a good time to increase exposure to stocks," she says, "both because of the strong underlying fundamentals and because the softness in the labor market bodes well for corporate profitability." (Accompanying picture: smiling Abbie Joseph Cohen.)

Each of these stories could easily have been written by a major U.S. newspaper. In fact, stories like these, in light of public disclosure of identical facts, are written every day. Neither story says anything false, yet they give radically different impressions. Each cites an authority, without acknowledging that a comparably respectable authority might have exactly the opposite interpretation of the news. Each omits some aspect of the data: the first by neglecting to mention the starting point of the unemployment rate, the second by ignoring unemployment levels. Each uses a headline, and a picture, to persuade readers who do not focus on the details. Each, in other words, slants the news by not telling the whole truth, but the articles are slanted in opposite directions.⁵

Our model of the market for news combines the assumption of readers preferring stories consistent with their beliefs with the assumption that newspapers can slant stories toward specific beliefs. We examine two crucial aspects of this environment. First, we consider two alternative assumptions about the nature of competition: monopoly versus duopoly. Our model of media competition is analogous to a Hotelling model of product placement (Jean Tirole, 1988, ch. 7). Newspapers locate themselves in the product space through their reporting strategies (i.e., how they slant). Readers' beliefs determine their "transportation" costs, since they face psy-

⁵ Persuasion can also work through outright fabrication of news, as was done routinely by the Communist press, and occasionally even in Western newspapers (e.g., Jason Blair's reporting for the *New York Times*.)

chic costs of reading papers whose reporting does not cater to their beliefs. We ask whether competition by itself eliminates or reduces the slanting of news, as economists often argue. We show that the answer for biased readers is clearly *no*. Competition generally reduces newspaper prices, but does not reduce, and might even exaggerate, media bias.

Second, we study heterogeneity of reader beliefs. What effect does such heterogeneity have on the nature of slanting and the overall accuracy in media? What is the impact of competition on media accuracy when reader beliefs are heterogeneous, as in the case of beliefs about President Clinton? To answer this question, it is crucial to distinguish between an average reader, who reads one source of news, and a hypothetical conscientious reader, who reads multiple sources. In general, competition with heterogeneous readers increases the slanting by individual media sources. But with heterogeneous readers, the biases of individual media sources tend to offset each other, so the beliefs of the conscientious reader become more accurate than they are with homogeneous readers. Our central finding is that reader heterogeneity plays a more important role for accuracy in media than does competition.

At a broader level, this paper contributes to one of the central issues in economics, namely whether the presence of rational, profit-maximizing firms eliminates any effect of irrational participants on market "efficiency." In the context of financial markets, Milton Friedman (1953) argued long ago that it does, and that rational arbitrageurs keep financial markets efficient. Subsequent research, however, has proved him wrong, both theoretically and empirically (Shleifer, 2000; Markus Brunnermeier and Stefan Nagel, 2004). One finding of this research is that, in some situations, such as stock market bubbles, it might pay profit-maximizing firms to pump up the tulips rather than eliminate irrationality (Brad DeLong et al., 1990). Subsequent research has considered the interaction between biased individuals and rational entrepreneurs in other contexts, such as the incitement of hatred (Glaeser, 2005), political competition (Kevin Murphy and Shleifer, 2004), and product design (Xavier Gabaix and Laibson, 2004). Here we ask a closely related question for the market for news: does competition among profit-maximizing news providers

eliminate media bias? We find that the answer, in both financial and political markets, is *no*. Powerful forces motivate news providers to slant and increase bias rather than clear up confusion. The crucial determinant of accuracy is not competition, *per se*, but consumer heterogeneity.

I. Model Setup

Readers are interested in some underlying variable t , such as the state of the economy, which is distributed $N(0, v_t)$. Let $p = 1/v_t$ denote the precision. Readers hold a belief about t that may be biased; beliefs are distributed $N(b, v_t)$. Thus, readers are potentially biased about the expected value of t , but have the correct variance.

Newspapers are in the business of reporting news about t . They receive some data $d = t + \varepsilon$, where $\varepsilon \sim N(0, v_\varepsilon)$. In the example from the introduction, these data might be an unemployment rate release. We assume that the papers then report the data with a slant s , so the reported news is $n = d + s$. For most of the paper, the exact technology of slanting is not important, but in Section V we study a specific one.

A. Reader Utility

Suppose readers are rational and unbiased. All they want is information. They dislike slanting because it is costly both in effort and the time it takes to read slanted news and figure out the "truth." In the BLS example, the report of the first newspaper does not tell the reader how much the unemployment rate changed, while that of the second newspaper does not contain the unemployment rate. To get a full picture, the reader needs more information. We assume that a rational reader's utility is decreasing in the amount of slanting. So, if he reads a newspaper, his utility is:

$$(1) \quad U_r = \bar{u} - \chi s^2 - P$$

where P is the paper's price. If he does not read the newspaper, he receives utility 0.

Biased readers, on the other hand, get disutility from reading news inconsistent with their beliefs. We model consistency as the distance between the news and the reader's beliefs, b , measured as $(n - b)^2$. In the BLS example, a

reader optimistic about the economy experiences disutility when reading stories that suggest a recession. At the same time, even biased readers dislike blatant and extreme slanting, at least in the long run. Holding constant the consistency with beliefs, they prefer less slanted news.⁶ So, if he reads the newspaper, the overall utility of a biased reader is:

$$(2) \quad U_b = \bar{u} - \chi s^2 - \phi(n - b)^2 - P$$

where $\phi > 0$ calibrates his preference for hearing confirming news.

B. Newspaper Strategy

Before seeing the data d , a newspaper announces its slanting strategy $s(d)$ and the price P it charges. Potential readers buy the paper if the price P is lower than the expected utility associated with reading the paper, $E_d[U(s(d))]$. To form expected utility, expectations are taken over d and are assumed to be the true expectations ($d \sim N(t, v_d)$) rather than the biased ones. This approach crudely captures the idea that this is a long-run game. Readers get a general sense of how much pleasure the paper provides them and make their purchasing decisions accordingly. It then makes more sense to think of expected utility using the empirical distributions. Practically, in the model both assumptions about expectations produce the same results.

Once readers decide whether to buy the paper, the paper observes its signal d and reports $n = d + s(d)$. Readers read the news and receive their utility. Timing of the full game is as follows:

- (a) The newspaper announces a strategy $s(d)$ for how to report the news. When there are two papers, both announce strategies simultaneously.
- (b) Price P is announced. When there are two papers, both announce prices simultaneously, after the other paper has revealed its strategy.

⁶ This assumption is immaterial to our results. All we require is that newspapers face some quadratic cost of slanting. This cost could just as easily arise on the supply side, with firms facing a technological or private reputational cost of slanting, and the results would be the same. The necessary feature is that firms cannot slant freely.

- (c) Individuals decide whether to buy the paper based on average utility associated with its strategy $s(d)$ and price P .
- (d) Newspaper receives data d and reports news $d + s(d)$. If there are two papers, they receive the same data d and report $d + s_j(d)$ where $j = 1, 2$.
- (e) If individuals buy the paper, they read the news and receive utility.

C. Cases Considered

We consider two different distributions of reader beliefs: homogeneous and heterogeneous. Homogeneity means that all readers hold the same beliefs b with precision p . For example, all or nearly all readers in the United States might believe that the Russians are corrupt or that the French are anti-American. Heterogeneity means that there is a distribution of reader beliefs. Such heterogeneity could come from political ideology. For example, opinions about U.S. presidents often divide along party lines. We assume that heterogeneous beliefs are distributed uniformly between b_1 and b_2 where $b_1 < b_2$ and $b_2 > 0$. Readers in this uniform distribution are indexed by $i \in [1, 2]$ so that reader i holds belief b_i . All readers hold their beliefs with precision p . We denote by \bar{b} the average of b_1 and b_2 . We also denote reader i 's utility function as $u_i(d)$ or $u_{b_i}(d)$, depending on context. The homogeneous and heterogeneous cases are designed to capture two different types of issues: ones on which there is consensus in the population and ones where there is substantial disagreement.

We also examine two cases of industry structure. In the first case, there is a single monopolistic newspaper. In the second, there are two newspapers, indexed by $j = 1, 2$, each seeing the same data d . For a monopolist, s_{hom}^* and s_{het}^* denote the optimal slanting strategy for the homogeneous and heterogeneous case. Similarly, P_{hom}^* and P_{het}^* denote optimal price in these cases. For duopolists, $s_{j,\text{hom}}^*$ and $s_{j,\text{het}}^*$ denote the optimal strategy of paper $j = 1, 2$ in the homogeneous and heterogeneous cases, respectively. Similarly, $P_{j,\text{hom}}^*$ and $P_{j,\text{het}}^*$ denote each duopolist's optimal price in these two cases.

This formalism of industry structure is similar in spirit to a Hotelling model. Readers' beliefs resemble consumers' preferred locations. Their dislike of inconsistent news resembles

transportation costs. Firms' choice of a slanting rule resembles their choice of location. In this context, our utility function implies quadratic transportation costs and our distribution of reader beliefs in the heterogeneous case corresponds to a uniform distribution of consumers. Consequently, many of our proofs resemble the proofs for the Hotelling models in this case (Claude d'Aspremont et al., 1979).⁷

D. Defining Bias

We are interested in the extent of newspaper bias in the market. We measure this by the average bias of the newspapers in the market, weighted by their market share. In the homogeneous case, where there is only one kind of reader, we simply define bias as

$$(3) \quad \text{ARB}_{\text{hom}} = E_d[(n - d)^2]$$

where n is the news read by these readers. So bias is defined as the average amount by which the news read deviates from the data for the average reader.

In the heterogeneous case, let n_i be the news read by reader $i \in [1, 2]$. Bias is then defined as:

$$(4) \quad \text{ARB}_{\text{het}} = \int_i E_d[(n_i - d)^2].$$

This measures the average bias that readers encounter.

II. Rational Readers

When readers are rational, newspapers face only a disincentive to slant. The following proposition summarizes the outcomes for different cases.

PROPOSITION 1: *Suppose readers are rational. Then, whether readers are homogeneous or*

⁷ As with all Hotelling models, the assumptions on transportation costs matter. With linear transportation costs, an equilibrium does not exist. But while the results depend on nonlinear transportation costs, they are not specific to the quadratic. Other convex functions produce similar results (Nicholas Economides, 1986). See Steffen Brenner (2001) for a survey. Similarly, as with all Hotelling models, the assumption of Bertrand competition is key to our results.

heterogeneous, the monopolist does not slant and charges the same price:

$$(5) \quad s_{\text{hom}}^* = s_{\text{het}}^* = 0$$

and

$$(6) \quad P_{\text{hom}}^* = P_{\text{het}}^* = \bar{u}.$$

In the duopolist case as well, papers do not slant and once again charge the same price:

$$(7) \quad s_{j,\text{hom}}^* = s_{j,\text{het}}^* = 0$$

and

$$(8) \quad P_{j,\text{hom}}^* = P_{j,\text{het}}^* = 0$$

for all j on the equilibrium path. The only effect of competition is to lower prices.

PROOF:

See Appendix for all proofs.

Proposition 1 illustrates the normal logic of economists' thinking about the media. When readers seek accuracy in news, newspapers pass on, without slant, the information they receive. Since perfect quality is achieved even without competition, the effect of competition is to reduce the price that readers pay. With both monopoly and duopoly, consumers get what they want and there is no media bias.⁸ In the rest of the paper, we focus on the case of biased readers.

III. Homogeneous Biased Readers

The following proposition summarizes the monopolist's behavior with homogeneous readers.

PROPOSITION 2: *A monopolist facing a homogeneous audience chooses:*

$$(9) \quad s_{\text{hom}}^*(d) = \frac{\phi}{\chi + \phi} (b - d)$$

$$(10) \quad P_{\text{hom}}^* = \bar{u} - \frac{\chi\phi}{\chi + \phi} [b^2 + v_d]$$

if $\bar{u} > [\chi\phi/(\chi + \phi)][b^2 + v_d]$. If not, there exists no slanting strategy that results in the news being read.

Because the monopolist can capture all surplus through the price he charges, to maximize profits he merely maximizes expected utility. The news he reports is:

$$(11) \quad n = \frac{\phi}{\chi + \phi} b + \frac{\chi}{\chi + \phi} d.$$

The reported news is a convex combination of bias and data, with weights given by utility parameters. In this case, we say the monopolist "slants toward b ." Since this linear slanting strategy will reappear throughout the paper, we define:

$$(12) \quad s_b(d) \equiv \frac{\phi}{\chi + \phi} (B - d).$$

With this notation, the proposition above can be rewritten as $s_{\text{hom}}^*(d) = s_b(d)$. The monopolist chooses this linear form because expected utility functions are separable in the value of d . The monopolist maximizes utility for every given value of d , which leads him to slant toward a biased reader's beliefs.⁹

The following corollary derives comparative statics for the magnitude of slanting.

COROLLARY 1: *In the homogeneous reader case, slanting increases with the reader preference for hearing confirmatory news and declines with the cost of slanting:*

$$(13) \quad \frac{\partial |s_{\text{hom}}^*(d)|}{\partial \phi} > 0$$

$$(14) \quad \frac{\partial |s_{\text{hom}}^*(d)|}{\partial \chi} < 0.$$

⁹ Even when $b = 0$, there is slanting. This is because even a reader who has zero bias ex ante does not want to change his mind ex post. Consequently, the monopolist slants news toward the reader's bias, 0.

⁸ As is clear from the proof of the proposition, this result generalizes trivially to $J > 2$ newspapers.

Proposition 2 suggests a theory of spin. Suppose that a politician, or some other figure of authority, has a first mover advantage, i.e., can choose which data d gets presented to the media first. The papers slant the data toward reader beliefs, but by Proposition 2, d will have significant influence on what papers report as compared to their getting data from an unbiased source. For example, by preemptively disclosing that a Chinese spy has been found in Los Alamos, a politician can focus the discussion on the risk to U.S. security from Chinese espionage, rather than on the administrative incompetence in the Department of Energy. This effect becomes even more powerful in a more general model of sequential reporting. In this case, the initial spin may shape reader priors, which future papers face and consequently slant news toward. The initial spin would then be reinforced even by ideologically neutral papers.

The condition $\bar{u} > [\chi\phi/(\chi + \phi)][b^2 + v_d]$ guarantees that this reader's reservation utility \bar{u} is high enough that he prefers reading the optimally biased news to no news. From now on, we assume that this condition holds.

ASSUMPTION 1: *Reader utility from news is high enough that readers prefer the equilibrium news to no news:*

$$(15) \quad \bar{u} > \frac{\chi\phi}{\chi + \phi} [b^2 + v_d].$$

With this assumption in place, we now turn to competition. How does competition between two newspapers affect the results above?

PROPOSITION 3: *Suppose duopolists face a homogeneous audience. Then there is an equilibrium in which duopolists choose on the equilibrium path:*

$$(16) \quad s_{j,\text{hom}}^*(d) = \frac{\phi}{\chi + \phi} (b - d)$$

and prices

$$(17) \quad P_{j,\text{hom}}^* = 0$$

for both $j = 1, 2$. Readers are indifferent between the two papers.

With a homogeneous audience, competition is Bertrand-like: it simply drives prices down to zero.¹⁰ Each duopolist's slant is exactly equal to the monopolist's slant, and they split the readers between them. The following corollary summarizes the impact of competition on bias in the homogeneous case.¹¹

COROLLARY 2: *For a homogeneous audience, both monopoly and duopoly produce the same amount of average reader bias:*

$$(18) \quad \text{ARB}_{\text{mon}}(v_d) = \text{ARB}_{\text{duo}}(v_d).$$

Propositions 2 and 3 are the first critical results of the paper. They show that when readers have homogeneous biases, competition does not eliminate them—it only leads to price reductions. Both monopolists and duopolists cater to reader prejudices. These propositions basically say that one cannot expect accuracy—even in the competitive media—on issues where the readers share beliefs. One example of such uniformity might be foreign affairs, where there may be a great deal of commonality of views toward a particular foreign country, such as Russia, China, or France. Another example is law enforcement, where most readers might sympathize with efforts by the government to prosecute members of a disliked group (e.g., the Arabs or the rich).

IV. Heterogeneous Biased Readers

What happens when readers differ in their beliefs? Newspapers must now decide which one of the heterogeneous reader groups is its target audience.

PROPOSITION 4: *Suppose a monopolist faces a heterogeneous audience with $\bar{b} = 0$. There exists a C_m , which depends on the parameters of the model, that determines the monopolist's strategy. If $b_2 - b_1 < C_m$, the monopolist maximizes profits by choosing:*

¹⁰ For this same reason, and as is clear from the proof of the proposition, this result holds for any number of newspapers $J \geq 2$.

¹¹ The stated equilibrium for the duopolists is not unique because any strategy profile that differs on a set of measure zero would also be an equilibrium.

$$(19) \quad s_{\text{het}}^* = s_b^*(d) = \frac{\phi}{\chi + \phi} (\bar{b} - d) = -\frac{\phi}{\chi + \phi} d$$

$$(20) \quad P_{\text{het}}^* = \bar{u} - \frac{\phi\chi}{\chi + \phi} v_d - \phi^2 b_2^2.$$

If $b_2 - b_1 > C_m$ the monopolist chooses not to cover the market, i.e., not all readers read the paper.

According to Proposition 4, the monopolist covers the market if the dispersion of reader beliefs is small enough. If beliefs are too far apart, readers on either extreme will not read the paper.¹²

Duopolists, in contrast, respond completely differently to heterogeneity. For tractability, we now consider only the situation where duopolists choose linear strategies.

PROPOSITION 5: *Suppose duopolists choose linear strategies of the form $s_b(d) = [\phi/(\chi + \phi)](B - d)$ and that $\bar{b} = 0$. Then there exists a constant*

$$(21) \quad C_d = \sqrt{\frac{4}{33} \left[\frac{\phi + \chi}{\phi^2} \bar{u} - \frac{\chi}{\phi} v_d \right]}$$

such that if $b_2 < C_d$ duopolists choose:

$$(22) \quad s_{1,\text{het}}^*(d) = \frac{\phi}{\chi + \phi} \left(\frac{3}{2} b_1 - d_1 \right)$$

$$(23) \quad s_{2,\text{het}}^*(d) = \frac{\phi}{\chi + \phi} \left(\frac{3}{2} b_2 - d_2 \right)$$

$$(24) \quad P_{j,\text{het}}^* = \frac{6\phi^2}{\chi + \phi} b_2^2$$

where we assume, without loss of generality,

¹² If $\bar{b} = 0$, but $b_2 - b_1 > C_m$, the monopolist would use the same slanting strategy as in Proposition 4, but would charge a high enough price that not all people read the paper. The case where $\bar{b} \neq 0$ is more complicated. The monopolist would not slant toward \bar{b} anymore. Instead, he would slant toward a point between \bar{b} and 0. This is because readers closer to the origin enjoy higher overall surplus from reading the paper (see Lemma (A1)). Consequently, the monopolist would prefer a distribution of readers closer to the origin so as to be able to charge higher prices.

that firm 1 slants toward the left and firm 2 slants toward the right. All readers read the newspaper.

Each duopolist positions himself as far away from the other as possible. The reported news in this case equals

$$(25) \quad n_j = d + s_{j,\text{het}}^*(d) = \frac{\phi}{\chi + \phi} \frac{3}{2} b_j + \frac{\chi}{\chi + \phi} d_j.$$

The reported news is a weighted average of the actual data d and $\frac{3}{2} b_j$, where b_j is the endpoint of the reader bias distribution. So duopolists are slanting news toward $\frac{3}{2} b_j$, points that are *more extreme* than the most extreme readers in the population.

This is analogous to the standard Hotelling result with uniform distributions and quadratic transportation costs (Tirole, 1988; d'Aspremont et al., 1979). As in the standard Hotelling model, the monopolist caters to both audiences unless they are too far apart, while duopolists maximally differentiate. But in the standard Hotelling model, firms are constrained to choose within the preference distribution. In our model, they can choose positions outside the distribution of reader bias, and in equilibrium choose very extreme positions.¹³

To see why this occurs, consider a simple case where $\phi = 1$, $\chi = 1$, $b_2 = 1$ and $b_1 = -1$. With these parameters, suppose the firms locate at $z_1 \leq z_2$.¹⁴ Equilibrium prices then equal (see the proof of Proposition 5):

$$(26) \quad P_1^*(z_1, z_2) = \Delta z \left(1 + \frac{\bar{z}}{3} \right)$$

$$(27) \quad P_2^*(z_1, z_2) = \Delta z \left(1 - \frac{\bar{z}}{3} \right)$$

where $\Delta z = z_2 - z_1$ and $\bar{z} = (z_1 + z_2)/2$. The more differentiated the duopolists (the greater is

¹³ If $\bar{b} = 0$ but $b_2 > C_d$, the duopolists differentiate less than stated in Proposition 5. The participation constraint of the reader with bias 0 begins to bind and the duopolists locate closer together than in the proposition. If b_2 is sufficiently large, the duopolists would even end up inside the distribution of reader beliefs so that $|z_j| < |b_j|$.

¹⁴ Recall that "located at z " means the paper biases according to the rule $s_z(d) = [\phi/(\chi + \phi)](z - d)$.

Δz), the higher the prices they can charge. Differentiation softens price competition because the temptation to undercut each other diminishes as the firms move farther away from the marginal consumer (who is located between them).

Now consider firm 1's choice of where to locate. When biasing toward z_1 , firm 1 captures all readers between -1 and $x^*(z_1, z_2) = \bar{z}/3$. Hence its profits equal $P_1^*(1 + \bar{z}/3)$. Differentiating with respect to z_1 gives the first-order condition

$$(28) \quad \frac{\partial P_1^*}{\partial z_1}(x^*(z_1, z_2)) + P_1^*\left(\frac{\partial x^*}{\partial z_1}\right) = 0$$

$$(29) \quad \frac{\partial P_1^*}{\partial z_1}\left(1 + \frac{\bar{z}}{3}\right) + P_1^*\left(\frac{1}{3}\right) = 0.$$

Increasing z_1 (that is, moving closer to the origin) has two effects on profits. The first is a price effect; there is a change in profits because changing position affects the equilibrium prices. The second is a market share effect; there is a change in profits because moving closer to the origin raises market share.

Papers slant toward positions well beyond the extreme consumers because the price effect dominates the market share effect until firms are very far apart. Focusing on the symmetric case with $\bar{z} = 0$, the price effect is $\partial P_1^*/\partial z_1 = \Delta z/6 - 1$. The price effect is negative as long as $\Delta z < 6$, in other words, until the difference in firm locations is three times as high as the difference in most extreme readers ($3(b_2 - b_1) = 6$). The market share effect, on the other hand, is $P_1^*/3 = \Delta z/6$. These two effects offset each other to produce an optimum when $\Delta z/6 - 1 + \Delta z/6 = 0$ or $\Delta z = 3$. At the symmetric equilibrium, the optimum is reached at $\Delta z = -2z_1 = 3$ or $z_1 = -3/2$. The distance between the newspapers ($z_2 - z_1 = 3$) is greater than the distance between the most extreme readers ($b_2 - b_1 = 2$).

In short, when choosing how to slant, duopolists maximally differentiate themselves.¹⁵ Practically, this means that news

¹⁵ This analysis also illustrates why Proposition 5 is about competition, per se, and not about variety alone. A monopolist who could start two newspapers does not need to differentiate to increase market power. He would differentiate simply to cater to reader tastes, but would not go beyond the most extreme readers as duopolists would.

sources can be even more extreme than their most biased readers. One cannot, therefore, infer reader beliefs directly from media bias.

Another point is worth noting:

$$(30) \quad E[(s_{j,\text{het}}^*(d))] \geq E[(s_{\text{het}}^*(d))].$$

Duopolists always slant more than the monopolist when readers are heterogeneous. In this sense, competition tends to polarize the news. The following corollary summarizes the impact of competition on bias.

COROLLARY 3: *Suppose $b_1 - b_2 < C_m$. In the heterogeneous reader case, competition increases the bias of the average reader:*

$$(31) \quad \text{ARB}_{\text{mon,het}}(v_d) < \text{ARB}_{\text{duo,het}}(v_d).$$

Corollary 3 shows that, with heterogeneous readers, competition by itself polarizes readership and, if anything, raises the average reader bias. Entry of a left-wing newspaper or a TV station into a local market previously dominated by a moderate or slightly right-wing monopolist might cause this monopolist to shift his reporting to the right.

Corollary 3 might shed light on the growing controversy in the United States about media bias. Several recent books have angrily attacked media outlets for having a left-wing bias (e.g., Goldberg, 2002; Coulter, 2003). Several equally angry books have responded that other media outlets have an even stronger right-wing bias (Alterman, 2003; Franken, 2003). We suspect that there is a grain of truth in all these books, and that the growing partisanship of alternative media sources is a response to the growth in competition, and market segmentation, in the media. Changes in media technology have led to significant entry, especially in television. If these media sources divide the market along ideological lines, we expect them to become more biased than they were in the regime of moderate competition. This is perhaps what the various commentators are recognizing.

Corollary 3 may also have implications for the effects of entry of new media outlets on the nature of reporting. In a provocative recent study, Gentzkow and Shapiro (2004) examine the responses to a Gallup poll by residents of nine Muslim countries about such topics as the

United States, terrorism, responsibility for 9/11, and so on. The authors document a striking pattern of factually inaccurate beliefs, but also suggest that the media have a strong effect on these beliefs. In particular, those who watch al-Jazeera (Arab television) are much more likely to hold factually false beliefs (as well as anti-American ones) than those watching CNN.¹⁶ In concluding their paper, Gentzkow and Shapiro appear to endorse recent proposals favoring an expansion of Western news in the Arab world, because such news is likely to moderate opinions and beliefs.

Our model suggests that caution is appropriate. The people who watch or listen to Western news are already sympathetic to its perspective and might already watch CNN, so they are unlikely to be strongly affected. Additional entry might cause al-Jazeera and similar networks to further differentiate their product by advancing yet more extreme views. The effect might be to radicalize, rather than moderate, their audience.

V. Reader Heterogeneity and Accuracy in Media

Our results so far focus on how an average reader in the population is affected. We can also look at the impact of reporting on a *conscientious* reader, a hypothetical reader who reads all the news available but is too small to affect what is reported. The interesting insights arise in the duopoly case where the hypothetical conscientious reader reads both papers. Since both papers are reporting on the same event, the conscientious reader might in principle be able to use the two to undo the slanting. To understand this process we need a precise model of slanting.

¹⁶ These results are not unique to the Muslim world. Steven Kull et al. (2003) document significant confusion among large percentages of U.S. respondents on such questions as Saddam Hussein's culpability in 9/11 and the discovery of weapons of mass destruction in Iraq. The study also finds that those who get their news from Fox News are less well informed about these issues than those who get their news from PBS and NPR.

A. Technology of Slanting

Following Hayakawa (1940), we assume that newspapers slant by selectively omitting specific bits of news, i.e., not reporting the whole truth.¹⁷ To formalize this idea, suppose that, rather than simply receiving a composite $d = t + \varepsilon$, the newspaper receives a sequence of positive and negative "bits" or facts. In the example from the introduction, these facts could be the unemployment rate, the unemployment rate in the past, expert opinions, other relevant economic indicators, and so on. These bits or facts are modeled as a length L string f consisting of positive (+1), negative (-1), or nonexistent (\emptyset) pieces of news. At each position, the probability of each of these values is a function of d , so now instead of simply seeing the composite d , the paper sees all the bits of facts that constitute it. The probability that the piece of news in position i , denoted f_i , is positive, negative, or nonexistent is given by the distribution function:

$$(32) \quad \Pr(f_i) = \begin{cases} +1 = qg(d) \\ -1 = q(1 - g(d)) \\ \emptyset = (1 - q) \end{cases}$$

where $g(\cdot)$ is a continuous and increasing function that is bounded between 0 and 1, and $0 < q \leq 1$. With probability $1 - q$, there is no news at position i . If there is news, it is positive with probability $g(d)$ and negative otherwise. Conditional on d , these probabilities are iid across different bits on a string. With multiple papers, we assume that they all see the same string f .

A newspaper that does not slant at all would simply report the string f without alteration. A reader who sees the string f can draw inferences from the number of +1's and -1's, which we define as $N_+(f)$ and $N_-(f)$, respectively. By the Law of Large Numbers:

$$(33) \quad \frac{N_+(f)}{N_-(f) + N_+(f)} = g(d) + \eta \rightarrow g(d)$$

where η is a noise term that converges to zero as the length of the string $L \rightarrow \infty$. Consequently, for large L , the information the reader receives

¹⁷ Importantly, newspapers do not slant by simply manufacturing evidence.

is well approximated by the case in which he simply observes d since $g^{-1}[N_+(f)/(N_-(f) + N_+(f))] \rightarrow d$.

In this formalism, a newspaper slants the signal by selectively omitting positive or negative bits of information. To slant upward, for example, a newspaper drops negative bits. Instead of reporting $+1, -1, -1, \emptyset, +1, -1, \dots$ it reports $+1, \emptyset, \emptyset, \emptyset, +1, -1, \dots$, for example. A paper that wishes to slant upward by $s > 0$ produces a string f' by dropping enough negative bits to guarantee

$$(34) \quad g^{-1}\left(\frac{N_+(f')}{N_-(f') + N_+(f')}\right) \approx d + s.$$

Likewise, a paper that wishes to slant negatively by $s < 0$ simply drops enough positive bits. As $L \rightarrow \infty$, the paper can choose to drop bits to approximate better and better any given slant s .

For simplicity, assume that newspapers omit facts in fixed ways. To slant positively, a paper omits the lowest indexed negative bits until it approximates the desired fraction. To slant negatively, a paper omits the lowest indexed positive bits until it reaches the desired fraction. This assumption is simply one way of formalizing the idea that two papers wishing to slant in a particular direction do so similarly.

B. Cross-Checking

By cross-checking the facts in the two newspapers, a conscientious reader may be able to reduce the effect of slanting. Suppose each paper receives string f , which can be thought of as implying data $d = t + \varepsilon$, and paper j reports string f_j . There are now several cases. If the implied slants for both papers are positive and $s_1 > s_2 > 0$, then every fact that paper 1 reports, paper 2 also reports. Moreover, because paper 2 is slanting less, it reports some facts that paper 1 does not. Consequently, a conscientious reader would interpret the news as if she had read only paper 2. The case where $0 > s_2 > s_1$ is similar. On the other hand, if the two papers are on opposite sides of the issue so that $s_1 > 0 > s_2$, paper 1 omits some negative details to slant upward and paper 2 omits some positive details to slant downward. The conscientious reader, however, can cross-check both papers. Paper 1 reports the positive facts, which paper 2 omits, and paper 2 reports the negative facts,

which paper 1 omits. By cross-checking, the conscientious reader gets all the facts, as if she were able to read an unslanted newspaper. Define $xc(\cdot)$ to be the cross-checking function:

$$(35) \quad xc(s_1, s_2) = \begin{cases} \min\{s_1, s_2\} & \text{if } s_1 > 0, s_2 > 0 \\ \max\{s_1, s_2\} & \text{if } s_1 < 0, s_2 < 0 \\ 0 & \text{otherwise.} \end{cases}$$

This function summarizes how the conscientious reader can cross-check the two papers.¹⁸

Define n_c to be the news the conscientious reader is effectively exposed to:

$$(36) \quad n_c = \begin{cases} n & \text{if one newspaper} \\ d + xc(s_1, s_2) & \text{if two newspapers.} \end{cases}$$

We then define conscientious reader bias analogously to the average reader bias:

$$(37) \quad CRB = E_d[(n_c - d)^2].$$

This definition of conscientious reader bias is independent of heterogeneity of reader beliefs. However, CRB does depend on the equilibrium news reporting, which in turn may depend on the heterogeneity of reader beliefs.

As the discussion on cross-checking suggests, reader heterogeneity can help the conscientious reader quite a bit. To formalize this, let us compare the case of homogeneous readers with bias b to the case of heterogeneous readers with beliefs distributed uniformly on $[b - \delta, b + \delta]$. The following corollary summarizes our principal finding:

COROLLARY 4: *The interaction of reader heterogeneity and duopoly lowers conscientious reader bias. When readers are heterogeneous, conscientious reader bias is lower under duopoly than monopoly:*

$$(38) \quad CRB_{\text{het,duo}} < CRB_{\text{het,mon}}.$$

¹⁸ The extreme cross-checking depends on the two papers slanting stories using the same rule. It is necessary for our results only that the papers use similar rules. Suppose that when one paper omits a fact, it appears in an oppositely slanted paper only with probability z . In this case, the cross-checking function becomes $(1 - z)s_1 + (1 - z)s_2 + zxc(s_1, s_2)$. Thus, the qualitative statements we make are preserved.

Under duopoly, conscientious reader bias is lower under heterogeneity than homogeneity:

$$(39) \quad CRB_{\text{het,duo}} < CRB_{\text{hom,duo}}.$$

Corollary 4 is the final result of our paper and its bottom line. It points to the absolutely central role that heterogeneity of reader beliefs plays in assuring accuracy in media. We have shown that when readers are homogeneous, competition results in lower prices, but not in accurate news reporting. When readers are heterogeneous, the news received by the average reader might become even more biased as competitive media outlets segment the market. Such market segmentation, however, benefits a conscientious reader, who can then aggregate the news from different sources to synthesize a more accurate picture of reality. When newspapers are at different sides of the political spectrum, the conscientious reader gets all the facts. While individual news sources slant even more when faced with a heterogeneous public, the aggregate picture becomes more clear. In this respect, reader heterogeneity is the crucial antidote to media bias.

This analysis indicates which issues are more likely to receive accurate media coverage, at least for the conscientious reader. Almost surely, the most likely domain of reader heterogeneity is domestic politics, where readers have diverse beliefs and media coverage is correspondingly diverse. Such dispersion of reader beliefs could come from their self-interested economic and social preferences, what used to be called "class differences." But, as Glaeser (2005) argues, such differences are reinforced by political entrepreneurs, who have an incentive to create particular beliefs that would bring them support, especially if these beliefs distinguish them from the incumbent. Newspapers would then follow these entrepreneurs in mirroring and reinforcing the beliefs of their supporters. In fact, in many countries today, and in the United States 100 years ago, newspapers were affiliated with political parties (Hamilton, 2003). Reader diversity, and newspaper diversity, are partly a reflection of underlying political competition. In other areas of competition, such as sports, we likewise expect local papers to support local teams, thereby creating diversity of reporting across cities reflecting the diversity of reader beliefs.

Perhaps the clearest illustration of this corollary is the coverage of the Monica Lewinsky affair during the Clinton presidency. The left-wing press presented an enormous amount of information designed to expiate the president's sins, while the right-wing press dug out as many details pointing to his culpability. In the end, however, as Posner (1999) remarks in his book, much of the truth has come out and a conscientious reader could get a fairly complete picture of reality.

VI. Conclusion

We have examined the roles of two forces in promoting accuracy in media: competition and reader diversity. We have found that competition by itself is not a powerful force toward accuracy. Competition forces newspapers to cater to the prejudices of their readers, and greater competition typically results in more aggressive catering to such prejudices as competitors strive to divide the market. On the other hand, we found that reader diversity is a powerful force toward accuracy, as long as accuracy is interpreted as some aggregate measure of revelation of information to a reader who takes in all the news. Greater partisanship and bias of individual media outlets may result in a more accurate picture being presented to a conscientious reader.

Reader heterogeneity comes in part from underlying political competition, whereby political parties, movements, and individual entrepreneurs attempt to generate support by presenting their points of view. If they can generate enough interest, media outlets will try to cater to the very same audiences that the political entrepreneurs attract, and diversity in media coverage will arise endogenously. In contrast, when potential audiences share similar beliefs, and when there is no advantage from political entry, such as the coverage of foreign countries or crime, we do not expect to see diversity of media reports or accuracy in media.

Political competition is only one source of underlying reader diversity. We can also imagine entrepreneurs starting newspapers on their own and, as long as they have deep enough pockets, creating enough demand for unorthodox views to broaden the range of opinions (and slants) that are being covered. Ideological diversity of entrepreneurs themselves may be the source of diversity of media coverage.

We have studied competitive persuasion in the market for news. Our principal finding is that, when competitors can create or reinforce differences of opinion, they will do so in order to divide the market and reap higher profits. There will be no convergence in reporting to the median reader (as in a Downsian median voter framework). We believe that this consequence of competitive per-

suasion is more general, and that attempts to differentiate competitively by moving toward extreme positions will arise in both political (Murphy and Shleifer, 2004) and product (Gabaix and Laibson, 2004) markets. In these and other domains, the influence of audience heterogeneity and competition on the content of persuasive messages remains to be fully explored.



Magazine | Video | LIFE | Person of the Year

Ideas

QUICK AND EASY GARDENING
The No-Fuss Guide to a Beautiful, Stylish Garden



BUY NOW

BUSINESS & TECH

Confidence Woman

By Belinda Luscombe | March 07, 2013 | 37 Comments

Sheryl Sandberg's first employees, according to her family, were her siblings David and Michelle. "Initially, as a 1-year-old and 3-year-old, we were worthless and weak," they said in a toast at her wedding. But by elementary school the person who is currently the chief operating officer of Facebook, and arguably one of the most powerful women in America, had whipped them into shape, teaching them to follow her around the house and shout "Right!" after each of her orations. Was this a game? Sort of. "To the best of our knowledge Sheryl never actually played as a child," they said. "[She] really just organized other children's play."

Sandberg tells these stories about herself early in her first book, a memoir-slash-sort of feminist manifesto in which she enjoins women to pursue their careers with more rigor, to engage more energetically in the corporate cook-off, to Lean In—as the book is titled—to the opportunities and challenges of becoming a boss. She says she had misgivings about sharing these family fables because they make her seem bossy, a term she takes issue with. "I notice bossy is applied almost always to little girls," says Sandberg over lunch (she ordered a Wagyu hot dog with no bun and no relish). "It's just not used for men."

In person, Sandberg does not give the impression that she's bossy. She gives the impression that she was born 43, that she was delivered preloaded with the capacity and will to order people around but also the capacity and will to ensure that they thrive. Now that she is really 43, she has so perfected these skills that merely helping run a \$66 billion tech company is not quite enough of a challenge. So Sandberg has taken on a new mission: to change the balance of power. That quest and her plan of attack have brought out the broadsides.

(MORE: [TIME's Complete Coverage on Sheryl Sandberg](#))

It would be un-Sandbergian to write a book and just leave it at that. Her campaign comes with [LeanIn.org](#), a nonprofit foundation with corporate partnerships, online seminars and guidelines for establishing support groups. It's probably not an overstatement to say Sandberg is embarking on the most ambitious mission to reboot feminism and reframe discussions of gender since the launch of Ms. magazine in 1971.

The thing is, she's in a pretty good position to pull it off. She's the co-pilot of the biggest network of humans the world has ever seen: Facebook's roughly 1 billion members, most of whom are female, at least in the U.S. She's worth hundreds of millions of dollars. And she has an undeniable record of knowing how to get things done. Her résumé, with its by-the-book stints at Harvard Business School, McKinsey and the Treasury Department, does not reek of revolutionary, but in the lineage of key feminist figures, she may well turn out to be pivotal. "In a sense it's almost like Betty Friedan 50 years ago," says author and historian Stephanie Coontz. "She's talking to a particular audience, but they really need this message."

(MORE: [Sandberg Exclusive Excerpt: 'Why I Want Women to Lean In'](#))

Midflight Stall

Why, almost exactly 44 years after Lorena Weeks became the first woman to use the Civil Rights Act to win the right to be promoted, at Southern Bell, are we still arguing about women and success? Only flat-earthers and small boys don't believe that women can lead huge Western democracies (thanks, Margaret Thatcher), head companies (thanks, Indra Nooyi), play exciting sports (thanks, Billie Jean King), rise to the rank of four-star general (thanks, Ann Dunwoody), change the world, trade cattle futures and be funny (thanks for all three, [Hillary Clinton](#)).

But women's journey to the top is having an altitude problem. Young female executives begin on the same career staircase as men, but it's almost as if the stairs change direction, Hogwarts-like, and take them somewhere else. For three decades, more women than men have graduated from college, but



PETER BOHLER FOR TIME

Strategy meeting Sandberg with LeanIn.org team members

RELATED

[Chore Wars](#)

[What Betty Friedan Saw Coming](#)

[View point: Will Family Issues Finally Get Addressed?](#)



that academic dominance has not led to corresponding business or political success. There are currently only 17 heads of state out of 195 who do not have a Y chromosome. Women hold about 20% of all seats in parliaments globally. Slightly more than 4% of Fortune 500 companies are headed by women, and women hold 17% of board seats. Worse, these numbers aren't changing very fast. Ten years ago, 14% of board seats were held by women. A decade has passed, and women have gotten a few inches farther into the boardroom. "Women are not making it to the top of any profession in the world," says Sandberg. "But when I say the blunt truth is that men run the world, people say, 'Really?' That, to me, is the problem."

Few people disagree that somewhere on the climb between the graduation podium and the C-suite, women are getting lost. The contentious issue is what—or who—is keeping them down. Fingers are pointed in every direction: the U.S. has primitive maternity-leave laws on par with those of Papua New Guinea—a country that still has actual cannibals. Women are hitting their childbearing deadlines around the time future executives are being winnowed out from regular management. Turnover at corporate boards, which are heavily male-dominated, is very slow; most have a mandatory retirement age of 72. American companies structure their workers' days around the expectation that someone else is handling the home front. Men have welcomed women into the workplace, but housework, cooking and child-rearing duties are still borne largely by women.

(MORE: [Forget About Mentors — Women Need Sponsors](#))

Sandberg acknowledges all these obstacles but drills down on one in particular, the one she says receives the least attention: the invisible barrier in women's minds. "Compared to our male colleagues, fewer of us aspire to senior positions," she writes.

It's not exactly that they're to blame, she notes. Females are raised from birth to have

different expectations. There's an ambition gap, and it's wreaking havoc on women's ability to advance. "My argument is that getting rid of these internal barriers is critical to gaining power. We can dismantle the hurdles in ourselves today. We can start this very moment."

Do women want that kind of power? Are men hardwired to want the big paycheck, the high-horsepower career more? How much of women's tendency to lean back stems from something deep in the DNA? Research findings suggest that women are as ambitious as men but that their ambition expresses itself in a different way. For Sandberg, these are not relevant issues, just as it's unclear whether humans are genetically predisposed to eat too much or do so because of the food around them. Either way, it's causing obesity and needs to change. "We have to evolve to meet new circumstances," she says. "I'd like to see where boys and girls end up if they get equal encouragement—I think we might have some differences in how leadership is done."

Sandberg's critics are quick to remark, Easy for you to say. She has two Harvard degrees, a rich but menschy CEO husband, vast personal wealth, all the household help she needs and a flexible workplace. She walked into two of the right companies—Google and Facebook—at the right time. Women lower on the scale of money and education may wonder just how Sandberg expects them to lean in to their paycheck jobs. And for her to suggest that other women aren't doing the right things to be successful, well, it's what many people are calling ballsy, as in that's what a guy would say. Her thesis has already drawn the ire of other women working in the same field. (Men have been less voluble. This is no-win territory for them.)

(MORE: [Dominique Browning: More Ways Women Sabotage Themselves](#))

"Are we going to spend another 20 years trying to make women adapt to a system that doesn't fit them?" asks Avivah Wittenberg-Cox, who runs a global management consultancy, 20-first, that helps companies achieve greater gender balance. She points to data from McKinsey that businesses with more women on their boards are more profitable. "Companies need women. It's a problem for them if women aren't advancing." She thinks Sandberg's message is the wrong one. "It's insulting to women to say they need to become more like men to succeed."

To be fair, that's not exactly what Sandberg is saying. For all her success, she's nothing like a man. She may currently have thousands of people saying "Right!" to her, but she's refined her technique since elementary school. Now it blends an overwhelming amount of data with a weapons-grade ability to nurture and an exquisite organizational acumen. She's like an escapee from a Star Trek episode in which Spock sired a child with an empath.

Take her role at Facebook. COOs aren't usually the rock stars in an organization. They're the nuts-and-bolts guys—usually guys—executing the CEO's will and hoping to get the top job. Sandberg's approach has been a little different.

"She built the whole business part of Facebook," says Mark Zuckerberg, the social juggernaut's hoodie-wearing CEO. "I didn't know anything about running a company. [We] knew where we wanted to get, but we were lacking someone who was a visionary at how you work at large scale." The company had about 70 million users and \$150 million in revenue before she joined in 2008. Now it has a billion users and recently reported revenues of \$1.59 billion for the quarter. "Some people emanate 'I'm a pro at what I do. And I'm such a pro that when you're around me, you're going to want to be more of a pro too,'" says Chris Cox, Facebook's implausibly young, handsome and Zen VP of product. "And that's how it felt when she showed up."

Nobody at Facebook has an office. Sandberg sits two desks down from Zuckerberg in a corner of one of the social network's parking garage-size open-plan buildings in Menlo Park, Calif. Next to her is a pillar with "I Love You, Mom" painted in childish letters, created during a visit to Mom's workplace by her 7-year-old son and 5-year-old daughter. Opposite her sits her longtime assistant Camille Hart, who works on the multicolored megascreen spreadsheet that is her boss's schedule. When Sandberg wants to talk to Zuckerberg, which is often, she spins around on her chair and literally leans in.

Passionate even for Facebook, where messianic is the default attitude, Sandberg's a huge fan of the word huge. As in, "That is huge." "It's a huge problem." "This is hugely important." Her second favorite word seems to be genuinely, although to be fair, she's partial to all adverbs. She gestures continually, with her fingers bent at the second knuckle, as if she's mixing pizza dough or winding yarn. She's an ardent listmaker and is never without a little notebook. Each page is either a project or a person, and she rips them out when the tasks are done. "I feel my to-do list," she says.

(MORE: [Caitlin Flanagan: What About the Children?](#))

Combined with her efficiency is her emotional quotient (EQ), an uncanny grasp of how people feel. In a meeting to discuss the purchase of a Web-design company—a process known as acqui-hiring, in which the deal is mainly aimed at bringing in new talent—Sandberg reminds her team that the firm's founder is about to have a birthday and wants to get the deal done before the big day. "I think that birthday helps us," she says. As Zuckerberg puts it, "She's unique in that she has an extremely high IQ and EQ, and it's just really rare to get that in any single person."

Sandberg doesn't like to call what she does management. It seems too clinical. She has the gift of making others feel their contribution is significant. (Two people told me they were the first to take Sandberg's kids to a farm.) She believes in crying in the office and devotes a chapter in her book to honest communication at work. "We argue pretty vehemently," says Cox. "One thing I appreciate about Sheryl—when it's about to get heated, we'll call each other. We don't raise our voices. We have a different tone."

Meetings are the vertebrae of any executive's day, and Sandberg runs a brisk one. In the pre-Sandberg era, they didn't always start on time. And there weren't always notes. "Sheryl's able to get a diverse set of people to get to a decisive position very quickly," says Mike Schroepfer, VP of engineering. "She's famously impatient." She's also practical, making sure people aren't meeting on an empty stomach. "After Sheryl came to Facebook, I got a lot less hungry," recalls Zuckerberg.

The Sandberg Way

After running thousands of meetings and hiring, directly or indirectly, thousands of people, Sandberg feels she's in a position to comment about the way women work. And here's what she's noticed: it's not their fault exactly, but they aren't pursuing their careers in the most efficient way. Inefficiency is abhorrent to Sandberg. She has a sign in her conference room that reads, "Ruthlessly Prioritize".

Of course, we can't all be Sheryl Sandberg. In fact, none of us can be Sheryl Sandberg. To understand why, it helps to know how she got to be who she is.

"I was raised [to believe] that going into business was a bad thing," says the oldest daughter of Joel and Adele Sandberg, an ophthalmologist and teacher from Florida. "You were supposed to be a doctor or work for the government or a nonprofit." (Both her siblings went into medicine.) Sandberg thought she was going to be a lawyer. In sixth grade she took second place in a Florida-wide oratory contest, even though all the other speakers were in high school. That she couldn't see over the lectern without a step stool didn't diminish the impact of her speech about the folktale of the little red hen and the importance of everyone's doing their bit for America.

It was Sandberg's parents who first demonstrated the power of the network. Joel is the efficient, competitive one, Adele the passionate, nurturing one. In the '70s they were activists for Soviet Jews who were trying to emigrate to Israel. If one of the refuseniks, as they were known, was arrested or sent to a labor camp, the community reached out to a guy in London. He then called a bunch of supporters all over the world, including Adele Sandberg, and they activated a telegram program and called their local politicians. By 1987, partly as a result of pressure from Western nations and networks like the Sandbergs', Jews were allowed to leave the Soviet Union. "My biggest concern for my kids was that they grow up to be a mensch," says Adele. "If she ended up turning into a snob, I would not be proud of her."

After topping her public school and getting her undergraduate degree at Harvard, Sandberg was accepted into its law school. Despite the thriving aerobics class she'd started on campus—where she says she learned to smile even when she didn't mean it—she went to work for the World Bank for her former professor Larry Summers. He had been her thesis adviser (she wrote about the economics of spousal abuse) and says, "I noticed her because she was the best student out of 75 or 80 in my undergraduate class." After two years of international aid work, partly on Summers' advice, Sandberg decided to skip law and do an M.B.A.

(MORE: [The Pay Gap Is Not as Bad as You \(and Sheryl Sandberg\) Think](#))

Success vs. Popularity

Sandberg learned one of the key lessons in her book during her time in business school (Harvard, again), and not in a good way. After her first year, she won a fellowship. The guys who won all talked about it. But Sandberg sensed it was better to keep quiet. "Female accomplishments," she writes, "come at a cost." And that cost is people declining to click on the Like button.

Sandberg often refers to a 2003 experiment that found that students considered a successful entrepreneur in a case study more likable when her name

was changed to a man's. "The data says clearly, clearly, clearly that success and likability are positively correlated for men and negatively correlated for women," says Sandberg. Finding that out "was the aha moment of my life." It explains, she believes, why women who will negotiate ruthless deals for their clients will not do the same for themselves. It accounts for why women are less eager than men to trumpet their management triumphs or put themselves forward for leadership positions. Because women are supposed to be nurturing and peacemaking, not aggressive. When she clues in managers on the success-and-likability conundrum, "it completely changes the way they review women," she says.

Awkwardly, it turns out, women don't particularly like successful women either. Sandberg points to how quickly people criticized her friend Marissa Mayer, the CEO of Yahoo, who went back to work two weeks after having a child and recently appeared to make Yahoo's work practices a lot less flexible. "No one knows what happened there," she says. "I think flexibility is important for women and for men. But there are some jobs that are superflexible and some that aren't." Regardless, she believes no man who ordered the same policies would have come under fire the way Mayer has.

(MORE: [Judith Warner: Why Sandberg Matters for Real Women](#))

Sandberg, too, has drawn her share of opprobrium. After Anne-Marie Slaughter, an academic and former State Department honcho, criticized her in a much-talked-about essay on why women can't have it all, Sandberg sent her an e-mail, which Slaughter talked about to a newspaper. Sandberg, the reigning world champion in finding a positive thing to say about everyone, initially declined to comment on this episode. The two have now made up.

At least one prominent feminist is supportive. "Every group of people that has been systematically told they were supposed to play a limited role internalizes that role," says Gloria Steinem. "She's saying we have to both fight against the barriers and get them out of our consciousness."

Sandberg's peers are generally supportive but guarded. "The most crucial thing for a woman to have if she's going to get to the top is a sponsor," says Ann Lee, author of *What the U.S. Can Learn from China* and a contemporary of Sandberg's at Harvard Business School. "I was not terribly surprised at Sheryl's success, because I knew Larry Summers had taken her under his wing." In fact, after a short stint at McKinsey in 1996, Sandberg went to work with Summers again, this time at the Treasury Department. When he became the Treasury Secretary, she was his 29-year-old chief of staff. "I was hugely lucky, and that explains most of my success," says Sandberg, "just like every man."

Her next move, to a small but energetic company called Google in 2001, took people more by surprise. Wayne Rosing, who now runs an astrophysics nonprofit, was vice president of engineering at the time and one of the people who interviewed Sandberg for the job. "She was such a Google type: smart, articulate, passionate and able to work through a problem she had never encountered before," he says. What Rosing didn't notice, however, was her passion for women's rights: "She was just one of the guys, one of the players." In fact, it was only after she got very sick while pregnant (the Sandberg women all had nine months of morning sickness) that she got the firm to put in special parking spots for expectant moms.

"I never called myself a feminist or gave speeches on women as late as five years ago," says Sandberg, whose interest in women's leadership coincided with her joining Facebook in 2008. Until the week before *Lean In* came out, she was the only woman on Facebook's board and had been there less than a year, and she's still the only woman among its top executives. Since the day she took Facebook public in a much hyped IPO, the stock has yet to rise above its offering price; investors are skittish, and advertisers are skeptical. The company needs a steady presence and a cohesive face as it moves forward. This might explain why Sandberg's nearly omnipresent Facebook handlers are quick to insist that *Lean In* is not a company project or a distraction and is definitely not Sandberg's exit strategy. The only time Zuckerberg looked at one of the two p.r. reps present during our interview was when he was asked how irreplaceable she was. He finally came up with: "She has irreplaceable qualities."

Other employees are less cautious. "I have not thought about Facebook without Sheryl," says Cox. "That would suck." He'd respond, he says, by trying to get as good at writing noncheesy thank-you notes as Sandberg is. "If Sheryl were to leave, a bunch of us would say I need to absorb that and honor that," he says. These people take their social networking seriously.

(VIDEO: [Sheryl Sandberg Leans Forward](#))

How She Does It

Among the myths that circle around Sandberg is that she leaves the office at 5:30 p.m. Actually, that is true. But after putting in some time with her family, she returns to work with a vengeance. She's one of those work-hard, play-hardly-ever types. She usually goes to parties only to work the room or if she's holding a gathering of women at her home. She and her husband Dave Goldberg try never to schedule dinners on the same night. If that does happen, she often calls on her sister. "She lives a mile away, and the answer is always yes," Sandberg says.

On their first outing, years before they started dating, Sandberg fell asleep on Goldberg's shoulder during a movie. "I was smitten, but I found out later she does this to everyone," he says. Her favorite film is 1994's *The Shawshank Redemption*. The last time she picked a movie for a group of friends, she chose *Fame*. As punishment, the group made her sit through the whole film. And not sleep.

In many ways her domestic life is very traditional. The family plays a lot of games; Zuckerberg recently taught them the *Settlers of Catan*. Her kids already get their own breakfasts and make their own school lunches (with help). Sandberg says studies that show working moms of today are as engaged with their kids as traditional moms of yore "make me feel so good, so much better." She declines to answer questions about her domestic help, saying it's not a question you would ask a man, then declines my offer to ask Goldberg the same question.

Chapter 8 of *Lean In* claims that one of the most important career choices a woman makes is whom to marry. She and Goldberg, who's as laid-back and

genial as Sandberg is intense and energetic, dated after several years of friendship, during which time Sandberg was briefly married. Four years ago Goldberg left a big job at Yahoo so the family could be together in Northern California. He took over SurveyMonkey, which at the time had 14 employees. "That was hard," he says. "But what Sheryl has been supergreat about is that there may be a time when we're going to move someplace for my career."

(MORE: [Readers Respond: How to Get Ahead at Work](#))

The job change hasn't held Goldberg back. SurveyMonkey now has a staff of 200 and 14 million users, and he just completed a recapitalization of the company that values it at \$1.35 billion.

Sandberg urges women to negotiate shared household duties with their spouses early and often. "We have areas of responsibility. I do travel. I do anything electronic, computers, cars," says Goldberg. "I do photos and videos. We share the child care 50-50. Although it's not like we keep score." And he does the finances. Since Facebook went public, his wife has cashed out about \$90 million worth of shares, according to a schedule that was set before the IPO, and she still has almost 18 million shares left. But she demurs when asked how much she's worth, claiming that that's Goldberg's area. "He manages our money," she says. "I have essentially no interest."

There is always chatter, especially among Californians, that Sandberg, who's a big Democratic fundraiser, will return to the public sector. She has the contacts and skill set. "I really loved being in the government," Sandberg says. "I won't rule out that I would ever want to go in again, but not run for office. But, not now. It's not the right time for my family." According to her father Joel, public policy was always her first love, but he's not sure she isn't there already. "Turns out that she probably has a better platform for doing it this way," he says.

Sandberg doesn't act as if she wants to leave her current job, even though it's almost impossible for her to become CEO. "Ironically, having written a book about women and leadership, having, like, the top leadership role is not the most important thing to me," she says. "I could have done that on the way out of Google. I had those offers."

It may be that solving the problem of fade-out in women's potential is enough of a mission for Sandberg, at least for now. It has proved to be a significant challenge for many of the corporations and governments that have tried to address it. But it's possible that in amassing circles of women and giving them simple empowering tools, she's putting the infrastructure and players in place for a much more ambitious trophy than a seat in the corporate boardroom. Getting women to the highest echelons of business might be her idea of getting them to the starting line. After the women get the power, well, then she can really let loose.

MORE: [TIME's Complete Coverage on Sheryl Sandberg](#)



Belinda Luscombe [@youseless](#)

Luscombe, an editor-at-large at TIME, writes about cultural trends and social sciences. The views expressed are solely her own.

↵

2. Contracting — Due: week 2

References:

- Cooter, Robert, and Thomas Ulen (2008), *Law & Economics*, Pearson International, Boston, MA, pp. on standard form contracts (302-304).
- Katz, Avery W. (1998), "Standard Form Contracts," in Peter Newman, ed., *The New Palgrave Dictionary of Economics and the Law*, vol. 3, London, Macmillan, 502-505.
- Korobkin, Russell (2003), "Bounded Rationality, Standard Form Contracts, and Unconscionability," *The University of Chicago Law Review*, vol. 70, 1203-95 (excerpts).

Cooter, Robert, and Thomas Ulen (2008), *Law & Economics*, Pearson International, Boston, MA, pp. on standard form contracts (302-304).

1. Fill in a Form: Contracts of Adhesion Most written contracts use standard forms. Some terms in a standard-form contract are fixed; others may be variable. For example, the legal staff of an automobile manufacturer may provide its salespersons with form contracts that stipulate the warranty (fixed terms) and leave the price open for negotiation (variable term). Some standard forms do not allow the parties to vary any terms. In an extreme situation, one party makes a take-it-or-leave-it offer, meaning that the other party must sign the standard form or not make a contract.

Many fixed terms in standard-form contracts are uniform throughout an industry. For example, many automobile manufacturers promise to repair certain problems with their new cars within the first 5 years or 50,000 miles of the car's life. When terms are uniform, sellers do not compete over them. Narrowing the scope of competition can reduce its intensity.

To see why, consider cartels. The members of a cartel agree to keep prices up, which profits the members as a group. Each individual member, however, profits even more by undercutting the cartel's price and luring buyers away from other members. To prevent such "cheating," the cartel must punish members who undercut the cartel's price. Uniform, fixed terms in contracts prevent sellers from offering special concessions to buyers. Consequently, the cartel can focus on determining whether all members charge the cartel's price. Monitoring "cheating" in the cartel is much easier when all sellers use the same contract with fixed terms.

In an influential article, Friedrich Kessler called take-it-or-leave-it agreements "contracts of adhesion." (Friedrich Kessler, *Contracts of Adhesion: Some Thoughts About Freedom of Contract*, 43 COLUM. L. REV. 629 (1943).) This term suggests that standard-form contracts indicate the existence of a monopoly, which deprives buyers of bargaining power. Consequently, courts sometimes use "contract of adhesion" as a term of opprobrium to undermine the enforceability of a contract.

This court practice can be justified when sellers use standard-form contracts to reduce competition. However, this court practice is unjustified when sellers use standard-form contracts to increase the efficiency of exchange. Standard-form contracts narrow the scope of bargaining, which can promote efficiency in two ways. First, standard-form contracts can promote price competition by reducing product differentiation. To see why, consider an analogy. Toothpaste comes in different sizes, shapes, colors, textures, tastes, and smells. Manufacturers tinker with these differences in an attempt to attract customers by differentiating their product. Product differentiation complicates price comparisons. Price competition would be more intense if all toothpaste were the same. Similarly, uniformity reduces differences among contracts and intensifies the competition over price.

Second, standard-form contracts reduce transaction costs. The parties can bargain over variable terms, such as the price, and the parties cannot bargain over fixed terms. Instead of bargaining, buyers choose among standardized contracts with different price and non-price terms. Seller may build an actual contract by plugging "modules" of language into a universal form. Thus, standard forms reduce the number of terms requiring drafting, bargaining, and agreement.

One of the standard assumptions of a perfectly competitive market is that transaction costs are zero. Standard-form contracts can move a market closer to the perfectly competitive ideal by reducing transaction costs. The availability of

substitutes in perfectly competitive markets prevents anyone from bargaining over price. Similarly, the availability of substitutes in perfectly competitive markets prevents anyone from bargaining over contracts. In general, substitutes turn everyone into "takers" of the price. The fact that many firms use the same standard form may indicate a high level of competition among them. Take-it-or-leave-it contracts can indicate perfect competition rather than monopoly.

Because standard-form contracts can increase competition and efficiency in exchange, the phrase "contract of adhesion" should not be applied to standard-form contracts. Rather, the phrase should be reserved for monopoly contracts. The relevant question is whether a market is competitive or monopolistic. The fact that a contract was made on a standard form does not establish a presumption in either direction.

What should courts do with the terms in monopoly contracts? In monopoly contracts, the price is too high. Courts, however, usually do not think that adjusting the prices in a contract is their job. Courts are more willing to adjust the non-price terms. Should they?

To answer this question, we first ask whether the nonprice terms of monopoly contracts are efficient or inefficient. The abstract answer given by economic theory is simple. The nonprice terms of a contract typically create incentives that affect the size of the surplus from exchange, and efficient nonprice terms maximize the surplus from exchange. In contrast, the price terms typically distribute the surplus between the parties. Sometimes the monopolist can use its power to extract the entire surplus from each exchange. A monopolist with this power will maximize its profits by maximizing the surplus from each exchange. In brief, a monopolist who can extract all of the surplus from each exchange by controlling the price will choose efficient nonprice terms.

In contrast, a monopolist who cannot extract all of the surplus from exchange by controlling the price may adopt inefficient nonprice terms in order to increase its control over the price terms. (These propositions can be restated in familiar jargon for economists.)⁴⁵ For example, a monopoly supplier of software may increase its power to over-price by contracts that prohibit resale.

Besides monopoly, another defect in markets can cause inefficient standardization of contracts. Lawyers often use the term "contract of adhesion" when a seller takes advantage of a buyer's ignorance. Thus contracts often stipulate a process for resolving future disputes that favor sellers, such as compulsory arbitration before a board organized by the association of sellers. The buyer often fails to read the contract with sufficient care to be aware of such terms, or the buyer is aware

⁴⁵In economic jargon, a perfectly discriminating monopolist sets efficient nonprice terms in its contract. Otherwise a monopolist may use inefficient nonprice terms to increase price discrimination. To illustrate the latter, assume that buyers who are willing to pay a lot for a product also prefer a strong warranty, whereas buyers who are willing to pay a little prefer a weak warranty. Recognizing this fact, the monopolist might offer two contracts: a high-price-strong-warranty contract and a low-price-weak-warranty contract. The difference in warranties helps to separate the two consumer groups so the monopolist can charge them different prices. Without the two nonprice terms, the monopolist cannot tell the two groups apart.

but does not appreciate the term's significance. When such a contract results in a legal dispute, the buyer's lawyer will argue that the court should void the contract because the standardized form prevented the buyer from bargaining. (Remember that according to the bargain theory of contract, which many judges accept in some form, "no bargain" implies "no contract.") This argument, however, misleads. If buyers are informed and markets are competitive, the standardized terms in form contracts will be efficient, not biased against buyers, without any bargaining. The real problem with this kind of contract is the buyer's ignorance, not the absence of bargaining.

QUESTION 7.39: Explain how uniformity can reduce price competition by strengthening cartels or increase price competition by reducing product differentiation.

QUESTION 7.40: Competition drives prices down to costs, whereas monopolies price above cost. California banks have paid large damages for allegedly charging fees greater than the cost of certain services that they provide. Suppose a car manufacturer charges an additional \$450 for an automatic transmission in a new car. What inefficiencies would result if the consumer could sue the manufacturer and make the company prove that \$450 is not disproportionately above the actual cost of the automatic transmission?

QUESTION 7.41: Monopoly distorts contracts by making prices too high. Why would a monopolist ever want to distort the nonprice term by, say, limiting liability for harm caused by defective products?

QUESTION 7.42: Assume that two kinds of buyers purchase contracts from a monopolist who promises to deliver goods in the future. One kind of buyer values the good more highly than the other. The monopolist would like to charge a higher price to the buyers who value the good more highly, but he cannot identify who they are. To overcome this problem, he offers two different contracts. One contract charges a high price and offers to pay high damages in the event that the seller fails to deliver the goods. The other contract charges a low price and offers to pay low damages in the event that the seller fails to deliver the goods. Explain why the two kinds of buyers might prefer different contracts. Explain why the monopolist might gain from offering two kinds of contracts. (In economic jargon, the "menu" of contracts "separates" the "pool" of buyers and permits "price discrimination.")

Wilson, D.S. and Sober, E. 1994. Reintroducing group selection to the human behavioral sciences. *Behavioral and Brain Sciences* 17: 585-608.

Wilson, E.O. 1971. *The Insect Societies*. Cambridge, MA: Harvard University Press.

stakeholders. See STANDING.

standard form contracts. Standard form contracts are an inevitable byproduct of a mass production economy. Just as fixed costs and scale economies in production lead manufacturers to develop standardized goods, those regularly engaged in business find it advantageous to standardize the terms on which they conduct exchange, because this saves the expense of negotiating arrangements for each individual transaction. Among legal commentators, however, form contracts have long been received with distrust, and the rules governing their interpretation have engendered considerable controversy.

The main reasons for such distrust are twofold. First, most persons presented with standardized forms do not bother to familiarize themselves with the specific contents, relying instead on the drafter's reputation and on the knowledge that other contracting parties regularly do business on like terms; this reaction is reinforced by the fact that form contracts typically contain abstruse language and are printed in small type. None of these observations fits well with traditional justifications for legal enforcement of contracts based on voluntary and knowing consent. Second, because the advantages of form contracts would be lost if bargains were open to routine renegotiation, their users are often unwilling to do business on other than standard terms. This take-it-or-leave-it aspect of standard forms has seemed to many lawyers to give the drafting party an inequitable degree of control over the bargaining process. Indeed, since Kessler (1943), standard forms have commonly been disparaged as 'contracts of adhesion', and associated both by courts (e.g., *Henningsen v. Bloomfield Motors*) and legal commentators (e.g., Kessler 1943; Rakoff 1983) with monopoly and market power.

While economic analysis has little to say regarding the libertarian objection to standard form contracts or their relationship to personal autonomy, it can help evaluate their effects on efficiency and the distribution of the gains from trade. From such a perspective, standard forms should be analysed like any other productive input, comparable to design, marketing, and technical support. Whether their use raises any special regulatory or policy concerns, therefore, depends on their implications for the standard litany of market failures: scale economies, monopoly, externalities, imperfect information, and the like.

1. MONOPOLY AND ECONOMIES OF SCALE. The popular association of standard form contracts with market power in legal circles and the related claim that they are thus likely to contain suboptimal terms are not well founded. Such arguments are based on two misconceptions. First, the use of standard forms need not indicate market power. Because standardization, like other types of mass production, lowers the per-unit cost of contracting, competitive

firms as well as monopolists have an incentive to use it. Indeed, standardized contracts are widely used in many industries conservative and liberal economists would agree are workably competitive. Even the fact that competing firms offer similar terms is no evidence of collusion (contrary to the supposition of the *Henningsen* court), since such a congruence would also be observed in perfectly competitive markets.

Second and more importantly, the suggestion that a monopolist would want to offer lower-quality contract terms than a competitive firm depends upon a mistaken analogy between quantity and quality. The reason a monopolist finds it profitable to produce an inefficiently low quantity of goods is that, by doing so, it forces consumers to compete against each other for the reduced supply, bidding up the price. If the monopolist tries to reduce quality, in contrast, whether of contract terms or of the underlying goods, it lowers the amount consumers are willing to pay. As Spence (1975) has demonstrated (see Comanor 1985 and Craswell 1991 for legally oriented expositions) a profit-seeking monopolist will thus want to choose the level of product quality that best suits the preferences of the marginal consumer, for this maximizes the markup he can charge. If all consumers have the same willingness to pay for quality, the monopolist will do best to provide a product with socially optimal quality and to extract available profits through a high price. Only if the tastes of the marginal consumer are unrepresentative of others who buy will suboptimal quality be offered. For example, if willingness to pay for quality is positively correlated with willingness to pay for the underlying good (as will be the case when quality and the good itself are both normal goods), the quality of monopolistically supplied goods will be too low, while if that same correlation is negative, quality will be too high.

Similarly, while critics of form contracts have expressed concern that oligopolists might use standard forms to collude on non-price terms, cartel profits are ordinarily maximized by choosing optimal quality. The proper question to ask in the oligopoly setting, rather, is whether standard forms hinder price competition. Antitrust lawyers are appropriately wary of agreements among competitors not to compete on non-price terms, since such agreements and the institutions set up to enforce them can make it easier to engage secretly in price-fixing. Since standard forms do not require substantial ongoing communication among their users, however, and since price is typically one of the few terms they leave open for bargaining, they do not raise the same risks as do explicit non-competition agreements. Furthermore, to the extent that standard forms facilitate comparison shopping on consumers' part, they may increase the intensity of price competition (though the inability to disguise price cuts by making them under non-price cover could strengthen cartels by making it more difficult for individual members to defect without detection).

A potentially more serious antitrust problem is that incumbent firms may have an incentive to distort standard contract terms in order to raise barriers to potential competition. This could result in provisions that provide too much quality from a short-run perspective, such as generous return or repair policies that require entrants to

maintain a vast inventory and expensive service department in order to break into the market. The danger here is analogous to that arising from product tying (Adams and Yellen 1976; Whinston 1990). Alternatively, it could result in too little quality as measured by the parties' formal legal rights, as in the case of liability disclaimers that increase buyers' reliance on seller reputation, which a newcomer cannot match. This is also a form of tying, since it effectively brings dispute resolution services in-house, forcing buyers to purchase them from the seller. Its success in deterring entry will depend on entrants' abilities to develop and market contracts of their own, and to get those contracts enforced in the public legal system.

A more fundamental link between market structure and form contracts arises from the scale economies that motivate standard forms in the first place. Because of the fixed costs associated with developing and marketing new forms, not all contractual terms will be offered in an unregulated market. The same is true of the underlying goods being sold, of course, since some market niches are too small to be profitably served. Thus, at least some consumers with different tastes for quality (and for contractual terms) will find themselves buying a single undifferentiated product from the same seller. In such a context, the proper regulatory question is whether the market provides the right number and selection of contractual products.

From an efficiency standpoint, a new product should be introduced if and only if the additional consumer surplus it generates justifies the extra costs of production, which include both the costs of the new product itself and any additional costs incurred by reducing the market for existing products. A monopolist able to engage in perfect price discrimination and having control over all possible substitutes, accordingly, would always provide an optimal selection of products, but firms unable to capture the full increment of consumer surplus may not. As Spence (1976) and Dixit and Stiglitz (1977) demonstrated, how close actual markets come to the optimum depends on the curvature of demand and cost curves, the degree of substitutability among competing brands, the feasibility of nonlinear pricing, and market structure.

In principle, then, government regulation of form contracts, by influencing the variety of terms offered, could improve market efficiency or redistribute welfare between marginal and inframarginal consumers. In order for the state to do a better job than the market in choosing products, however, it would either have to have access to superior information regarding inframarginal consumer preferences or be better able to bring purchasing power to bear on their behalf. The former possibility seems unlikely, and the latter would probably be best effectuated not by regulation, but by a direct subsidy to whatever product lines or terms are deemed to be underprovided. Such a subsidy would also be the most direct way to redistribute in favour of whatever inframarginal or distinct consumer groups were considered worthy of state support (for example, those with special needs due to physical disability).

2. EXTERNALITIES AND PUBLIC GOODS. A second and related reason why unregulated markets might produce a suboptimal variety of contract terms is externality. Indeed,

the product selection problem discussed in the previous section can be recast as an externality, in that competitors' sales are reduced when a firm introduces a new product. Because monopolistic competitors set price above marginal cost, this is a real and not just a pecuniary externality; the amount that consumers would pay to switch from one product to another does not properly reflect marginal social benefit.

Additionally, standard form contracts generate a classic network externality, in that the value of a particular contractual form tends to vary directly with the number of its users. As Klausner (1995) has argued in the context of corporate charters, the less familiar a contractual term, the harder it is to predict how courts and third parties will interpret and react to it, not to mention how the contracting parties themselves will understand it. Accordingly, rational contractors have an incentive to adopt customary forms in order to save on costs of legal advice, accounting services, and third-party financing. Anyone who chooses a familiar form, or who invests in expertise relating to its use, thus confers a positive benefit on others who use that form and a negative externality on those who use other forms.

These externalities are exacerbated by the fact that contractual innovations are public goods. Like information generally, they are nonrival in use and largely nonexcludable from those who wish to copy them without payment. Prevailing regimes of intellectual property, furthermore, do not accord them much legal protection; while the particular form in which a contract term is expressed can be copyrighted under US law, for instance, the underlying substance cannot. As a result, innovation in form contracts is likely to be inefficiently undersupplied.

Such factors suggest a legitimate role for the state in encouraging the creation and development of contractual forms. One way to do this would be to lessen doctrinal barriers that currently raise the costs of writing and enforcing forms, but these rules are also needed to cope with problems of fraud and asymmetric information in contracting (see the next section, *infra*). Another approach would be to grant form contracts stronger intellectual property protection, though the proper scope of such protection in the presence of network externalities is currently a matter of controversy. And a third would be for the government to supply such forms itself, through the promulgation of default rules of interpretation and legal terms of art. This last approach has substantially influenced the US drafters of the Uniform Commercial Code. For instance, UCC §2-205 provides a formal method for merchants to make irrevocable offers to buy or sell goods without consideration; §§1-201(3), 1-205 and 2-208 provide that the parties' agreement is to be read as incorporating course of dealing, course of performance, and usage of trade; and §§2-316 and 2-319 establish explicit verbal formulas for disclaiming warranties and allocating the risk of loss for goods in transit.

Whether judicial or statutory provision of contractual forms actually increases the variety of terms available on the market, however, is an open empirical question. As Goetz and Scott (1985) have argued, the availability of the UCC's 'off-the-rack' default terms may deter private parties from trying to formulate terms of their own; and if

the goods produced by private parties are better suited to their needs than those supplied by the state, the net effect, could be to reduce efficiency. Such an effect, if it exists, would be a special case of the more general phenomenon (Bergstrom et al. 1986; Andreoni 1988) whereby government provision of public goods can crowd out private provision of those same goods. Indeed, such crowding out could, in the limit, entirely neutralize the effects of government action, though this is a polar case depending on special assumptions. It seems likely, rather, that there is at least some room for state-supplied contract terms to improve market efficiency, since the public-good nature of authoritative legal pronouncements is presumably the chief justification for government enforcement of contracts in the first place.

3. ASYMMETRIC INFORMATION. Perhaps the most plausible rationale for government regulation of standard form contracts, and the one most closely corresponding to traditional legal objections to their use, is asymmetric information. As is well known, market equilibrium is inefficient when transacting parties are differentially informed regarding the value of the good being exchanged. In the 'lemons' model of Akerlof (1970), for instance, where sellers know the precise quality of their product and buyers know only the average quality of all goods being sold, low-quality goods tend to drive high-quality goods off the market. This is because in the absence of special information, a typical buyer will assume she is dealing with an average-quality seller and will pay no more than her reservation price for average-quality goods. Sellers of high-quality goods, therefore, will be unable to recover their costs. The lemons model applies quite straightforwardly to the case of form contracts, since such contracts vary substantially in their terms and the drafting party (effectively, the 'seller' of the contract) knows much more about those terms than the nondrafting party.

Actual markets have developed a variety of methods for addressing the lemons problem, including most prominently the contractual warranty (Grossman 1981), under which a seller undertakes to reimburse losses resulting from low quality. But warranties and analogous promises only work to the extent the parties communicate about them; and in the form contract setting the effective cost of communication is often high. One approach to this problem is to reduce the degree of asymmetric information by making it harder for drafters to include self-serving terms; the doctrine of unconscionability and the maxim that contracts are to be construed against the drafter take this tack. Another is to deny enforcement of certain terms thought problematic unless the drafter has incurred the cost of calling such terms to the other party's attention. For instance, Article 2 of the UCC requires that warranty disclaimers be 'conspicuous' (§2-316) and that form-contract terms barring oral modifications be 'separately signed' (§2-209). Such rules can be thought of as 'penalty defaults' (Ayres and Gertner 1989, 1992), putting the burden of communication on the party who can undertake it most cheaply.

So long as buyers of form contracts face any communication costs at all, however (and even when a strange term

is called to one's attention it is usually still necessary to spend time and effort assessing its import), such penalty default rules will be less than fully effective. The reason is that the buyer's investigation expenditure is a relation-specific investment subject to expropriation by the seller (Katz 1990a, 1990b). Specifically, the buyer must decide whether to invest resources evaluating a standardized form before she knows its terms. Because her costs of becoming informed are sunk once incurred, she can wind up in a situation where she just barely wants to accept, but wishes she had not bothered to become informed. The seller's optimal course of action, furthermore, is to choose terms that place the buyer in precisely this situation. Since the buyer can anticipate this turn of events, she prefers not to read in the first place. (The argument extends to situations in which buyers attach different levels of importance to quality, as buyers who value quality the highest will find that it is not worthwhile to read, but then the seller will want to reduce quality to the reservation level of the next-most-sensitive buyers, and so on.) In equilibrium, therefore, buyers will not read, sellers will offer the lowest possible quality terms, and buyers will refuse to pay more than fly-by-night prices. While the last two results may be moderated in actual markets by sellers' concerns for reputation (Klein and Leffler 1981; Baird and Weisberg 1982), the first is consistent with empirical behaviour.

A legal rule that imposes a duty to read on the buyer makes no difference to this logic. While sellers might wish to offer higher quality contracts at higher prices, they have no means to commit themselves to doing so, since they can always act opportunistically by sneaking disclaimers into the fine print. Rules providing implied warranties and refusing to give effect to fine-print terms, however, can improve the efficiency of exchange by restricting the scope for such opportunism.

Such a conclusion of course depends on regulators' ability to set the implied and unenforceable terms correctly; if warranties are set too high, the result could be worse than *laissez-faire*. So long as state-set terms are on average more efficient than the quality levels that would be set in an unregulated market, however, and so long as there is a cheap way to contract around the legal default, implied warranties will likely improve the functioning of the market. US warranty law under the UCC arguably approximates such a policy in non-consumer settings, combining an implied warranty of merchantability defined by ordinary market expectations with reasonably straightforward requirements for disclaimer. The practical difficulties of predicting when courts will find a breach of warranty in specific cases and the substantial quantity of litigation arising as a result, however, leaves this conclusion open to objection. In the consumer area, in contrast, public policy has been for the most part less controversial, even though under federal law (e.g., the Magnuson-Moss Warranty Act and regulations promulgated under its authority) most consumer warranties are not disclaimable. This is due to the fact that the lemons problem is worsened in consumer markets by the higher degree of asymmetric information and by consumers' relatively lesser sophistication and access to legal assistance.

4. CONCLUSION. Standard form contracts arise from technological features of the world that depart from the neoclassical model of perfect competition as well as from the classical model of contract. Thus, one should expect some inefficiencies to arise from their use, at least in the first-best sense that the price of goods or of particular contract terms diverges from marginal cost. Whether standard forms benefit contracting parties overall, however, is a problem of the second best, depending upon whether these inefficiencies are outweighed by the savings they bring to the process of exchange. And as with product standardization more generally, they almost certainly are. No modern commentator proposes a return to a classical regime based on individualized bargaining. Rather, the relevant policy question is whether state regulation of form contracts can improve on the unregulated market.

The empirical factor most relevant to answering this question is the extent and effectiveness of reputation. If reputational concerns lead drafters of forms to moderate their opportunism, regulation may be largely unnecessary. But in markets where reputation is insufficient to address the lemons problem – where purchase is infrequent, parties are isolated, and sellers can easily discriminate between sophisticated and unsophisticated buyers – state intervention can lower the efficiency costs of form contracts, making them more useful to sellers and buyers alike. Thus, a measure of special legal treatment for standard form contracts is appropriate on economic grounds.

AVERY WIENER KATZ

See also COGNITION AND CONTRACT; CONSUMER PROTECTION; CONTRACT FORMATION AND INTERPRETATION; NETWORK EFFECTS AND EXTERNALITIES; TRADEMARKS; TYING; UNCONSCIONABILITY.

Subject classification: 5c(1).

STATUTES

- UCC §1-201(3) ('General Definitions – Agreement') (1990).
- UCC §1-205 ('Course of Dealing and Usage of Trade') (1990).
- UCC §2-205 ('Firm Offers') (1990).
- UCC §2-208 ('Course of Performance and Practical Construction') (1990).
- UCC §2-209 ('Modification, Recission and Waiver') (1990).
- UCC §2-316 ('Exclusion or Modification of Warranties') (1990).
- UCC §2-319 ('F.O.B. and F.A.S. Terms') (1990).
- Magnuson-Moss Warranty Act, 15 USC §§2301–2312 (1996).

CASES

Henningsen v. Bloomfield Motors, Inc., 161 A2d 69 (NJ 1960).

BIBLIOGRAPHY

- Adams, W.J. and Yellen, J.L. 1976. Commodity bundling and the burden of monopoly. *Quarterly Journal of Economics* 90: 475–98.
- Akerlof, G. 1970. The market for 'lemons': quality uncertainty and the market mechanism. *Quarterly Journal of Economics* 84: 488–500.
- Andreoni, J. 1988. Privately provided public goods in a large economy: the limits of altruism. *Journal of Public Economics* 35: 57–73.
- Ayres, I. and Gertner, R. 1989. Filling gaps in incomplete contracts: an economic theory of default rules. *Yale Law Journal* 99: 87–130.
- Ayres, I. and Gertner, R. 1992. Strategic contractual inefficiency and the optimal choice of legal rules. *Yale Law Journal* 101: 729–73.

- Baird, D.G. and Weisberg, R. 1982. Rules, standards, and the battle of the forms: a reassessment of 2–207. *Virginia Law Review* 68: 1217–62.
- Bergstrom, T.C., Blume, L.E. and Varian, H.R. 1986. On the private provision of public goods. *Journal of Public Economics* 29: 25–50.
- Comanor, W. 1985. Vertical price-fixing, vertical market restrictions, and the new antitrust policy. *Harvard Law Review* 98: 983–1002.
- Craswell, R. 1991. Passing on the costs of legal rules: efficiency and distribution in buyer–seller relationships. *Stanford Law Review* 43: 361–98.
- David, P.A. 1985. Clio and the economics of QWERTY. *American Economic Review, Papers and Proceedings* 75(2): 332–7.
- Dixit, A.K. and Stiglitz, J.E. 1977. Monopolistic competition and optimum product diversity. *American Economic Review* 67: 297–308.
- Goetz, C.J. and Scott, R.E. 1985. The limits of expanded choice: an analysis of the interactions between express and implied contract terms. *California Law Review* 75: 261–322.
- Grossman, S. 1981. The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics* 24: 461–83.
- Katz, A. 1990a. The strategic structure of offer and acceptance: game theory and the law of contract formation. *Michigan Law Review* 89: 215–95.
- Katz, A. 1990b. Your terms or mine? the duty to read the fine print in contracts. *RAND Journal of Economics* 21: 518–37.
- Kessler, F. 1943. Contracts of adhesion – some thoughts about freedom of contract. *Columbia Law Review* 43: 629–42.
- Klausner, M. 1995. Corporations, corporate law, and networks of contracts. *Virginia Law Review* 81: 757–852.
- Klein, B. and Leffler, K. 1981. The role of market forces in assuring contractual performance. *Journal of Political Economy* 89: 615–41.
- Rakoff, T. 1983. Contracts of adhesion: an essay in reconstruction. *Harvard Law Review* 96: 1174–1284.
- Spence, A.M. 1975. Monopoly, quality, and regulation. *Bell Journal of Economics* 6: 417–29.
- Spence, A.M. 1976. Product selection, fixed costs, and monopolistic competition. *Review of Economic Studies* 43: 217–35.
- Whinston, M. 1990. Tying, foreclosure, and exclusion. *American Economic Review* 80: 837–59.

standards of care. See DUE CARE; LEGAL STANDARDS OF CARE.

standing. Standing to sue regulates access to court by determining who may bring a particular lawsuit. Standing is an important legal doctrine because access to court largely determines an individual's rights. An individual has a right when he may call upon coercive powers of the state to take a specified action. Under the Anglo-American legal tradition, one typically secures the coercive powers of the state by first going to court and petitioning a judge to order a governmental official to take a specified act. Thus, standing in many ways determines the rights of individuals and thereby substantially influences our social and economic interactions.

Most analyses of standing have two characteristics. First, they stress that it is hard to generalize about standing because it is allegedly amorphous, complex, and inconsistent. Second, the analyses tend to be normative, typically arguing for increased access to court. Yet when stripped to its essence – who has access to court – standing has simple, predictable consequences; these consequences are the focus of this essay.

Bounded Rationality, Standard Form Contracts, and Unconscionability

Russell Korobkin†

Economic theory suggests that, in most circumstances, market forces will ensure that standard form contracts contain terms that are not only socially efficient but also beneficial to non-drafting parties as a class compared to other possible combinations of price and terms. This analysis in turn suggests that courts should enforce all form terms or, at a minimum, all form terms that non-drafting parties read and understand. Relying on social science research on decisionmaking, this Article argues that non-drafting parties (usually buyers) are boundedly rational decisionmakers who will normally price only a limited number of product attributes as part of their purchase decision. When contract terms are not among these attributes, drafting parties will have a market incentive to include terms in their standard forms that favor themselves, whether or not such terms are efficient. Thus, there is no a priori reason to assume form contract terms will be efficient. The Article then argues that the proper policy response to this conclusion is greater use of mandatory contract terms and judicial modification of the unconscionability doctrine to better respond to the primary cause of contractual inefficiency.

INTRODUCTION

More than thirty years ago, W. David Slawson estimated that 99 percent of all contracts did not resemble the Platonic ideal of a list of jointly negotiated terms but were instead presented by one party to the other on a pre-printed form.¹ If anything, the dominance of form contracts over negotiated contracts has increased in the intervening decades. The terms of mergers, joint ventures, and very large transactions are sometimes dickered, one at a time in the classic fashion, but nearly all commercial and consumer sales contracts are form driven.² While a few terms—price often being one—might be negotiated on a

† Professor, UCLA School of Law. This Article benefited tremendously from the comments of workshop participants at the Max Planck Institute's Common Goods Project Group, the University of Pennsylvania Law School, the University of Southern California Law Center, and the University of Nevada-Las Vegas Boyd School of Law; as well as from Jennifer Arlen, Rachel Croson, Christoph Engel, Sam Fraidin, Chris Guthrie, Bentley MacLeod, Jeff Rachlinski, Dan Simon, Eric Talley, Tom Ulen, and Stephen Ware; and the research assistance of Heather Richardson and Dominik Sklenar.

¹ W. David Slawson, *Standard Form Contracts and Democratic Control of Lawmaking Power*, 84 Harv L Rev 529, 529 (1971).

² See, for example, Robert A. Hillman and Jeffrey J. Rachlinski, *Standard-Form Contracting in the Electronic Age*, 77 NYU L Rev 429, 435 (2002) ("People encounter standard forms in most of their contractual endeavors. . . . [S]tandard forms govern [most] contractual relationships.").

deal-by-deal basis, the boilerplate “fine print” usually specifies the breadth of the parties’ obligations to one another, including, to use some prominent examples, terms that govern the extent of the seller’s warranties, which party will bear the risk of various types of losses, the extent to which the buyer or seller may recover damages in the event of breach, and the type and location of forums available to resolve disputes between the parties.³ Such forms are often referred to as “contracts of adhesion,” as one party presents the terms to the other on a take-it-or-leave-it basis with no opportunity for negotiation,⁴ although form terms are not necessarily adhesive and not all adhesive terms are presented on a pre-printed form.⁵

Contract law generally provides for the enforcement of the terms in form contracts,⁶ thus essentially allowing the drafting party (almost always the seller in consumer contracts but sometimes the buyer in commercial contracts) to create its own private law to govern its transactions.⁷ If the non-drafting party indicates his general assent to the form, courts will enforce the terms contained therein whether or not that party approves of the terms provided, understands those terms, has read them, or even has the vaguest idea what the terms might be about. Limited exceptions are made to this rule, most notably if the terms are found to be “unconscionable.”⁸

The prevailing rule of form-term enforcement upsets many scholars, who recommend law reform.⁹ These critics decry the “unfairness”

³ In 1999, the New Jersey Law Revision Commission surveyed the terms of fifty common types of form contracts and determined that the following terms appeared with regularity and could potentially be abusive: (1) warranty; (2) damages; (3) attorneys’ fees; (4) refund and repair; (5) indemnification; (6) risk of loss; and (7) waiver of rights. See John J.A. Burke, *Contract as Commodity: A Nonfiction Approach*, 24 Seton Hall Leg J 285, 293 (2000).

⁴ See, for example, *Rudbart v North Jersey District Water Supply Commission*, 127 NJ 344, 605 A2d 681, 685 (1992); Todd D. Rakoff, *Contracts of Adhesion: An Essay in Reconstruction*, 96 Harv L Rev 1174, 1177 (1983).

⁵ See Arthur Allen Leff, *Unconscionability and the Code—The Emperor’s New Clause*, 115 U Pa L Rev 485, 505 (1967) (noting that a form contract is a symptom of adhesion but not its “essence”).

⁶ See, for example, *Graham v Scissor-Tail, Inc.*, 28 Cal 3d 807, 623 P2d 165, 172 (1981) (“[A] contract of adhesion is fully enforceable . . . unless certain other factors are present which, under the established legal rules . . . operate to render it otherwise.”). For a more complete discussion of the exceptions to the rule, see Part IV.

⁷ See Slawson, 84 Harv L Rev at 536 (cited in note 1).

⁸ See Part IV. Courts occasionally invoke other related doctrines such as “reasonable expectations” or “public policy” to invalidate terms in form contracts, but to the extent that those doctrines relate specifically to form contract terms, they have for the most part either become part of unconscionability analysis or indistinguishable from it. Thus, these “related doctrines” are discussed along with unconscionability in Part IV.

⁹ See, for example, W. David Slawson, *The New Meaning of Contract: The Transformation of Contracts Law by Standard Forms*, 46 U Pitt L Rev 21, 23 (1984) (recommending that the reasonable expectations of the parties be enforced); Rakoff, 96 Harv L Rev at 1180–83 (cited in note 4) (arguing that form terms should be presumptively unenforceable); K.N. Llewellyn, Book Review, 52 Harv L Rev 700, 704 (1939) (arguing that unreasonable form terms should not be en-

of the ability of drafting parties (hereinafter “sellers,” although the drafting party is not always a seller) to impose adhesive terms on non-drafting parties (hereinafter “buyers” or “purchasers”). Two specific objections to the enforceability of terms embedded in form contracts are most plausible.

First, some critics argue that the enforcement of form terms is objectionable because it undermines individual autonomy, as the buyer finds herself obligated to terms to which she did not voluntarily agree.¹⁰ Freedom of contract demands freedom *from* contract, and just as no party has the ability to force another into a contract,¹¹ no party should have the ability to force another party to accept specific terms.¹²

The problem with this argument is that, given the complexity of modern commerce, the alternative to form contracts is almost certainly not the resurgence of fully dickered, obligationally complete contracts, but rather law-imposed default terms invoked to fill gaps in the contract the parties negotiate.¹³ Actual assent to each contract term in a transaction of any complexity simply is not possible; if terms are not imposed on one party by the other, some terms will almost certainly be imposed on both parties by the government.

The alternative objection is consequentialist in nature: The routine enforcement of form terms results in contracts being less favorable to buyers than they otherwise would be. By seeding the “fine print” with pro-seller terms and then refusing to negotiate those terms, sellers capture more of the cooperative surplus created by the agreement than they would if terms were negotiated.

forced), reviewing O. Prausnitz, *The Standardization of Commercial Contracts in English and Continental Law* (Sweet and Maxwell 1937); Michael I. Meyerson, *The Reunification of Contract Law: The Objective Theory of Consumer Form Contracts*, 47 U Miami L Rev 1263, 1299 (1993) (claiming that consumers should be bound only to the terms they know and understand); Alex Y. Seita, *Uncertainty and Contract Law*, 46 U Pitt L Rev 75, 132 (1984) (proposing that contracts should be governed by default terms some of which may only be overcome when the disadvantaged party has given “intelligent and meaningful approval”); Jeffrey L. Harrison, *Class, Personality, Contract, and Unconscionability*, 35 Wm & Mary L Rev 445, 489 (1994) (calling for an “expanded notion of unconscionability” to prevent “uneven exchanges”).

¹⁰ See, for example, Rakoff, 96 Harv L Rev at 1237 (cited in note 4) (“[E]nforcing boilerplate terms trenches on the freedom of the adhering party.”); Slawson, 84 Harv L Rev at 530, 542 (cited in note 1) (stating that most standard form contracts do not embody the democratic consent of the parties). But see Charles Fried, *Contract as Promise: A Theory of Contractual Obligation* 1–2 (Harvard 1981) (arguing that preserving party autonomy should be the primary goal of contract law).

¹¹ But see *Hobbs v Massasoit Whip Co*, 158 Mass 194, 33 NE 495, 495 (1893) (Holmes) (stating that a course of dealings between parties might change the background rule that a recipient of unrequested goods is not forced to return them or pay for them).

¹² See, for example, Rakoff, 96 Harv L Rev at 1238 (cited in note 4) (arguing that a concern for the freedom of the adhering party “push[es] toward the conclusion that such terms should be completely unenforceable”).

¹³ See, for example, *id* at 1246–47.

The problem with this argument is that standard law-and-economics reasoning suggests that, if buyers and sellers behave in accordance with assumptions of rational choice theory, the operation of the market usually will provide drafting parties with an incentive to include only efficient terms in form contracts.¹⁴ Counterintuitively, a well-functioning market should ensure that buyers and sellers actually prefer the *same* contract terms. If they do, buyers are best served if courts enforce all terms in form contracts, even when those terms are adhesive. Consequently, to establish that form terms disadvantage buyers, and thus make out a prima facie case that a policy of routinely enforcing form terms is undesirable on that ground, a theory of market failure is required that explains why, contrary to the predictions of standard economic theory, sellers would have a profit incentive to place inefficient rather than efficient terms in form contracts.

This Article provides such a theory. Terms that govern the contractual relationship between buyers and sellers are attributes of the product in question, just as are the product's price and its physical and functional characteristics. Because buyers are boundedly rational rather than fully rational decisionmakers, when making purchasing decisions they take into account only a limited number of product attributes and ignore others. While sellers have an economic incentive to provide the efficient level of quality for the attributes buyers consider ("salient" attributes), they have an incentive to make attributes buyers do not consider ("non-salient" attributes) favorable to themselves, as doing so will not affect buyers' purchasing decisions. Assuming that price is always a salient product attribute for buyers, market competition actually will *force* sellers to provide low-quality non-salient attributes in order to save costs that will be passed along to buyers in the form of lower prices. Ironically, the consequence of market forces in a world of boundedly rational buyer decisionmaking is that contracts will often include terms that are socially inefficient, leave buyers as a class worse off (judged from the perspective of buyers' subjective preferences)¹⁵ than they would be if their contracts included only efficient terms, and leave sellers as a class worse off as well.

Courts can increase utility for buyers and sellers, as well as promote social efficiency, by enforcing efficient terms in form contracts and refusing to enforce inefficient terms. Courts' present use of unconscionability and related doctrines to strike objectionable terms

¹⁴ See Part I.A.

¹⁵ Although the argument can be made that individuals are incapable of determining what is best for them, and thus the paternalistic intervention of the state is appropriate. See, for example, Duncan Kennedy, *Distributive and Paternalist Motives in Contract and Tort Law, with Special Reference to Compulsory Terms and Unequal Bargaining Power*, 41 Md L Rev 563, 624-29 (1982).

from form contracts is not well calibrated to produce this outcome, as the factual circumstances that trigger findings of unconscionability under the doctrine are, at best, weakly correlated with the main cause of inefficiency in form terms. By recognizing purchasers' bounded rationality as the most important root cause of inefficiency in form contracts, courts can modify their use of unconscionability analysis to increase both social welfare generally and buyer welfare specifically.

Courts' initial analytical step should be an analysis of whether a challenged contract term is salient to a significant number of buyers. When a contract term is salient to purchasers, the market can be trusted to provide an efficient version of the term: Absent fraud, duress, or significant third-party externalities, no judicial intervention is necessary. When a contract term is non-salient to most purchasers, the market check on seller overreaching is absent, and courts should be suspicious of the resulting term. Put slightly differently, whenever a term in a form contract is non-salient to most purchasers, those purchasers are incompetent to protect their interests vis-à-vis that term. In that situation, legislatures should mandate efficient terms ex ante when possible, and courts should police ex post for clearly inefficient terms.

This Article presents this argument in the following manner. Part I describes how, assuming basic postulates of economic analysis, the market should ensure that terms in form contracts are both socially efficient and desirable for both buyers as a class and sellers as a class. Without market failure, there is no valid consequentialist argument for non-enforcement of any contract terms, whether provided on a pre-printed form or offered on an adhesive basis.

Part II argues that the reason form terms deserve scrutiny is that buyers are not fully rational, but rather make decisions in a boundedly rational manner, and that this provides sellers with an incentive to draft non-salient contract terms to their own advantage, whether or not such terms are efficient.

Part III considers alternative conceptual approaches to policing the terms of form contracts in light of the incentives created by buyer bounded rationality. It concludes that ex ante legislative regulation of form contracts by promulgating mandatory terms should be a part of the response but cannot be the complete response. In addition, ex post regulation of form terms by courts is also necessary.

Part IV critically examines the doctrinal tools that courts currently use to police the enforcement of form contract terms—most prominently the unconscionability doctrine—and finds that current judicial doctrine is not well calibrated to the goal of mitigating the pernicious effects of form contracts.

Part V provides specific recommendations for how courts can and should modify the unconscionability doctrine to better police inefficient form terms. It contends that (1) “procedural unconscionability” analysis should be motivated by an inquiry into a term’s salience, (2) “substantive unconscionability” determinations should depend on whether terms are more costly to buyers than they are beneficial to sellers *ex ante*, (3) courts should require buyers to meet an exacting burden of proof before finding a term unconscionable under this criterion, and (4) courts should liberally refuse to enforce terms found unconscionable under this standard, and even refuse to enforce entire contracts on some occasions, in order to provide an incentive to sellers to draft efficient form contract terms *ex ante* when the market fails to provide such an incentive.

This Article concludes with brief discussions of whether “bounded rationality,” as the term is used here, is or is not “rational” behavior and the importance of comparative institutional analysis in devising any legal policy response to bounded rationality in a variety of contexts.

I. RATIONAL ACTORS AND FORM CONTRACT TERMS

Economic analysis suggests that in a perfectly functioning market with complete information contracts between buyers and sellers will contain only efficient terms,¹⁶ defined as those for which the differential between benefits and costs is greatest, regardless of how distributed between buyers and sellers.¹⁷ Economic theory also suggests that substituting an inefficient term into the contract would make both buyers and sellers worse off.¹⁸ The implication of these two propositions is that, in the absence of significant negative externalities to third parties, courts should never refuse to enforce contract terms, even if the terms are embedded in pre-printed forms and offered on an adhesive basis. To do so would be socially inefficient, and it would make buyers, as well as sellers, worse off than they otherwise would be.¹⁹ Sec-

¹⁶ See, for example, R. Ted Cruz and Jeffrey J. Hinck, *Not My Brother's Keeper: The Inability of an Informed Minority to Correct for Imperfect Information*, 47 *Hastings L J* 635, 638 (1996); Eric A. Posner, *Contract Law in the Welfare State: A Defense of the Unconscionability Doctrine, Usury Laws, and Related Limitations on the Freedom to Contract*, 24 *J Legal Stud* 283, 284 (1995).

¹⁷ For the definition of efficiency in this context, see Richard Craswell, *Passing on the Costs of Legal Rules: Efficiency and Distribution in Buyer-Seller Relationships*, 43 *Stan L Rev* 361, 363 (1991).

¹⁸ In economic language, terms for which the net marginal benefits outweigh the net marginal costs will be Pareto efficient as well as Kaldor-Hicks efficient.

¹⁹ See Hillman and Rachlinski, 77 *NYU L Rev* at 432 (cited in note 2) (“[F]ailure to enforce standard terms can harm both consumers and businesses.”); Alan Schwartz and Louis L. Wilde, *Imperfect Information in Markets for Contract Terms: The Examples of Warranties and Security Interests*, 69 *Va L Rev* 1387, 1392–93 (1983) (“[A]ssuming a given distribution of wealth,

3. Public sector reform — Due: week 3

References:

Jensen, Michael C., and William H. Meckling (2009), “Specific Knowledge and Divisional Performance Measurement,” *Journal of Applied Corporate Finance*, 21(2), 49-57.

Arruñada, Benito (2013), “Markets for Public Services: Strong Incentives vs. Comprehensive Planning?” Universitat Pompeu Fabra, March.

Specific Knowledge and Divisional Performance Measurement

by Michael C. Jensen, Harvard Business School, and William H. Meckling, University of Rochester*

Performance measurement is one of the critical factors that determine how individuals in an organization behave. Performance measurement includes subjective as well as objective assessments of the performance of both individuals and subunits of an organization such as divisions or departments. Besides the choice of the performance measures themselves, performance evaluation involves the process of attaching value weights to the different measures to represent the importance of achievement on each dimension.

A company's performance measurement and evaluation system is one of the three important aspects of organizational design that we refer to collectively as the organizational "rules of the game." The other two are the reward and punishment system—including nonmonetary rewards and promotions as well as salary changes and bonuses—and the system for assigning "decision rights" to individuals in an organization. By "decision rights," we mean the rights to decide on and take an action.¹

These three aspects of the rules of the game are obviously related. If the performance measures are to have the desired effects on the behavior of an organization's members, the reward and punishment system must link rewards with performance in a clear and consistent way. Furthermore, the performance measures should also be consistent with the ways in which decision rights are allocated throughout an organization. For example, it is less important—and may even be counterproductive—to measure the efficiency with which a manager employs plant and equipment using return on assets (as opposed to, say, total dollars of profits) if the manager does not influence decisions to invest in plant and equipment and there is nothing he or she can do to offset the effect that asset decisions have on performance.

In this paper, we examine five common divisional perfor-

mance measurement methods—cost centers, revenue centers, profit centers, investment centers, and expense centers—and furnish the outlines of a theory that attempts to explain when each of these five methods is likely to be the most efficient. As discussed below, the central insight of our theory is that each of these methods can be seen as providing an alternative way of aligning corporate decision-making authority with valuable "specific knowledge" inside the organization. Stated as briefly as possible, our theory suggests that cost and revenue centers work best in cases where headquarters has (or can readily obtain) good information about cost and demand functions, product quality, and investment opportunities. Decentralized profit and investment centers will tend to supplant revenue and cost centers when the managers of business units have a significant informational advantage over headquarters.

Specific and General Knowledge

Perhaps the most important consideration in designing a performance measurement system is the cost of acquiring and transferring knowledge among decision agents—that is, managers and employees. We define specific knowledge as knowledge that is costly to transfer among agents and that is not easily observable by other agents (particularly, by managers higher in the organization's hierarchy). General knowledge is information that is transferable among agents at low cost or is easily observable by other agents.²

Idiosyncratic knowledge about people, machines, organizations, customers, and suppliers, as well as knowledge of time and place, are all examples of specific knowledge. Such knowledge is difficult or impossible to aggregate—and information about time and place is by its very nature destroyed by the process of aggregation. Specific knowledge is also often obtained at low cost by individuals in an organization as a byproduct of other activities—for example, the idiosyn-

*The original version of this article was published as "Divisional Performance Measurement" in Michael C. Jensen, *Foundations of Organizational Strategy*, Chapter 12 (Harvard University Press, 1998). A working version was first presented at the Harvard Colloquium on Field Studies in Accounting, June 18-20, 1986. Research for this paper was supported by the Division of Research, Harvard Business School, and the Managerial Economics Research Center, University of Rochester.

1. This includes not only sole and complete control over a given decision, but a range of possibilities for influencing the decision. In large organizations, it is common for no single person to have all the decision rights necessary to undertake a major project. Instead, there is a complex process that brings many people into the decision-making function, a process that breaks the simple notion of a decision right into many components that are allocated to various decision agents. A common breakdown of such components is as follows: (1)

initiation rights—the right to initiate resource allocation proposals; (2) notification rights—the right to be notified of the actions or proposed actions of others and the right to provide information or recommendations about those proposals; (3) ratification rights—the right to review and ratify or veto the resource allocation recommendations of others; (4) implementation rights—the right to implement the ratified proposals; and (5) monitoring rights—the right to monitor the implementation of ratified proposals, including the rights to measure and evaluate performance and to determine rewards and punishments. For a discussion of initiation, ratification, implementation, and monitoring rights, and how their assignment can reduce agency costs, see Eugene Fama and Michael Jensen, "Separation of Ownership and Control," *Journal of Law and Economics* 26 (June 1983).

2. The terms specific and general knowledge are used to characterize the two ends of a continuum that measures the cost of transferring knowledge between agents. >>

cratic knowledge about a machine that its operator gains over time. In contrast, prices and quantities are examples of general knowledge that is easily aggregated and inexpensive to transmit among agents.

Achieving effective use of information in decision-making is a major problem in organizations. The literature in computers and information systems views the problem as one of finding ways to transfer knowledge relevant to a decision to the agents involved in the decision. This makes sense when the knowledge is general, or when the problem is one of discovering new technology (for example, in computing or communications) that will effectively convert specific to general knowledge. But when the relevant knowledge is specific and when technology is unable to lower the cost of transfer substantially, this approach will fail.

The alternative to moving the knowledge is to move the decision rights to those agents who possess the relevant specific knowledge. Although this will sometimes mean greater coordination, or centralization, of decision-making (for example, in the case of an industry-consolidating merger), in the majority of cases the transfer of decision rights is likely to be away from headquarters—that is, a decentralization of decision-making. But even though decentralization has the potential to improve decision-making, it is not without costs of its own. The costs incurred are those that arise from the potential conflict between the private interests of managers and employees and the goals of the organization. As the decision rights are transferred to managers and employees at lower levels in the organization, such agents are likely to use the decision rights in ways that benefit themselves at the expense of the performance of the organization. This potential conflict of interest makes it necessary to devote resources to controlling the costs associated with the inconsistent objectives of agents in the organization—what have come to be called “agency costs.”

The concept of agency costs, as we defined it in our 1976 paper in the *Journal of Financial Economics*, include the following: the costs of devising and enforcing contracts with agents; the costs of monitoring the agents’ behavior; the bonding costs incurred by the agent to help assure the principal that he or she will not engage in opportunistic behavior; and, finally, the “residual loss”—that is, the reduction in firm value resulting from the reality that it is not cost-effective (if indeed possible at all) to define and enforce contracts perfectly. This residual loss arises because it pays companies to incur additional monitoring, bonding, and

contracting costs only to the point where the improvements in the decision process just pay for themselves. Thus, not all counterproductive behavior is eliminated.

Alternative Divisional Performance Measures

There are five major categories of performance measurement systems: (1) cost centers, (2) revenue centers, (3) profit centers, (4) investment centers, and (5) expense centers. We briefly discuss each of these measurement systems and then turn to an analysis of the conditions under which each will tend to be an efficient system.

Cost Centers

Cost centers are designed to encourage managers to focus on increasing the efficiency of the production process without the distractions caused by changes in demand conditions that would affect them if revenues were included in the performance measure. But, as we argue below, cost centers are likely to be effective only in certain, fairly restrictive, circumstances.

There are three alternative ways to define the objective for a manager whose performance measurement and evaluation system is set up as a cost center:

1. Minimize costs for given output.
2. Maximize output for given total cost.
3. Minimize average costs (with no quantity constraint).

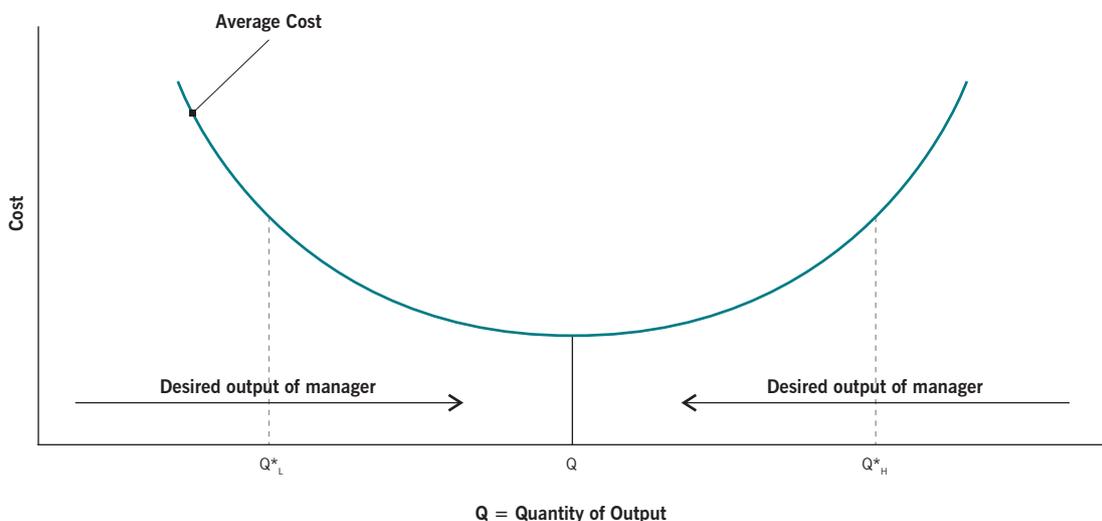
Rules 1 and 2 are logically equivalent and, provided management chooses the right level of output or the optimal total cost constraint, they are both consistent with maximizing the value of the firm. Rule 3, however, is logically inconsistent with maximizing the value of the firm because it motivates the cost center manager to achieve a level of output that minimizes average cost. And, as illustrated in Figure 1, that level of output will not be the value-maximizing level except by accident.

Figure 1 illustrates the point for a manufacturing division with a standard U-shaped average cost function that is evaluated as a cost center. The figure portrays two alternative optimal output levels, Q_L^* and Q_H^* , for two alternative sets of demand conditions (“L” denoting low demand and “H” denoting high demand). Since minimum average cost occurs at output level Q , regardless of whether demand for the product is high or low, that is where the divisional manager will choose to operate. The manager will produce more than the optimal level of output when demand is low, and too little output

The importance of the costs of transferring knowledge was suggested by our reading of Friedrich von Hayek’s seminal article “The Use of Scientific Knowledge in Society,” *American Economic Review* 35, No. 4 (September 1945). Although he used the terms scientific and particular knowledge, we believe specific and general knowledge defined in terms of cost of transferring knowledge between agents capture the important dimensions of Hayek’s discussion of the role of knowledge. Oliver Williamson uses the term “information impactedness” to characterize a similar phenomenon (see *Markets and Hierarchies*:

Analysis and Antitrust Implications, New York: Free Press, 1975). That term, however, does not suggest a continuum in which the costs of information transfer can vary, and this seriously limits the effectiveness of the analysis. The notion of “asymmetric information” widely used in the principle/agent literature deals with the same issue and has the same problem (see, for example, M. K. Harris, C.H. Kriebel, and A. Raviv, “Asymmetric Information, Incentives and Intrafirm Resource Allocation,” *Management Science* 28, no. 6 June 1982).

Figure 1 **Desired Output of Manager Evaluated as a Cost Center with No Quantity Constraint***



* Q^*_L and Q^*_H are two alternative optimal outputs.

when demand is high. And the company as a whole will sacrifice the profits that could have been earned from operating at the optimal level of output.

Moreover, even if the division manager does not have the rights to set the output level unilaterally but has input into the decision, he will tend, other things being equal, to provide a constant source of pressure to move the planned output level closer to Q , the minimum average cost output level. In the situation where optimal output is higher than the minimum average cost point, the manager will tend to take actions that reduce output unexpectedly—for example, claiming machine breakdowns or labor or material shortages (that could have been avoided with better planning). And, if it is difficult for those at higher levels in the hierarchy to distinguish the reasons for these events (because the information required to do so is specific and located in the manufacturing division), it will be difficult to eliminate these counterproductive effects from the system as long as the manufacturing division is a cost center.

Good knowledge of the minimum obtainable cost functions would allow the evaluation mechanism to adjust for differences in quantity of output and therefore eliminate the problems associated with incentives to game the system—at least on the quantity dimension. With such knowledge, the evaluation system would measure performance as deviations from the minimum obtainable cost function. In practice, however, knowledge of the cost functions will in general be unavailable or very costly. Standard cost systems are at best a crude attempt to control for the effects of quantity changes. And further reducing their usefulness in such situations,

they make the correct adjustments only when marginal cost is constant—a condition that is unlikely to hold for large changes in output either up or down.

In addition to higher-level uncertainty about optimal quantity and standard costs, another potential problem with cost centers is that a cost center manager has incentives to reduce quality below the optimal level—because that too reduces measured cost. This incentive the system provides to lower quality means that cost centers will tend to work well only when it is inexpensive to ascertain quality as well as the optimal quantity and quality, and how “true” standard costs vary with output. Moreover, for some functions, the measurement of quantity is as difficult as the measurement of quality. Consider, for example, the computer services supplied by a centralized service bureau in a firm. There is no simple, unique (or non-arbitrary) way to measure quantity in such a multi-dimensional environment.

And there is another potential limitation of cost centers. If a division produces different products, the product mix decision will pose serious difficulties in this structure because the relative amounts of each product to be produced must be determined outside of the division and given to the cost center manager as a constraint to be met. This is another example of the need to control the quantity decision for a cost center manager.

As stated earlier, the general principle in assigning decision rights is to attempt to “co-locate” the decision rights with the relevant specific knowledge. The cost center manager is typically given decision rights over the choice of factor inputs, operating procedures, technology, and so on, all of which

generally require a great deal of knowledge that is specific to the local situation. The advantages of this system, when it can be implemented, come from the specialization it encourages. As we noted above, the cost center manager can focus on increasing the efficiency of the production process without the distractions caused by changes in demand conditions.

In sum, *cost centers will tend to be the most efficient performance measurement system when the optimal quantity, quality, and product mix decisions can confidently be made outside the division.* But, when it is difficult to measure quantity and quality, and when the knowledge required to make the optimal quantity, quality, and product mix decisions is specific and inaccessible to those higher in the hierarchy, it will be difficult to operate the division as a cost center. In such cases, as discussed later, profit or investment centers are likely to work better.

Revenue Centers

Revenue centers are the logical complement to cost centers. Revenue centers have essentially the same advantages as cost centers—the greater focus and inducement to specialization associated with greater controllability. But they also have similar problems, and thus limitations on their effectiveness, stemming from lack of specific knowledge at headquarters.

Although the main performance measure in such centers is total revenue, the objective can take one of three logical forms:

1. Maximize total revenues for a given price.
2. Maximize total revenues for a given quantity of unit sales.
3. Maximize total revenues (with no quantity constraint).

Again, the first two of these options are logically the same and, given the correct choice of price or quantity, they are both consistent with maximizing the value of the firm. But, as we saw in the case of cost centers, the revenue center managers cannot be allowed to determine the quantity or they will simply go to the quantity where revenue is maximized—that is, the point where marginal revenue is zero. As long as marginal costs are positive, this will exceed the profit-maximizing quantity.

The product mix decision is a particular problem in revenue centers because of the tendency for the performance measure to become total revenues from *all* products. If so, other things equal, the manager will substitute sales efforts from lower-priced to higher-priced products; and, unless the profit margins (including the associated capital costs) are the same for all products, the higher revenue from such substitution will come at the expense of overall profits. In

this situation, a better performance measure is *gross margin*, defined as the dollar difference between total revenues and total variable costs (including a capital charge).

The advantage of the revenue center is that the manager can specialize in the marketing and sales effort without concern for the factors that influence production cost. To do so the manager will generally be given decision rights over marketing and sales issues that require considerable specific knowledge (available only at the local level), but not the right to decide on quantity or product mix. This means that, *when the knowledge required to make the quantity and product mix decision is available at low cost at higher levels in the hierarchy, the revenue center structure will tend to be efficient.*

Profit Centers

A divisional profit center is evaluated on the difference between its revenues and costs as defined by the measurement system. Although “profit center” refers to a performance measurement system, it is also widely used to describe a divisional structure in which the profit center manager is given a broader set of decision rights. But this, of course, does not prevent profits from being used as a measure of performance in divisions run by managers with limited decision rights. We use the term here to describe a system in which a division’s performance is measured by its profits.

When the knowledge required to make decisions about the product mix, quantity, and quality is specific to the division and therefore costly or impossible for managers at higher levels in the hierarchy to obtain, the profit center can be an effective performance measurement system. In these cases it is desirable to use profits as a performance measure while giving profit center managers decision rights over factors such as the product mix, quantity, and quality.

Nevertheless, the profit center structure also has potentially serious problems of its own. It is well known that maximization of profits for each division does not lead to maximum profits for the firm as a whole, except in the special circumstance in which there are no interdependencies between divisions. Such interdependencies can take the form of:

- Interfirm transactions in which one or more divisions buys the product of another, and therefore the price paid by the buying division affects its costs and pricing decisions (the “transfer pricing” problem);³
- Interdependent demands (e.g., Pontiac and Oldsmobile, or film and cameras), where demand for one or more of the firm’s products depends on the policies for the other products (e.g., pricing, quality, or technology); or
- Interdependent supply or cost functions, where the cost

3. See J. Hirschleifer, “Internal Pricing and Decentralized Decisions,” in *Management Controls: New Directions in Basic Research*, ed. C.P. Bonini, H.M. Wagner (New York: McGraw-Hill, 1964).

of producing a product depends on the production decisions for other products (e.g., gasoline and kerosene, since more gasoline production means less kerosene obtained from a barrel of crude oil).

When interdependencies among centers are significant, profit center performance measurement can lead to seriously suboptimal behavior on the part of divisional managers.

One solution to the first kind of interdependencies—interfirm transactions that give rise to the transfer pricing problem—is for corporate headquarters to set a transfer price that is equal to the marginal cost of the producing division at the optimal quantity of output. But this requires top management to know both the revenue and cost functions in detail (in order to determine the optimal output level in each period and the marginal cost at the optimal output level). If the information necessary to know both revenue and cost functions is specific to the operating divisions, it will be difficult for top management to set the optimal transfer price. On the other hand, if close substitutes for the good being traded internally are also traded in outside competitive markets, the outside market price is a likely candidate for the optimal transfer price. But even in such cases, if there are important synergies in the form of shared costs or benefits among different profit centers, the optimal transfer price could deviate significantly from the market price.

In short, there is no simple solution to the problems caused by interdependencies in demand or cost functions.⁴ If these interdependencies are significant and there is no simple way to coordinate the actions of the two divisions, one possible solution is to merge them into a single division and then apply the profit measure to the sum of the two divisions rather than to either separately. Another possibility is to evaluate and reward division managers based in part on the performance of “related” business units as well as the performance of their own division.

Investment Centers and EVA

Investment centers are a variation on the profit center structure in which the manager is evaluated on some measure that relates profits to the assets (and underlying capital) used to generate them. As such, investment centers are performance measurement systems that take into account the efficiency of asset utilization.

Investment centers are likely to be most effective when managers of the division have the specific knowledge required to decide the optimal level of investment, when they are given (or acquire) decision rights over investment and asset levels, and when the costs associated with suboptimal asset utiliza-

tion are likely to be large. U.S. companies have commonly taken asset utilization into account by using rate-of-return measures such as return on assets (ROA) or return on equity (ROE). But, as we argue below, both of these measures are highly susceptible to gaming and tend to provide counterproductive incentives when managers have decision rights over the level of investment or assets.

Again, as we saw in the case of cost centers and revenue centers, the objective function for an investment center can take one of three forms:

1. Maximize the percentage return on assets for given total assets.
2. Maximize total assets for given total percentage return.
3. Maximize total percentage return on assets (with no constraint on total assets).

Forms 1 and 2 can be consistent with maximizing the value of the firm if the constraints on total assets or total percentage return are chosen correctly, and this can work if top management has the relevant specific knowledge to set the correct constraints. However, a common form for this objective function to take is the unconstrained Form 3, and this is inconsistent with maximizing the value of the firm. A manager evaluated on maximizing the total percentage return on assets has an incentive to reduce assets to the point where the firm owns no assets other than the single asset with the highest return. This, of course, is not consistent with maximizing value or wealth. A 100% return on \$1,000 of assets is \$1,000, while a 30% return on \$100,000 of assets is \$30,000.

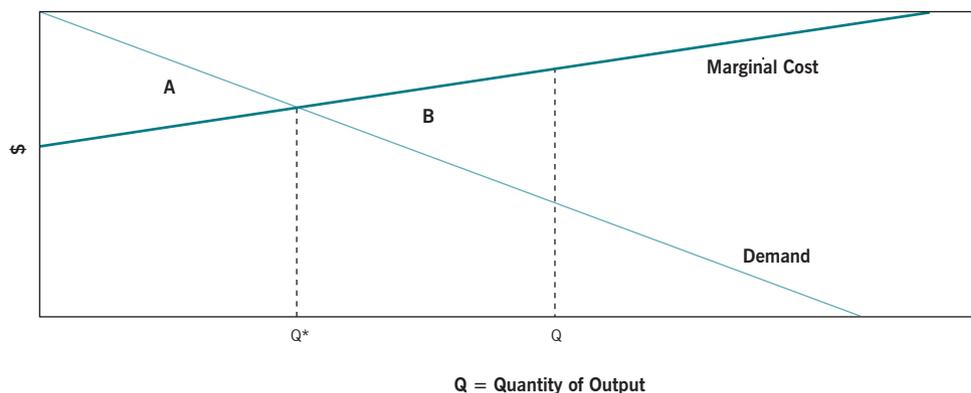
Economic Value Added, or EVA for short,⁵ is an alternative single-period performance measure that eliminates this incentive for underinvestment. EVA is defined as net cash flow in a period less a capital charge equal to the cost of capital times the dollar value of the assets employed in the business. This “residual income,” as it used to be known in the accounting literature, has none of the disadvantages mentioned above of ROA or ROE. Because it is total *dollars* of net cash inflow less the total dollar charges for capital used in the business, EVA is an appropriate number to maximize. EVA also has the advantage of revealing to managers the real cost of capital used in a business. Because accounting statements reflect the cost of debt, but not the cost of equity, used in a business, managers often think and behave as if equity has no cost. EVA accounting statements show a loss when net cash flows are not sufficient to cover the full cost of an organization’s capital.

But EVA is not a panacea. Like all single-period, or flow, measures of performance, EVA fails to solve what we call

4. See Hirshleifer (1964).

5. For a detailed description of EVA and its uses, see G.B. Stewart, *The Quest for Value: A Guide for Senior Managers* (New York: Harper Business, 1991). See also the next article in this issue, Al Ehrbar and Bennett Stewart, “The EVA Revolution.”

Figure 2 The Firm Value Maximizing Level of Output, Q^* , for an Internal Service Center Organized as an Expense Center and the Center's Optimal Output Level, Q



the “capital value” problem.⁶ This problem arises for projects where early years’ EVA is negative, but the future annual EVA of a project is sufficiently large to justify the investment on a net present value basis. Managers evaluated solely on the basis of the current year’s EVA will not take such projects. In these cases, market value—or the discounted present value of net cash flows less the investment required to generate them—is the appropriate value to maximize.

Thus, while EVA is the best flow measure of performance currently known, it is not the universal answer to the search for the perfect performance measure. Perfect measures of capitalized value will never be found because value cannot be known with certainty until *after* a project has run its course to completion and shutdown.

Expense Centers

Expense centers are the private equivalent of the classic public bureaucracy.⁷ A division organized as an expense center generally produces services for the rest of the organization, and the consuming units are not charged for the services they consume. The providers of internal administrative services such as human resources, patent management, and public relations are commonly organized as expense centers.

Consider a division that negotiates a budget allocation from a central budget office at the beginning of each year and simultaneously makes a commitment to the quantity of services that will be provided. Figure 2 portrays the demand for the division’s services from the rest of the organization. This curve plots the marginal value to the organization as a

whole of the division’s output at various levels. (For simplicity, we assume the service is produced with no fixed costs and marginal costs given by the schedule in Figure 2.)

The profit-maximizing output level, Q^* , is the point where the marginal production cost of the service equals the demand price (which is the marginal benefit to the organization of an additional unit of the good or service). A potential problem with this system, however, is that if the manager of the expense center in Figure 2 is motivated to maximize the size of the division (if only because compensation schedules tend to increase rewards for jobs with larger budgets and more people), the equilibrium output level will be significantly higher than Q^* , as indicated by the point “ Q ” in Figure 2. (For simplicity the figure assumes that the relevant measure of size in the objective function of the divisional manager is the quantity of output.)

The point Q is determined under the assumption that the budget office can estimate reasonably well the value to the entire organization of the division’s *total* output, but has limited knowledge of the costs and benefits of individual units of output. This means the budget office will not authorize a budget for the division that exceeds the division’s total value to the organization. Q will thus be the maximum quantity that can be produced by the division subject to the constraints that (1) its total budget does not exceed the total value of its output to the organization (the area under the demand curve) and (2) its total cost of producing the output (the area under the MC curve assuming no fixed per-period costs) does not exceed its budget.

6. Stern Stewart & Co., the firm directly responsible for refining the residual income measure into EVA, attempts to address this capital value problem by establishing a “bonus bank” system that ties current bonus payments to a kind of rolling three-year average payout of previous years’ bonus awards. But, although this effectively extends managers’ performance horizon from one to three years, this does not completely solve

the problem. Significant amounts of (publicly traded) stock in the entity managed can accomplish this, but this of course is often not cost-effective.

7. See W.A. Niskanen, Jr., “The Peculiar Economics of Bureaucracy,” *American Economic Review* 58 (May 1968).

(This is the quantity for which the area of triangle A equals the area of triangle B.)

Based on this analysis, the incentives of an expense center manager can thus be summarized as follows: The manager of the center wants to produce as much as possible, and the total value (ignoring the cost) of the center to the organization rises as the promised output rises. Therefore the budget increases with increases in the promised level of output. The result is that the budgets and output of expense centers tend to become larger than the size that would maximize the value of the organization. But there is a limit to such growth because the manager cannot produce more than Q with the maximum total budget the budget office is willing to give him for that promised level of output. And, in this fashion, this promised level becomes the equilibrium output.

Internal Chargeback Systems: Decentralizing Part of The Control Function

The tendency of an expense center to overproduce is exacerbated by the fact that the consumers of the center's product are not charged for the services they consume. Therefore, consumers have no incentives to compare the cost of the services they consume with the value of the services to them. In addition, if the budget office attempts to cut the center's budget, the center will be able to obtain support from the consumers of its output to oppose such cuts. The center director reacts to budget cuts by threatening cuts in the most highly valued rather than the marginal services and this also motivates the center's users to lobby against such cuts. The fact that the users of the service do not pay for the output they consume also means they will tend to demand services of too high a quality.

Consider a situation where the knowledge required to evaluate the performance of a division that provides services or product to other units of an organization is (1) specific (that is, costly to transfer among agents), (2) not easily observable from higher levels in the hierarchy, and (3) located among users of the division's output. In this situation it can be desirable to transfer some of the control function to the users of the division. This can be done by instituting a charge system in which the users pay for the output of the producing division. When consumers must pay for a good or service rather than receiving it at no cost, they have incentives to compare the benefits of the goods with the prices they must pay for them. This will cause them to consume less of the goods or services, thereby reducing the overconsumption problem caused by the expense center structure.

If a chargeback system is to be effective as a decentralized control mechanism, the users must also have decision rights that give them effective choices—for example, the right to purchase the good outside the firm, to produce it themselves, or to buy it from another division that has gone into competition to produce and supply the good internally. Given these

decision rights, a buying division has incentives to compare the quality and prices of the goods offered by the supplying division to that which they can obtain from other suppliers or by making it themselves. This constant evaluation will then be reflected in the buyer's decision to purchase or not to purchase from the supplying division. This right to choose to buy elsewhere provides strong incentives for the buying division to monitor the hard-to-assess qualities of the product of the supplying division, and it will be able to use its specific knowledge of those qualities in its monitoring.

In such a system, the higher levels in the hierarchy have effectively delegated much of the monitoring of the supplying division to its customers. This in turn means that the overall divisional monitoring function can be accomplished at higher levels in the hierarchy simply by measuring the profits of the producing division. In so doing, headquarters is freed from much of the task of monitoring the quality and quantity of output.

Internal chargeback systems can be used not only for expense centers, but with any of the performance measurement methods thus far described. In each case, there are benefits to be obtained by soliciting the help of the buyers of the division's output in the monitoring function. But there is at least one potentially important obstacle to implementing such chargeback systems—namely, the possession of “monopoly” power by the selling division or “monopsony” power by the buying division. Selling division managers who have monopoly powers within the organization will charge prices that are too high, while buying divisions with monopsony power could force sellers to furnish the product or service at prices that are too low. In either case, the resulting output of the organization will be less than the optimal level. For this reason, chargeback systems will be more efficient when neither the sellers or buyers have internal monopoly powers.

Choice of Performance Measure

The choice of a performance measure requires a theory that predicts when one performance measure will provide more reliable incentives to maximize value than another. Our goal here is the construction of a theory of the determinants of performance measurement that enables one to predict when a division will be organized as a profit center, cost center, investment center, revenue center, or expense center.

The decision to operate a division as an expense center is in large part a decision to control and monitor the division directly from higher in the hierarchy. This centralized control and monitoring option will be more attractive when it is easier to evaluate the performance of the division from higher levels of the hierarchy, and when it is difficult to decentralize the monitoring function to users of the output of the division. For example, it is sometimes difficult to identify a set of users who could be charged for the output of the unit. Such users must be individuals or business units whose combined valua-

The Locus of Uncertainty Problem

Unfortunately, there are pressures inside companies that tend to exacerbate the monopoly problem while at the same time substantially reducing incentives for users to make effective use of their specific knowledge about the quality and quantity of the output of the producing division. One source of such pressure is what we have labeled the “locus of uncertainty” problem.

Organizations that institute chargeback systems as part of a decentralized control mechanism commonly inhibit the functioning of those systems by constraining the choices of the customers of the internal seller through such devices as line budgets or “funny money” allocations that cannot be spent on anything other than the good or service in question. Computer services are a good example. It is common for computer funds in the budgets of buyers of a centralized internal computer supplier to be constrained for use in purchases from the central facility only. Since the funds allocated in such line budgets have zero opportunity cost to the managers, the managers’ purchase decisions do not reflect their assessment of the value of the service relative to other uses of the funds. This means the purchasing decisions of users do not reflect their evaluation of the quality and quantity of the services supplied by the central facility in comparison to that available from alternative suppliers or from their own production of the service. Thus, one of the major benefits from introduction of a chargeback system, the revelation of such specific knowledge possessed by users, is lost to the organization.

In every situation in which a chargeback system is used, there is an individual (whom we will call the “budget officer”) who must bear a good deal of uncertainty in order for the organization to receive the benefits of the chargeback system. This is the person responsible for the budgets of both the selling and buying divisions. The problem arises from the fact that, at the beginning of the year, the same monies allocated to the selling division for use in the production of the service must also be allocated to the buying divisions. If the buying divisions choose to spend the resources on goods and services other than those forecast by the selling division for its product, the monies will have in effect been spent

twice and the budget officer will experience a deficit.

If the evaluation mechanism faced by the budget officer is not flexible enough to allow for such deficits, the budget officer has incentives to collaborate with the supplying division to make the latter an effective monopoly by forbidding the expenditure of funds allocated for its product on anything else. This is accomplished by line budget allocations. Such constraints destroy much of the benefits of the chargeback system—indeed, they make it a sham.⁸ Under what circumstances, then, does the use of line budgets make sense? The centralized restriction of choices through line budgets is likely to be efficient mainly when problems in measuring the performance of users make it difficult to ensure that users are generally reflecting the value of the good to the entire organization in their decisions. In practice, however, such restrictions are also widely used when there is no benefit to the organization and when they generate considerable costs.

In many cases, it is the CFO or Controller of the organization who is the budget officer, and who therefore faces the locus of uncertainty problem generated when he or she is faced with allocating a given amount of money twice. Doing so requires risking the possibility that the budget will be overspent when buying division managers reject the overpriced or low-quality services of the selling division. There is no simple solution to this problem. The top management of the company must simply recognize that, in order to get the benefits of a chargeback system, the company must be prepared to run the risk of overspent budgets.

Of course, budget overruns can cause problems for security analysts, who are mostly interested in having their forecasts proved correct. In this case capital markets can drive CFOs and their companies to forgo the benefits of a decentralized monitoring system for the false security of a more certain, but far less efficient, management-by-the-numbers” system. To the extent division managers have better information about their businesses than headquarters, performance evaluation systems that put primary emphasis on meeting budgeted targets lead to “sandbagging” and other forms of “gaming” in the budget-setting process, and this in turn leads to even more inefficiencies and waste.

8. For a case dealing with the locus of uncertainty problem, see G.P. Baker and K.B. Monsler, “San Francisco Bay Consulting,” Case no. 9-195-196, Harvard Business School Publishing, 1995.

tion of the center’s output is likely to equal the value to the organization as a whole.

Take the case of the patent services group in a large pharmaceutical company. If the scientists in the lab are

given the decision rights over when to use patent services and are charged for them, the organization will likely consume too little of the service. On the other hand, because of the significant time lag—often as much as ten years—between

a patent-and-disclosure decision and the outcome of that decision, it would also be a mistake to give the decision rights to the manufacturing or marketing divisions. The managers of these units are unlikely to have the scientific expertise to keep up with the multitude of developments in the lab and to foresee the commercial applicability of that subset which should be protected with patents or disclosure. They will tend to focus their attention instead on the struggle to contain the usual day-to-day emergencies in the firm's current markets.

In short, it will be difficult in many organizations to decentralize the monitoring of such services. In these situations, the major alternative is to organize the supply of such services as an expense center and monitor its performance directly. It will tend to have all the problems of expense centers, but the costs of these problems might nevertheless be the lowest attainable among all alternative organizational structures.

In general, a cost center will be more desirable the lower is the cost of obtaining good information about:

- quantity;
- quality;
- correct output mix; and
- cost functions.

Profit centers will tend to be more desirable when the above costs are high and when:

- the correct revenues for the division are relatively easy to identify;

- there are few interdependencies in cost and demand functions between divisions; and
- there are no major internal monopoly problems.

In cases of business units that provide product or services mainly to other units inside the firm, profit centers will tend to work best when they are combined with a decision rights assignment that decentralizes part of the monitoring function of the center to its customers through a chargeback system. It is important, however, that such chargeback systems give those customers effective alternatives and thereby provide potential or actual competition for the profit center.

Investment centers and EVA will tend to be more desirable when the activity is capital-intensive and when it is difficult to identify optimal divisional asset investments from higher in the hierarchy.

MICHAEL JENSEN is the Jesse Isidor Straus Professor of Business Administration Emeritus at Harvard Business School. He was Managing Director of the Monitor Group's Organizational Strategy Practice 2000-2007 and is now Senior Advisor. Jensen is also the Co-Founder, Chairman and Co-CEO of Social Science Electronic Publishing, or "SSRN".

WILLIAM MECKLING was Dean of the University of Rochester's Simon School of Business.

Markets for Public Services: Strong Incentives in the Liberal State vs. Comprehensive Planning in the Welfare State?

Benito Arruñada¹

March 26, 2013

Abstract

This paper suggests that the partial but strong incentives that characterized privately valuable public services in the liberal state might be more effective than the comprehensive but weak incentives introduced by the ‘internal markets’ created when reforming the welfare state. The paper compares three organizational forms: (1) the bureaucratic expense center used to provide privately valuable services such as healthcare through the organizations created by the welfare state; (2) the internal markets introduced to reform them; and (3) the hybrid solutions that have been used by the liberal state since the 19th century to provide such privately valuable services. This comparison suggests that market forces may play a better role in organizing public services when they are limited to a few variables, which makes stronger incentives possible and, at the same time, reduces the need for extensive planning and supervisory staff.

JEL codes: H11, H42, H51, H52, K23

Keywords: public management, competition, public services, expense centers, bureaucracy

¹ Pompeu Fabra University and Barcelona GSE. E-mail: benito.arrunada@upf.edu. This work has greatly benefitted from exchanges with Jürgen Backhaus, José Manuel Freire, Fernando Méndez, Stephen Hansen, Cándido Paz-Ares, Celestino Pardo, Yannick Perez, Carlos Rodríguez Braun, Jérôme Sgard, Barry Weingast and participants at the EIU Workshop on “Legal Order, The State and Economic Development” and the CUNEF Workshop on “The Economics of Institutions and Organizations,” as well as the support of the Spanish Ministry of the Economy and Competitiveness, through grant ECO2011-29445. Usual disclaimers apply.

1. Introduction: The organization of public services

Most public services are now organized as “discretionary expense centers” (Kaplan and Atkinson 1989:531-33), so that each supplying unit receives a budget with which it has to provide services, usually unpaid, to other departments or end users. When the performance of such units is evaluated, if ever, this is usually based on subjective judgment and grossly incomplete indicators. Many of the decisions on what to produce, for whom and even how are usually centralized or subject to detailed decision-making procedures.²

In business firms, the use of expense centers is usually limited to areas in which it is difficult to measure their activities, in terms not only of efficiency but also of effectiveness, and in which it is not only difficult to link the resources used with the services rendered but also to define and evaluate the degree to which the objectives are achieved. Such characteristics can be found, especially, in central departments and in the actual general management of firms. The extent to which it is difficult to make such measurements usually depends on the type of resources and technology used and the services provided. The resources may be difficult to evaluate because of their high human capital content and because the costs are shared amongst or can be allocated to a wide range of users and products. In addition, many of the production processes may present a low degree of standardization and provide a wide range of services to other departments within the same organization or to external customers. Finally, many services are of an intangible nature so, though their quality is very important, it is difficult to evaluate because subjective elements are involved. It is often the case that only the users are aware of the quality and this knowledge is difficult to pass on to the person who has the authority to take corrective action. In the worst possible case, not even the users are aware of the real quality. This is especially so when activities are intensive in human capital so their quality can only be evaluated by qualified personnel.

In public services, technology also limits the possibility of measurement because many activities present the same characteristics. This is especially the case in the production of public goods for which there is no rivalry and exclusion is impossible (mainly those characteristic of the “Liberal State” such as national defense or the police). It is much less difficult to measure the services that constitute the core of the “welfare state,” such as health, education or pensions and other types of insurance. These are privately valuable services in which it is possible to exclude people from consuming them (except for their external effects, which are of varying importance and could be handled in many other ways apart from direct state provision). However, in spite of their private nature, they have often been provided free of charge by state agencies which have been financed by taxes unrelated to consumption. Moreover, in many cases it was also decided that discretionary expense centers should be

² Many expense centers (especially in public administration) have little formal discretion. However, they have an information advantage that grants them considerable informal discretion for allocating resources internally and manipulating central decisions on resource allocation and choice of procedure. In the case of professional services, formal regulation of production processes is usually very detailed but officials still have considerable discretion over their own quality and productivity.

responsible for their own organization, with a budget being allocated to the civil servant or the department rendering the service without paying them according to performance. These two patterns are in contrast to the system usually adopted previously for public services having private utility. Especially in the services that were re-organized during the liberal state of the 19th century (justice, registers, pharmacies and even, to some extent, education), activities were highly regulated but explicit fees were charged to users, and high-powered incentives were offered to suppliers. In some cases, professionals were even treated as public franchisees and compensated with the residual profit of their unit, after paying for any other resources for which they were responsible.

Recent reforms in public services, which have for decades been provided free of charge by expense centers, aim to recreate a market, introducing both some form of user choice and suppliers' incentives, in particular increasing competition between suppliers. It is not a question of "inter-organizational" competition between different organizations but rather of the "intra-organizational" competition among the divisions or departments of a single public organization. Such intra-organizational competition within the Administration is similar to that which often exists in large corporations. The problems it creates are not completely new but they are more complex (Arrow 1970:229). The large multi-divisional corporations have extensive experience in dealing with it and have developed many formulae for this purpose. Like the Public Administration, such corporations are constantly suffering from the tendency for administrative units to grow excessively and to provide services inefficiently to the other divisions. At root, such failures can only be remedied by changing the incentives of the two parties. Firstly, those of the internal users, making them pay for the services they acquire from the expense center and, secondly, those of the internal suppliers, taking from them their monopoly for the provision of such services. Sometimes, an attempt is made to create a sort of "internal market" within the firm. The general management of the organization then has to act as regulator of the competition that arises among its internal divisions.

This work examines the difficulties arising when elements of this type of internal competition are introduced in Public Administration and suggests that the role of market forces may be more productive if limited to a few variables. This seems to make stronger incentives possible while reducing the need for extensive planning and supervisory staff.

The rest of the paper proceeds as follows. Section 2 describes how public services commonly organized as budgetary or 'discretionary expense' centers, which do not charge users, tend to be too large and inefficient. Section 3 examines the possibilities and difficulties arising from the transformation of such expense centers into units that are to some extent subject to market forces, with users facing opportunity costs and suppliers being subject to competitive pressure. Section 4 analyzes solutions that have been used since the inception of the liberal state for organizing legal and judicial services in the Spanish public sector. These solutions involve strong incentives with little management, hinting that, instead of spending huge amounts of money on developing comprehensive internal markets with large management staffs but weak incentives, effective solutions may instead be based on relatively automatic management of strong incentives based on a few key variables of performance. Section 5 concludes.

2. The pathology of bureaucracy

Whether located in the public or the private sector, expense centers tend to be chronically inclined towards oversizing and overspending (Niskanen 1968, 1971).³ This is partly the result of circumstances associated with the incentives of producers and users and with the lack of information available to the person in charge of controlling the activity.

With regard to incentives, it usually happens that the remuneration, power and promotion prospects of those responsible for the expense center—and, in general, for all those working within it—increase when the center's budget increases, which means that they all have a common interest in enlarging it. Moreover, when users do not pay for the goods and services they acquire or when what they do pay does not vary in line with their level of consumption, most of them tend to demand them beyond the optimum level, that is, the level at which the cost would be the same as the marginal value for the organization or group. Obviously, this possibility is greater for goods or services that show positive utility.

The person in charge of allocating and controlling the use of resources finds it difficult to determine the budget for expense centers as there is no information indicating what is valuable or not in the consumption of the services rendered by the centers. In addition, even if it were possible to establish the budget for each expense center at its optimal level, it would still be difficult to evaluate the way in which the budgets are spent. For example, neither a deficit nor a surplus in an expense center indicates either wastage or efficiency because the center's products are not measured. And the managers of expense centers are interested in maintaining the informational disadvantage of the person responsible for the budget. They will therefore resist any policy aiming to measure the amount or quality of the services. For the same reason, they will tend to conceal the availability of idle resources and are unlikely to present a budget surplus as this might be interpreted as meaning they have excess resources. They are more likely to spend all available resources, even on purchases offering limited utility. In a similar way, users will tend to exaggerate the value of what they consume.

Because of these information asymmetries, the most that can be expected is that the problem will be contained within sustainable limits, using palliatives of doubtful efficiency, such as supervision and budgetary reviews. An incremental procedure is often adopted, focusing not on volume or resources but only on any new budget allocations or cutbacks. In practice, attempts to draw up new budgets starting from zero and reviewing all expenditure are both costly and inefficient. Since expense centers have this tendency to overspend, they often find their budgets cut when the organization they belong to is going through hard times (Neuman 1975), whether it is a company, a town council or a government. Such cutbacks entail certain problems. It is often not known in which areas a reduction of expenditure would be most effective or, even if this is known, it may be impossible to apply cutbacks in one area

³ This description simplifies the conduct of expense centers. For example, bureaucrats often focus more on leisure than on growth (Peacock, 1983), and may aim to increase not the whole budget available but only those items which provide greater personal utility (Dunleavy 1985, 1991). Moreover, it assumes a degree of autonomy that only exists in certain administrations, especially in the USA (Peters 1996:26-27). However, this criticism pays greater attention to formal autonomy and its short-term effects than to the actual autonomy which creates the information asymmetry and, especially, the implicit collusion between suppliers and users, as well as the long-term effects which are similar, irrespective of the degree of formal autonomy. A more serious limitation is that this model describes privately valuable services better than purely public services. The theory therefore tallies with the argument whereby bureaucracies tend to provide too many private goods and too few public goods (Breton 1974).

and not another. Crises often lead to overall cutbacks, perhaps to contain the costs generated by the actual budgetary battle, when their effects are anyway likely to be temporary. The problem is likely to reappear soon, because cutbacks do not solve the root of the problem which lies in the interests of the centers and their clients and is aggravated by the information difficulties suffered by the person responsible for the budgetary allocation. It may even occur that the cutbacks do not reduce expenditure as they come up against all sorts of restrictions and defense strategies on the part of those affected. The threat of dumping dead bodies in the manager's office has been used so often that it features as a category in some texts on hospital sociology.

3. Internal competition as a solution to the problems of bureaucracy

As an alternative to such palliative treatments, a more radical way of dealing with expense centers is to change participants' incentives, so that users incur some cost when they demand services and suppliers are paid for their performance. Ideally, this should provide an automatic control of quantity and quality of service, eliminating surplus demand and utilizing the information of users and suppliers on utility and cost, without any need to transfer this information to the unit responsible for the budget. Ideally, the system should play the informational role of the market (Hayek 1945).

3.1. Users and suppliers' incentives

The most basic requirement for motivating users so that their decisions help to achieve better allocation of resources is that their consumption of the service should incur a positive opportunity cost. They will then reveal their evaluation of the service and, providing other conditions are met, the supplier will be encouraged by the competition.⁴

There are many possibilities for reform. The most obvious way of generating an opportunity cost is to charge a real price for internal transactions. In addition, the payment of a real price usually provides an intense incentive. However, instead of real prices, nominal prices are often used with the corresponding accounting entries in the accounts of suppliers.⁵ In the public sector, such nominal charges usually take the form of vouchers with which users can purchase services from specific public or private suppliers. The use of nominal prices reduces user motivation but will not necessarily destroy it altogether, provided that such

⁴ It is debatable if paying internal prices without any freedom of choice regarding either the amount of services or their price also generates incentives for control, as argued by Zimmerman (1979).

⁵ Users of many public sector services are not departments but final consumers or citizens. Therefore, there is no need to set up artificial mechanisms for performance evaluation and remuneration.

nominal prices reflect real opportunity costs, as when users are able to obtain other services with these resources or, at least, the same services but from different suppliers.

Under any circumstances, in order to be an efficient control mechanism, the supplier should be in a situation of competition or, at least, users should be able to use their resources for other purposes. If, for example, each department in a university is allocated 100,000 USD to be spent only on photocopying, such departments would not have much motivation for controlling a single internal supplier. If, however, they are allowed to spend this amount on other things, they would implicitly be controlling the supplier by moving their demand elsewhere even though, by so doing, they would be complicating the budget problem.

Regarding suppliers, the most important general options are in theory represented by the different possibilities of divisional organization of expense centers in the form of cost centers, revenue centers, profit centers and investment centers and even franchises, the latter being a hybrid formula between organization and market. As decisions are increasingly delegated to the center in question, the performance indicator will have to be increasingly global (Kaplan and Atkinson 1989:529-33). This means that the person in charge of a cost center usually decides how resources should be used but not how many should be produced nor their quality, and this person is evaluated by some sort of indicator of production cost. In revenue centers, those in charge are evaluated by turnover or revenue, and a decision is freely taken on the amounts sold or the selling price, but not on both variables. A profit center manager must maximize some type of divisional accounting profit and is usually free to decide not only how resources should be used but also about production, quality and product prices. In investment centers, in addition to the attributes of profit centers, there is also freedom and responsibility regarding the use of larger amounts of resources as these include the capital used by the center. Finally, a franchise provides a hybrid solution in which the franchisee owns a large proportion of the assets and is paid from the profits of the local outlet.

The remuneration of whoever is in charge of such a center has to be linked to some sort of performance indicator so that they show interest in using their resources optimally. For example, the director of a school that has managed to raise demand for places or to improve students' academic performance could be rewarded with an annual bonus or promotion. And vice versa, if the performance indicators drop, not only would the director see a drop in income and fewer chances of promotion but the school might lose its independence or a substitute director might be called in to turn it around.

3.2. *The nature and costs of control*

In theory, there are many possibilities for this type of transformation of expense centers but the essential characteristics are defined by just two variables: freedom and responsibility. Freedom refers to the degree of discretion that participants are given in their decisions. Responsibility refers to the mechanisms used to evaluate and compensate their performance. Together, these give rise to a wide range of possibilities from which reformers must choose (Table 1).⁶ Freedom is introduced by redistributing decision rights: decisions that were

⁶ This analysis of the elements of organizational control is based on Arrow (1964) and, especially, Jensen and Meckling (1995). Arrow formulates organizational control as the interaction of operating rules and enforcement.

previously centralized are delegated to users and suppliers. Responsibility, which before hinged mainly on hierarchical or vertical control that aimed to evaluate compliance (rather than performance), starts to be based on some degree of horizontal and mutual control: that exerted by users on suppliers through their purchase decisions and by suppliers on users through their pricing policy.

Table 1. Design dimensions: participants' freedom and responsibility

	<i>Freedom: Decision-making rights of participants</i>	<i>Responsibility: Performance evaluation and incentives</i>
Users	Choice of supplier: <ul style="list-style-type: none"> - Where: inside or outside the organization - By whom: the users, a representative, a gatekeeper 	Opportunity costs: <ul style="list-style-type: none"> - Shadow invoice - Voucher - Price - Co-payment
Suppliers	Discretion for: <ul style="list-style-type: none"> - Organizing the activity - Transacting internally - Transacting externally 	Divisionalization: <ul style="list-style-type: none"> - Responsibility centers (for profit, cost, investment) - Franchised administration (units' managers hold property rights, hire employees and are paid with the unit's profit) Individual compensation function: <ul style="list-style-type: none"> - Pay for performance - Professional career - Units' profits (in franchised administration)

The existence of such horizontal control does not make vertical control less necessary, however, but it does transform its nature: it must ensure that mutual control between users and suppliers functions correctly so that they do not behave selfishly and their interaction is socially beneficial. The agency responsible for both users and suppliers must preclude that they use their freedom to serve only their own goals, sacrificing the common interest. Given that users and suppliers now enjoy more discretion and stronger incentives, vertical control must, in fact, be more effective (Milgrom and Roberts 1992:226-28).

Ideally, reform will re-create market functioning (which is why people often talk about "internal markets"[Enthoven 1991]), so will introduce competition, both among suppliers and among the ways in which users might allocate their resources. It might even go so far as to change the nature of the supplying unit, converting it into a profit center or even a franchise. However, such a radical solution usually requires considerable investments and expenditure for planning and managing the whole process (Arruñada and Hansen 2012). Its effectiveness, therefore, is doubtful. In fact, in order to supervise the reform and manage the internal market, a giant planning apparatus is often set up, a "Gosplan" which, in turn, has all the characteristics and defects of an expense centre (Arruñada 1997). An internal market is only a market in name, as there are no property rights and all prices are administered (Hayek 1945).

Jensen and Meckling express these operating rules as the allocation of decision rights (freedom) and distinguish two phases in enforcement (responsibility)—evaluation and compensation for performance.

4. The organization of privately valuable services in the liberal state

An alternative, seemingly more modest, solution is to apply just some market mechanisms, or just in some dimensions, but more forcefully. The patterns followed by the traditional organization of public services in the liberal state were often of this type, combining partial discretion by decision makers (limited, e.g., to certain dimensions such as the choice of supplier by the user or the tenure-based choice of specific jobs by professional civil servants) with powerful incentives (based, e.g., on user fees and pay for performance, a solution commonly found in the organization of notaries, registrars, judicial clerks or family doctors).

In order to illustrate the arguments, I review below three solutions adopted in the 19th century for judicial and quasi-judicial services in the Spanish Public Administration: notaries, registries and courts. These solutions differ drastically from both standard expense centers and internal markets. In contrast to the standard expense center, services are financed with user fees and at least some suppliers are paid for performance. In contrast with the internal market, incentives are much stronger, as users pay real money and suppliers are paid real bonuses. However, they are characterized by the limited size of the planning or supervisory agency. Instead of aiming for a complete artificial market, too costly to achieve, these admittedly suboptimal—in Simon’s (1956) terms “satisficing”—solutions seem to provide a better alignment of a few key dimensions of participants’ behavior.⁷

Strong user fees. Court users paid fees which financed a substantial proportion of costs. Elimination of these judicial rates in 1986 put an end to this system. This resulted in congestion, rationing, frivolous litigation, capacity increases and greater delays in civil than in criminal cases (Pastor, 1993). User fees were reintroduced for commercial cases in 2002 and generalized in 2012 (Gómez, Celentani and Ganuza 2012). In contrast, both notaries and registrars are paid explicit, regulated fees by one of the parties, and these fees finance services fully. Both notaries and registrars have shown considerable flexibility in adapting to drastic swings in market demand, caused by the real estate bubble in the period 2000-2007 and the adaptation of corporation law a decade earlier.

Limited user choice. Free choice of supplier is allowed only for essentially private services such as notaries’ services (Arruñada 1996), but not for registries and courts. Since the notary mainly serves the parties to the contract, it makes sense for them to freely choose the notary. It is also understandable that notaries tend to be flexible regarding the wishes of the parties, interpreting legal restrictions in the way that best suits them. Conversely, for courts and registries, freedom of choice would endanger their impartiality and their basic function of protecting third parties: e.g., free choice of land registry would not protect parties such as future land purchasers, who are unknown at the time of choosing (Arruñada 2003).

Strong incentives to suppliers. Incentives to notaries and registrars are as strong as they can be, as both notary’s offices and registers function as public franchisees. Each notary and each registrar is responsible for one office, hires their own employees and resources, and is paid (together with top employees) with the residual profit of the office, after the less professional staff and after any office costs have been paid. Conversely, judges are paid a

⁷ Arruñada and Hansen (2012) compare two of these solutions in depth.

fixed salary with substantial increases linked to tenure and promotion, a typical characteristic of judicial careers (Posner 1995). However, in a system that remained in place in many courts until the 1980s, though it was being phased-out since 1947, courts' clerks were paid substantial bonuses linked to processed cases.⁸ This motivated paperwork productivity without damaging the quality of court decisions. Court clerks controlled judges' productivity because, if the latter did not deal with cases expeditiously, they did not receive the full variable remuneration; and judges (who were often worse paid than the clerks so felt a degree of rivalry that was not always negative) controlled the quality of the office's administrative work which, since it tended to be done fast, was not always as it should be. In all cases, the key elements of quality control were the deferred nature of remuneration and personal liability in professional decisions. This latter aspect is specially important for registrars, who are subject to a standard of strict (i.e., non-negligent) liability.

Role of automatic control. Instead of specialized control by a supervisory bureaucracy, incentives are arranged in a way that favors mutual control by participants. Not only users control suppliers and suppliers control users but complementary suppliers control each other (e.g., registries control notaries, and vice versa, and, similarly, court clerks and judges control each other). Such controls between complementary suppliers are enhanced by using different compensation functions. For instance, both notaries and registrars are paid with net user fees; however, notaries are chosen by users and compete with each other while registrars enjoy territorial monopolies. Consequently, notaries strive to satisfy their clients while registrars strive for legal quality and the protection of third parties different from notaries' clients. Similarly, court clerks were paid a variable fee for performance, motivating them to speed up case paperwork while judges were paid a fixed salary and worried about possible appeals that might damage their reputation and chances of promotion. In both cases, opposite compensation functions create tension between complementary suppliers, providing some degree of automatic control as well as two sources of competitive information for regulators. A similar formula was used in the regulation, administrative appeal and inspection of notary's offices and registries by concentrating these supervisory tasks in the hands of a tiny and specialized body of civil servants at the General Directorate for Registries and Notarial Offices. These civil servants were paid a fixed salary below the variable compensation of the notaries and registrars they were inspecting, so they tended to be stern about any slackness.

5. Concluding remarks

This paper compares three ways of organizing the provision of privately valuable public services: budgetary bureaucracy, comprehensive internal markets and traditional hybrid solutions based on partial market incentives and first used by the liberal state in the 19th century. These hybrid solutions, which have survived for many decades and produced relatively effective outcomes, contrast starkly not only with those of bureaucracy but also with those found in modern internal markets. Whereas internal markets strive to develop

⁸ In 2004, to deal with court congestion, the Government introduced short-term performance targets, which caused an increase in average measured productivity and reduced the productivity of top performing judges (Bagues and Esteve-Volart 2010). They were opposed by judges' associations and may have damaged professional morale.

comprehensive measures of performance but provide weak incentives, these liberal state solutions rely on partial measures of performance but provide strong incentives. Moreover, instead of requiring large staffs to manage suppliers and their interaction with users in the internal market, such hybrid solutions work in a regime of “automatic management” and are therefore frugal in their use of planning and supervision resources.

The purpose of this paper is simply to call attention to the difference between two organizational solutions which, for different reasons, are considered close to the market: on the one hand, “internal market” solutions that claim to be, and are usually seen as, attempts to introduce competitive forces in bureaucratic public services, to the extent that they are sometimes considered the result of “neoliberal” policies; and, on the other, the solutions adopted by the liberal state in the 19th century to provide privately valuable services.

The paper also suggests the need for further analysis based on wider evidence and considering solutions in other services and countries (e.g., examining traditional hybrid solutions adopted for healthcare in those countries which have opted for private provision of publicly-financed healthcare). It also poses many other questions. For instance, it would be of interest to identify which factors determined the choice between the expense center and the hybrid solutions as the organizational form for providing the typical services of the welfare state. If still present today, these factors might inform us as to the true nature and prospects of internal market initiatives. In particular, they might tell us if there is any prospect of such internal markets developing into real markets instead of just remaining giant bureaucracies.

6. References

- Arrow, K. J. (1964), “Control in Large Organizations,” *Management Science*, 10(3), 397-408.
- Arruñada, B. (1996), “The Economics of Notaries,” *European Journal of Law and Economics*, 3(1), 5-37.
- Arruñada, B. (1997), “Internal Markets in the Reform of the Spanish NHS: A Comment on the Planners’ Latest Fantasy,” in J. G. Backhaus, (ed.), *Essays in Social Security and Taxation*, Metropolis, Marburg, 429-44.
- Arruñada, B. (2003), “Property Enforcement as Organized Consent,” *Journal of Law, Economics, and Organization*, 19(2), 401-44
- Arruñada, B. (2012), *Institutional Foundations of Impersonal Exchange: The Theory and Policy of Contractual Registries*, University of Chicago Press, Chicago.
- Arruñada, B., and S. Hansen (2012), “Organizational Architecture and Public Sector Reform,” Universitat Pompeu Fabra, Department of Economics and Business, July.
- Bagues, M., and B. Esteve-Volart (2010), “Performance Pay and Judicial Production: Evidence from Spain,” *Working Paper*, June (<http://www.manuelbagues.com/bagues%20&%20esteve-volart%20-%20judges.pdf>, visited March 13, 2013).
- Breton, A. (1974), *An Economic Theory of Representative Government*, Aldine, Chicago.
- Dunleavy, P. (1985), “Bureaucrats, Budgets and the Growth of the State,” *British Journal of Political Science*, 15, 299-328.

- Dunleavy, P. (1991), *Democracy, Bureaucracy and Public Choice*, Harvester Wheatsheaf, Brighton.
- Enthoven, A. C. (1991), "Internal Market Reform of the British National Health Service," *Health Affairs* 10(3), 60-70.
- Gómez, F., M. Celentani and J.J. Ganuza (2012), "A vueltas con las tasas judiciales," *Nada es gratis*, December 18 (<http://www.fedeablogs.net/economia/?p=26980>, visited March 13, 2013).
- Hayek, F.A. (1945), "The Use of Knowledge in Society," *American Economic Review*, 35(4), 519-30.
- Jensen, M. C., and W. H. Meckling (1995), "Specific and General Knowledge, and Organizational Structure," *Journal of Applied Corporate Finance*, 8(2), 4-18.
- Kaplan, R. S., and A. A. Atkinson (1989), *Advanced Management Accounting*, 2nd ed., Prentice-Hall, Englewood Cliffs, NJ.
- Milgrom, P., and J. Roberts (1992), *Economics, Organization and Management*, Prentice-Hall, Englewood Cliffs, NJ.
- Neuman, J. L. (1975), "Make Overhead Cuts that Last," *Harvard Business Review*, May-June, 116-26.
- Niskanen, W. A. (1968), "Nonmarket Decision Making: The Peculiar Economics of Bureaucracy," *American Economic Review*, 58(1), 293-305.
- Niskanen, W. A. (1971), *Bureaucracy and Representative Government*, Aldine Atherton, Chicago.
- Pastor Prieto, S. (1993), *¡Ah de la justicia! Política judicial y economía*, Madrid, Cívitas.
- Peacock, A. T. (1983), "Public X-Inefficiency: Informational and Institutional Constraints," in H. Hanusch, (ed.), *Anatomy of Government Deficiencies*, Springer, Berlin, 125-37.
- Peters, B. G. (1996), *The Future of Governing: Four Emerging Models*, University Press of Kansas, Lawrence.
- Posner, R. A. (1995), "What Do Judges Maximize", in *Overcoming Law*, Harvard University Press, Cambridge, MA, 109-44.
- Simon, H. A. (1956), "Rational Choice and the Structure of the Environment," *Psychological Review*, 63(2), 129-38.
- Zimmerman, J. L. (1979), "The Costs and Benefits of Cost Allocations," *Accounting Review*, 54(3), 504-21.

4. Other topics — Due: week 4

References:

- Pinker, Steven (1997), “Hotheads,” chapter 6 of *How the Mind Works*, Norton, New York, 363-424.
- Pinker, Steven (1997), “Family Values,” chapter 7 of *How the Mind Works*, Norton, New York, 425-520.
- Fama, Eugene F., and Michael. C. Jensen (1983), “Agency Problems and Residual Claims,” *Journal of Law and Economics*, 26(2), 327-49.
- Hayek, Friedrich A. (1988), “The Weasel Word ‘Social’,” in *The Fatal Conceit: The Errors of Socialism*, Routledge, London, 114-19.
- Friedman, Milton (1970), “The Social Responsibility of Business is to Increase its Profits,” *New York Times Magazine*, September 13.
- Echebarria, Koldo (2022), “La universidad, por detrás de la tecnología”, *LinkedIn*, January, 20, [link](#).
- Felten, Ed, Manav Raj, and Robert Seamans (2023), “How will Language Modelers like ChatGPT Affect Occupations and Industries?,” 20 March 2023, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4375268.

HOTHEADS

On March 13, 1996, Thomas Hamilton walked into an elementary school in Dunblane, Scotland, carrying two revolvers and two semiautomatic pistols. After wounding staff members who tried to tackle him, he ran to the gymnasium, where a kindergarten class was playing. There he shot twenty-eight children, sixteen fatally, and killed their teacher before turning the gun on himself. "Evil visited us yesterday, and we don't know why," said the school's headmaster the next day. "We don't understand it and I don't think we ever will."

We probably never will understand what made Hamilton commit his vile final acts. But the report of pointless revenge by an embittered loner is disturbingly familiar. Hamilton was a suspected pedophile who had been forced to resign as a Scout leader and then formed his own youth groups so he could continue working with boys. One group held its meetings in the Dunblane school's gymnasium until school officials, responding to parents' complaints about his odd behavior, forced him out. Hamilton was the target of ridicule and gossip, and was known in the area, undoubtedly for good reasons, as "Mr. Creepy." Days before his rampage he had sent letters to the media and to Queen Elizabeth defending his reputation and pleading for reinstatement in the scouting movement.

The Dunblane tragedy was particularly shocking because no one thought it could happen there. Dunblane is an idyllic, close-knit village where serious crime was unknown. It is far from America, land of the wackos, where there are as many guns as people and where murderous rampages by disgruntled postal workers are so common (a dozen inci-

dents in a dozen years) that a slang term for losing one's temper is "going postal." But running amok is not unique to America, to Western nations, or even to modern societies. *Amok* is a Malay word for the homicidal sprees occasionally undertaken by lonely Indochinese men who have suffered a loss of love, a loss of money, or a loss of face. The syndrome has been described in a culture even more remote from the West: the stone-age foragers of Papua New Guinea.

The amok man is patently out of his mind, an automaton oblivious to his surroundings and unreachable by appeals or threats. But his rampage is preceded by lengthy brooding over failure, and is carefully planned as a means of deliverance from an unbearable situation. The amok state is chillingly cognitive. It is triggered not by a stimulus, not by a tumor, not by a random spurt of brain chemicals, but by an idea. The idea is so standard that the following summary of the amok mind-set, composed in 1968 by a psychiatrist who had interviewed seven hospitalized amoks in Papua New Guinea, is an apt description of the thoughts of mass murderers continents and decades away:

I am not an important or "big man." I possess only my personal sense of dignity. My life has been reduced to nothing by an intolerable insult. Therefore, I have nothing to lose except my life, which is nothing, so I trade my life for yours, as your life is favoured. The exchange is in my favour, so I shall not only kill you, but I shall kill many of you, and at the same time rehabilitate myself in the eyes of the group of which I am a member, even though I might be killed in the process.

The amok syndrome is an extreme instance of the puzzle of the human emotions. Exotic at first glance, upon scrutiny they turn out to be universal; quintessentially irrational, they are tightly interwoven with abstract thought and have a cold logic of their own.

UNIVERSAL PASSION

A familiar tactic for flaunting one's worldhness is to inform listeners that some culture lacks an emotion we have or has an emotion we lack. Allegedly the Utku-Inuit Eskimos have no word for anger and do not feel the emotion. Tahitians supposedly do not recognize guilt, sadness, longing, or loneliness; they describe what we would call grief as fatigue, sickness, or

bodily distress. Spartan mothers were said to smile upon hearing that their sons died in combat. In Latin cultures, machismo reigns, whereas the Japanese are driven by a fear of shaming the family. In interviews on language I have been asked, Who but the Jews would have a word, *naches*, for luminous pride in a child's accomplishments? And does it not say something profound about the Teutonic psyche that the German language has the word *Schadenfreude*, pleasure in another's misfortunes?

Cultures surely differ in how often their members express, talk about, and act on various emotions. But that says nothing about what their people feel. The evidence suggests that the emotions of all normal members of our species are played on the same keyboard.

The most accessible signs of emotions are candid facial expressions. In preparing *The Expression of the Emotions in Man and Animals*, Darwin circulated a questionnaire to people who interacted with aboriginal populations on five continents, including populations that had had little contact with Europeans. Urging them to answer in detail and from observation rather than memory, Darwin asked how the natives expressed astonishment, shame, indignation, concentration, grief, good spirits, contempt, obstinacy, disgust, fear, resignation, sulkiness, guilt, slyness, jealousy, and "yes" and "no." For example:

(5.) When in low spirits, are the corners of the mouth depressed, and the inner corner of the eyebrows raised by that muscle which the French call the "Grief muscle"? The eyebrow in this state becomes slightly oblique, with a little swelling at the inner end; and the forehead is transversely wrinkled in the middle part, but not across the whole breadth, as when the eyebrows are raised in surprise.

Darwin summed up the responses: "The same state of mind is expressed throughout the world with remarkable uniformity; and this fact is in itself interesting as evidence of the close similarity in bodily structure and mental disposition of all the races of mankind."

Though Darwin may have biased his informants with leading questions, contemporary research has borne out his conclusion. When the psychologist Paul Ekman began to study emotions in the 1960s, facial expressions were thought to be arbitrary signs that the infant learns when its random grimaces are rewarded and punished. If expressions appeared universal, it was thought, that was because Western models had become universal; no culture was beyond the reach of John Wayne and Charlie Chaplin. Ekman assembled photographs of people express-

ing six emotions. He showed them to people from many cultures, including the isolated Fore foragers of Papua New Guinea, and asked them to label the emotion or make up a story about what the person had gone through. Everyone recognized happiness, sadness, anger, fear, disgust, and surprise. For example, a Fore subject said that the American showing fear in the photograph must have just seen a boar. Reversing the procedure, Ekman photographed his Fore informants as they acted out scenarios such as "Your friend has come and you are happy," "Your child has died," "You are angry and about to fight," and "You see a dead pig that has been lying there for a long time." The expressions in the photographs are unmistakable.

When Ekman began to present his findings at a meeting of anthropologists in the late 1960s, he met with outrage. One prominent anthropologist rose from the audience shouting that Ekman should not be allowed to continue to speak because his claims were fascist. On another occasion an African American activist called him a racist for saying that black facial expressions were no different from white ones. Ekman was bewildered because he had thought that if the work had any political moral it was unity and brotherhood. In any case, the conclusions have been replicated and are now widely accepted in some form (though there are controversies over which expressions belong on the universal list, how much context is needed to interpret them, and how reflexively they are tied to each emotion). And another observation by Darwin has been corroborated: children who are blind and deaf from birth display virtually the full gamut of emotions on their faces.

Why, then, do so many people think that emotions differ from culture to culture? Their evidence is much more indirect than Darwin's informants and Ekman's experiments. It comes from two sources that cannot be trusted at all as readouts of people's minds: their language and their opinions.

The common remark that a language does or doesn't have a word for an emotion means little. In *The Language Instinct* I argued that the influence of language on thought has been exaggerated, and that is all the more true for the influence of language on feeling. Whether a language appears to have a word for an emotion depends on the skill of the translator and on quirks of the language's grammar and history. A language accumulates a large vocabulary, including words for emotions, when it has had influential wordsmiths, contact with other languages, rules for forming new words out of old ones, and widespread literacy, which

allows new coinages to become epidemic. When a language has not had these stimulants, people describe how they feel with circumlocutions, metaphors, metonyms, and synecdoches. When a Tahitian woman says, "My husband died and I feel sick," her emotional state is hardly mysterious; we can bet she is not complaining about acid indigestion. Even a language with a copious vocabulary has words for only a fraction of emotional experience. The author G. K. Chesterton wrote,

Man knows that there are in the soul tints more bewildering, more numberless, and more nameless than the colours of an autumn forest; . . . Yet he seriously believes that these things can every one of them, in all their tones and semitones, in all their blends and unions, be accurately represented by an arbitrary system of grunts and squeals. He believes that an ordinary civilized stockbroker can really produce out of his own inside noises which denote all the mysteries of memory and all the agonies of desire.

When English-speakers hear the word *Schadenfreude* for the first time, their reaction is not, "Let me see . . . Pleasure in another's misfortunes . . . What could that possibly be? I cannot grasp the concept; my language and culture have not provided me with such a category." Their reaction is, "You mean there's a *word* for it? Cool!" That is surely what went through the minds of the writers who introduced *Schadenfreude* into written English a century ago. New emotion words catch on quickly, without tortuous definitions; they come from other languages (*enmu*, *angst*, *naches*, *amok*), from subcultures such as those of musicians and drug addicts (*blues*, *funk*, *juiced*, *wasted*, *rush*, *high*, *freaked out*), and from general slang (*pissed*, *bummed*, *grossed out*, *blown away*). I have never heard a foreign emotion word whose meaning was not instantly recognizable.

People's emotions are so alike that it takes a philosopher to craft a genuinely alien one. In an essay called "Mad Pain and Martian Pain," David Lewis defines mad pain as follows:

There might be a strange man who sometimes feels pain, just as we do, but whose pain differs greatly from ours in its causes and effects. Our pain is typically caused by cuts, burns, pressure, and the like; his is caused by moderate exercise on an empty stomach. Our pain is generally distracting; his turns his mind to mathematics, facilitating concentration on that but distracting him from anything else. Intense pain has no ten-

dency whatever to cause him to groan or writhe, but does cause him to cross his legs and snap his fingers. He is not in the least motivated to prevent pain or to get rid of it.

Have anthropologists discovered a people that feels mad pain or something equally weird? It might seem that way if you look only at stimulus and response. The anthropologist Richard Shweder points out, "It is a trivial exercise for any anthropologist to generate long lists of antecedent events (ingesting cow urine, eating chicken five days after your father dies, kissing the genitals of an infant boy, being complimented about your pregnancy, caning a child, touching someone's foot or shoulder, being addressed by your first name by your wife, ad infinitum) about which the emotional judgments of a Western observer would not correspond to the native's evaluative response." True enough, but if you look a bit deeper and ask how people *categorize* these stimuli, the emotions elicited by the categories make you feel at home. To us, cow urine is a contaminant and cow mammary secretions are a nutrient; in another culture, the categories may be reversed, but we all feel disgust for contaminants. To us, being addressed by your first name by a spouse is not disrespectful, but being addressed by your first name by a stranger might be, and being addressed by your religion by your spouse might be, too. In all the cases, disrespect triggers anger.

But what about the claims of native informants that they just don't have one of our emotions? Do our emotions seem like mad pain to them? Probably not. The Utku-Inuits' claim that they do not feel anger is belied by their behavior: they recognize anger in foreigners, beat their dogs to discipline them, squeeze their children painfully hard, and occasionally get "heated up." Margaret Mead disseminated the incredible claim that Samoans have no passions—no anger between parents and children or between a cuckold and a seducer, no revenge, no lasting love or bereavement, no maternal caring, no tension about sex, no adolescent turmoil. Derek Freeman and other anthropologists found that Samoan society in fact had widespread adolescent resentment and delinquency, a cult of virginity, frequent rape, reprisals by the rape victim's family, frigidity, harsh punishment of children, sexual jealousy, and strong religious feeling.

We should not be surprised at these discrepancies. The anthropologist Renato Rosaldo has noted, "A traditional anthropological description is like a book of etiquette. What you get isn't so much the deep cultural wisdom as the cultural clichés, the wisdom of Polonius, conventions in

the trivial rather than the informing sense. It may tell you the official rules, but it won't tell you how life is lived." Emotions, in particular, are often regulated by the official rules, because they are assertions of a person's interests. To me it's a confession of my innermost feelings, but to you it's bitching and moaning, and you may very well tell me to put a lid on it. And to those in power, other people's emotions are even more annoying—they lead to nuisances such as women wanting men as husbands and sons rather than as cannon fodder, men fighting each other when they could be fighting the enemy, and children falling in love with a soulmate instead of accepting a betrothed who cements an important deal. Many societies deal with these nuisances by trying to regulate emotions and spreading the disinformation that they don't exist.

Ekman has shown that cultures differ the most in how the emotions are expressed in public. He secretly filmed the expressions of American and Japanese students as they watched gruesome footage of a primitive puberty rite. (Emotion researchers have extensive collections of gross-out material.) If a white-coated experimenter was in the room interviewing them, the Japanese students smiled politely during scenes that made the Americans recoil in horror. But when the subjects were alone, the Japanese and American faces were equally horrified.

FEELING MACHINES

The Romantic movement in philosophy, literature, and art began about two hundred years ago, and since then the emotions and the intellect have been assigned to different realms. The emotions come from nature and live in the body. They are hot, irrational impulses and intuitions, which follow the imperatives of biology. The intellect comes from civilization and lives in the mind. It is a cool deliberator that follows the interests of self and society by keeping the emotions in check. Romantics believe that the emotions are the source of wisdom, innocence, authenticity, and creativity, and should not be repressed by individuals or society. Often Romantics acknowledge a dark side, the price we must pay for artistic greatness. When the antihero in Anthony Burgess' *A Clockwork Orange* has his violent impulses conditioned out of him, he loses his taste for Beethoven. Romanticism dominates contemporary American popular culture, as in the Dionysian ethos of rock music, the

pop psychology imperative to get in touch with your feelings, and the Hollywood formulas about wise simpletons and about uptight yuppies taking a walk on the wild side.

Most scientists tacitly accept the premises of Romanticism even when they disagree with its morals. The irrational emotions and the repressing intellect keep reappearing in scientific guises: the id and the superego, biological drives and cultural norms, the right hemisphere and the left hemisphere, the limbic system and the cerebral cortex, the evolutionary baggage of our animal ancestors and the general intelligence that propelled us to civilization.

In this chapter I present a distinctly unromantic theory of the emotions. It combines the computational theory of mind, which says that the lifeblood of the psyche is information rather than energy, with the modern theory of evolution, which calls for reverse-engineering the complex design of biological systems. I will show that the emotions are adaptations, well-engineered software modules that work in harmony with the intellect and are indispensable to the functioning of the whole mind. The problem with the emotions is not that they are untamed forces or vestiges of our animal past; it is that they were designed to propagate copies of the genes that built them rather than to promote happiness, wisdom, or moral values. We often call an act "emotional" when it is harmful to the social group, damaging to the actor's happiness in the long run, uncontrollable and impervious to persuasion, or a product of self-delusion. Sad to say, these outcomes are not malfunctions but precisely what we would expect from well-engineered emotions.

The emotions are another part of the mind that has been prematurely written off as nonadaptive baggage. The neuroscientist Paul MacLean took the Romantic doctrine of the emotions and translated it into a famous but incorrect theory known as the Triune Brain. He described the human cerebrum as an evolutionary palimpsest of three layers. At the bottom are the basal ganglia or Reptilian Brain, the seat of the primitive and selfish emotions driving the "Four Fs": feeding, fighting, fleeing, and sexual behavior. Grafted onto it is the limbic system or Primitive Mammalian Brain, which is dedicated to the kinder, gentler, social emotions, like those behind parenting. Wrapped around that is

the Modern Mammalian Brain, the neocortex that grew wild in human evolution and that houses the intellect. The belief that the emotions are animal legacies is also familiar from pop ethology documentaries in which snarling baboons segue into rioting soccer hooligans as the voice-over frets about whether we will rise above our instincts and stave off nuclear doom.

One problem for the triune theory is that the forces of evolution do not just heap layers on an unchanged foundation. Natural selection has to work with what is already around, but it can *modify* what it finds. Most parts of the human body came from ancient mammals and before them ancient reptiles, but the parts were heavily modified to fit features of the human lifestyle, such as upright posture. Though our bodies carry vestiges of the past, they have few parts that were unmodifiable and adapted only to the needs of older species. Even the appendix is currently put to use, by the immune system. The circuitry for the emotions was not left untouched, either.

Admittedly, some traits are so much a part of the architectural plan of an organism that selection is powerless to tinker with them. Might the software for the emotions be burned so deeply into the brain that organisms are condemned to feel as their remote ancestors did? The evidence says no; the emotions are easy to reprogram. Emotional repertoires vary wildly among animals depending on their species, sex, and age. Within the mammals, we find the lion and the lamb. Even within dogs (a single species), a few millennia of selective breeding have given us pit bulls and Saint Bernards. The genus closest to ours embraces common chimpanzees, in which gangs of males massacre rival gangs and females can murder one another's babies, and the pygmy chimpanzees (bonobos), whose philosophy is "Make love not war." Of course, some reactions are widely shared across species—say, panic when one is confined—but the reactions may have been retained because they are adaptive for everyone. Natural selection may not have had complete freedom to reprogram the emotions, but it had a lot.

And the human cerebral cortex does not ride piggyback on an ancient limbic system, or serve as the terminus of a processing stream beginning there. The systems work in tandem, integrated by many two-way connections. The amygdala, an almond-shaped organ buried in each temporal lobe, houses the main circuits that color our experience with emotions. It receives not just simple signals (such as of loud noises) from the lower stations of the brain, but abstract, complex information from the brain's

highest centers. The amygdala in turn sends signals to virtually every other part of the brain, including the decision-making circuitry of the frontal lobes.

The anatomy mirrors the psychology. Emotion is not just running away from a bear. It can be set off by the most sophisticated information processing the mind is capable of, such as reading a Dear John letter or coming home to find an ambulance in the driveway. And the emotions help to connive intricate plots for escape, revenge, ambition, and courtship. As Samuel Johnson wrote, "Depend upon it, sir, when a man knows he is to be hanged in a fortnight, it concentrates his mind wonderfully."

The first step in reverse-engineering the emotions is try to imagine what a mind would be like without them. Supposedly Mr. Spock, the Vulcan mastermind, didn't have emotions (except for occasional intrusions from his human side and a seven-year itch that drove him back to Vulcan to spawn). But Spock's emotionlessness really just amounted to his being in control, not losing his head, coolly voicing unpleasant truths, and so on. He must have been driven by *some* motives or goals. Something must have kept Spock from spending his days calculating pi to a quadrillion digits or memorizing the Manhattan telephone directory. Something must have impelled him to explore strange new worlds, to seek out new civilizations, and to boldly go where no man had gone before. Presumably it was intellectual curiosity, a drive to set and solve problems, and solidarity with allies—emotions all. And what would Spock have done when faced with a predator or an invading Klingon? Do a headstand? Prove the four-color map theorem? Presumably a part of his brain quickly mobilized his faculties to scope out how to flee and to take steps to avoid the vulnerable predicament in the future. That is, he had fear. Spock may not have been impulsive or demonstrative, but he must have had drives that impelled him to deploy his intellect in pursuit of certain goals rather than others.

A conventional computer program is a list of instructions that the machine executes until it reaches STOP. But the intelligence of aliens, robots, and animals needs a more flexible method of control. Recall that intelligence is the pursuit of goals in the face of obstacles. Without goals, the very concept of intelligence is meaningless. To get into my locked

apartment, I can force open a window, call the landlord, or try to reach the latch through the mail slot. Each of these goals is attained by a chain of subgoals. My fingers won't reach the latch, so the subgoal is to find pliers. But my pliers are inside, so I set up a sub-subgoal of finding a store and buying new pliers. And so on. Most artificial intelligence systems are built around means and ends, like the production system in Chapter 2 with its stack of goal symbols displayed on a bulletin board and the software demons that respond to them.

But where does the topmost goal, the one that the rest of the program tries to attain, come from? For artificial intelligence systems, it comes from the programmer. The programmer designs it to diagnose soybean diseases or predict the next day's Dow Jones Industrial Average. For organisms, it comes from natural selection. The brain strives to put its owner in circumstances like those that caused its ancestors to reproduce. (The brain's goal is not reproduction itself; animals don't know the facts of life, and people who do know them are happy to subvert them, such as when they use contraception.) The goals installed in *Homo sapiens*, that problem-solving, social species, are not just the Four Fs. High on the list are understanding the environment and securing the cooperation of others.

And here is the key to why we have emotions. An animal cannot pursue all its goals at once. If an animal is both hungry and thirsty, it should not stand halfway between a berry bush and a lake, as in the fable about the indecisive ass who starved between two haystacks. Nor should it nibble a berry, walk over and take a sip from the lake, walk back to nibble another berry, and so on. The animal must commit its body to one goal at a time, and the goals have to be matched with the best moments for achieving them. Ecclesiastes says that to every tiling there is a season, and a time to every purpose under heaven: a time to weep, and a time to laugh; a time to love, and a time to hate. Different goals are appropriate when a lion has you in its sights, when your child shows up in tears, or when a rival calls you an idiot in public.

The emotions are mechanisms that set the brain's highest-level goals. Once triggered by a propitious moment, an emotion triggers the cascade of subgoals and sub-subgoals that we call thinking and acting. Because the goals and means are woven into a multiply nested control structure of subgoals within subgoals within subgoals, no sharp line divides thinking from feeling, nor does thinking inevitably precede feeling or vice versa (notwithstanding the century of debate within psychology over

which comes first). For example, fear is triggered by a signal of impending harm like a predator, a cliff-top, or a spoken threat. It lights up the short-term goal of fleeing, subduing, or deflecting the danger, and gives the goal high priority, which we experience as a sense of urgency. It also lights up the longer-term goals of avoiding the hazard in the future and remembering how we got out of it this time, triggered by the state we experience as relief. Most artificial intelligence researchers believe that freely behaving robots (as opposed to the ones bolted to the side of an assembly line) will have to be programmed with something like emotions merely for them to know at every moment what to do next. (Whether the robots would be *sentient* of these emotions is another question, as we saw in Chapter 2.)

Fear also presses a button that readies the body for action, the so-called fight-or-flight response. (The nickname is misleading because the response prepares us for *any* time-sensitive action, such as grabbing a baby who is crawling toward the top of a stairwell.) The heart thumps to send blood to the muscles. Blood is rerouted from the gut and skin, leaving butterflies and clamminess. Rapid breathing takes in oxygen. Adrenaline releases fuel from the liver and helps the blood to clot. And it gives our face that universal deer-in-the-headlights look.

Each human emotion mobilizes the mind and body to meet one of the challenges of living and reproducing in the cognitive niche. Some challenges are posed by physical things, and the emotions that deal with them, like disgust, fear, and appreciation of natural beauty, work in straightforward ways. Others are posed by people. The problem in dealing with people is that people can deal back. The emotions that evolved in response to other people's emotions, like anger, gratitude, shame, and romantic love, are played on a complicated chessboard, and they spawn the passion and intrigue that misleads the Romantic. First let's explore emotions about things, then emotions about people.

THE SUBURBAN SAVANNA

The expression "a fish out of water" reminds us that every animal is adapted to a habitat. Humans are no exception. We tend to think that animals just go where they belong, like heat-seeking missiles, but the animals must experience these drives as emotions not unlike ours. Some

places are inviting, calming, or beautiful; others are depressing or scary. The topic in biology called "habitat selection" is, in the case of *Homo sapiens*, the same as the topic in geography and architecture called "environmental aesthetics": what kinds of places we enjoy being in.

Until very recently our ancestors were nomads, leaving a site when they had used up its edible plants and animals. The decision of where to go next was no small matter. Cosmides and Tooby write:

Imagine that you are on a camping trip that lasts a lifetime. Having to carry water from a stream and firewood from the trees, one quickly learns to appreciate the advantages of some campsites over others. Dealing with exposure on a daily basis quickly gives one an appreciation for sheltered sites, out of the wind, snow, or rain. For hunter-gatherers, there is no escape from this way of life: no opportunities to pick up food at the grocery store, no telephones, no emergency services, no artificial water supplies, no fuel deliveries, no cages, guns, or animal control officers to protect one from the predatory animals. In these circumstances, one's life depends on the operation of mechanisms that cause one to prefer habitats that provide sufficient food, water, shelter, information, and safety to support human life, and that cause one to avoid those that do not.

Homo sapiens is adapted to two habitats. One is the African savanna, in which most of our evolution took place. For an omnivore like our ancestors, the savanna is a hospitable place compared with other ecosystems. Deserts have little biomass because they have little water. Temperate forests lock up much of their biomass in wood. Rainforests—or, as they used to be called, jungles—place it high in the canopy, relegating omnivores on the ground to being scavengers who gather the bits that fall from above. But the savanna—grasslands dotted with clumps of trees—is rich in biomass, much of it in the flesh of large animals, because grass replenishes itself quickly when grazed. And most of the biomass is conveniently placed a meter or two from the ground. Savannas also offer expansive views, so predators, water, and paths can be spotted from afar. Its trees provide shade and an escape from carnivores.

Our second-choice habitat is the rest of the world. Our ancestors, after evolving on the African savannas, wandered into almost every nook and cranny of the planet. Some were pioneers who left the savanna and then other areas in turn, as the population expanded or the climate changed. Others were refugees in search of safety. Foraging tribes can't

stand one another. They frequently raid neighboring territories and kill any stranger who blunders into theirs.

We could afford this wanderlust because of our intellect. People explore a new landscape and draw up a mental resource map, rich in details about water, plants, animals, routes, and shelter. And if they can, they make their new homeland into a savanna. Native Americans and Australian aborigines used to burn huge swaths of woodland, opening them up for colonization by grasses. The ersatz savanna attracted grazing animals, which were easy to hunt, and exposed visitors before they got too close.

The biologist George Orians, an expert on the behavioral ecology of birds, recently turned his eye to the behavioral ecology of humans. With Judith Heerwagen, Stephen Kaplan, Rachel Kaplan, and others, he argues that our sense of natural beauty is the mechanism that drove our ancestors into suitable habitats. We innately find savannas beautiful, but we also like a landscape that is easy to explore and remember, and that we have lived in long enough to know its ins and outs.

In experiments on human habitat preference, American children and adults are shown slides of landscapes and asked how much they would like to visit or live in them. The children prefer savannas, even though they have never been to one. The adults like the savannas, too, but they like the deciduous and coniferous forests—which resemble much of the habitable United States—just as much. No one likes the deserts and the rainforests. One interpretation is that the children are revealing our species' default habitat preference, and the adults supplement it with the land with which they have grown familiar.

Of course, people do not have a mystical longing for ancient homelands. They are merely pleased by the landscape features that savannas tend to have. Orians and Heerwagen surveyed the professional wisdom of gardeners, photographers, and painters to learn what kinds of landscapes people find beautiful. They treated it as a second kind of data on human tastes in habitats, supplementing the experiments on people's reactions to slides. The landscapes thought to be the loveliest, they found, are dead ringers for an optimal savanna: semi-open space (neither completely exposed, which leaves one vulnerable, nor overgrown, which impedes vision and movement), even ground cover, views to the horizon, large trees, water, changes in elevation, and multiple paths leading out. The geographer Jay Appleton succinctly captured what makes a landscape appealing: prospect and refuge, or seeing without being seen. The combination allows us to learn the lay of the land safely.

The land itself must be legible, too. Anyone who has lost a trail in a dense forest or seen footage of sand dunes or snow drifts in all directions knows the terror of an environment lacking a frame of reference. A landscape is just a very big object, and we recognize complex objects by locating their parts in a reference frame belonging to the object (see Chapter 4). The reference frames in a mental map are big landmarks, like trees, rocks, and ponds, and long paths or boundaries, like rivers and mountain ranges. A vista without these guideposts is unsettling. Kaplan and Kaplan found another key to natural beauty, which they call mystery. Paths bending around hills, meandering streams, gaps in foliage, undulating land, and partly blocked views grab our interest by hinting that the land may have important features that could be discovered by further exploration.

People also love to look at animals and plants, especially flowers. If you are reading this book at home or in other pleasant but artificial surroundings, chances are you can look up and find animal, plant, or flower motifs in the decorations. Our fascination with animals is obvious. We eat them, they eat us. But our love of flowers, which we don't eat except in salads in overpriced restaurants, needs an explanation. We ran into it in Chapters 3 and 5. People are intuitive botanists, and a flower is a rich source of data. Plants blend into a sea of green and often can be identified only by their flowers. Flowers are harbingers of growth, marking the site of future fruit, nuts, or tubers for creatures smart enough to remember them.

Some natural happenings are deeply evocative, like sunsets, thunder, gathering clouds, and fire. Orians and Heerwagen note that they tell of an imminent and consequential change: darkness, a storm, a blaze. The emotions evoked are arresting, forcing one to stop, take notice, and prepare for what's to come.

Environmental aesthetics is a major factor in our lives. Mood depends on surroundings: think of being in a bus terminal waiting room or a lakeside cottage. People's biggest purchase is their home, and the three rules of home buying—location, location, and location—pertain, apart from nearness to amenities, to grassland, trees, bodies of water, and prospect (views). The value of the house itself depends on its refuge (cozy spaces) and mystery (nooks, bends, windows, multiple levels). And people in the unlikeliest of ecosystems strive for a patch of savanna to call their own. In New England, any land that is left alone quickly turns into a scruffy deciduous forest. During my interlude in suburbia, every weekend my fellow burghers and I would drag out our lawn mowers, leaf blowers,

weed whackers, limb loppers, branch pruners, stem snippers, hedge clippers, and wood chippers in a Sisyphean effort to hold the forest at bay. Here in Santa Barbara, the land wants to be an arid chaparral, but decades ago the city fathers dammed wilderness creeks and tunneled through mountains to bring water to thirsty lawns. During a recent drought, homeowners were so desperate for verdant vistas that they sprayed their dusty yards with green paint.

FOOD FOR THOUGHT

Great green gobs of greasy grimy gopher guts,
Mutilated monkey meat,
Concentrated chicken feet.
Jars and jars of petrified porpoise pus,
And me without my spoon!

—fondly remembered camp song, sung to the
tune of "The Old Gray Mare"; lyricist unknown

Disgust is a universal human emotion, signaled with its own facial expression and codified everywhere in food taboos. Like all the emotions, disgust has profound effects on human affairs. During World War II, American pilots in the Pacific went hungry rather than eat the toads and bugs that they had been taught were perfectly safe. Food aversions are tenacious ethnic markers, persisting long after other traditions have been abandoned.

Judged by the standards of modern science, disgust is manifestly irrational. People who are sickened by the thought of eating a disgusting object will say it is unsanitary or harmful. But they find a sterilized cockroach every bit as revolting as one fresh from the cupboard, and if the sterilized roach is briefly dunked into a beverage, they will refuse to drink it. People won't drink juice that has been stored in a brand-new urine collection bottle; hospital kitchens have found this an excellent way to stop pilferage. People won't eat soup if it is served in a brand-new bedpan or if it has been stirred with a new comb or fly-swatter. You can't pay most people to eat fudge baked in the shape of dog feces or to hold rubber vomit from a novelty store between their lips. One's own saliva is not disgusting as long as it is in one's mouth,

but most people won't eat from a bowl of soup into which they have spat.

Most Westerners cannot stomach the thought of eating insects, worms, toads, maggots, caterpillars, or grubs, but these are all highly nutritious and have been eaten by the majority of peoples throughout history. None of our rationalizations makes sense. You say that insects are contaminated because they touch feces or garbage? But many insects are quite sanitary. Termites, for example, just munch wood, but Westerners feel no better about eating them. Compare them with chickens, the epitome of palatability ("Try it—it tastes like chicken!"), which commonly eat garbage and feces. And we all savor tomatoes made plump and juicy from being fertilized with manure. Insects carry disease? So does all animal flesh. Just do what the rest of the world does—cook them. Insects have indigestible wings and legs? Pull them off, as you do with peel-and-eat shrimp, or stick to grubs and maggots. Insects taste bad? Here is a report from a British entomologist who was studying Laotian foodways and acquired a firsthand knowledge of his subject matter:

None distasteful, a few quite palatable, notably the giant waterbug. For the most part they were insipid, with a faint vegetable flavour, but would not anyone tasting bread, for instance, for the first time, wonder why we eat such a flavourless food? A toasted dungbeetle or soft-bodied spider has a nice crisp exterior and soft interior of soufflé consistency which is by no means unpleasant. Salt is usually added, sometimes chili or the leaves of scented herbs, and sometimes they are eaten with rice or added to sauces or curry. Flavour is exceptionally hard to define, but lettuce would, I think, best describe the taste of termites, cicadas, and crickets; lettuce and raw potato that of the giant *Nephila* spider, and concentrated Gorgonzola cheese that of the giant waterbug (*Lethocems indicus*). I suffered no ill effects from the eating of these insects.

The psychologist Paul Rozin has masterfully captured the psychology of disgust. Disgust is a fear of incorporating an offending substance into one's body. Eating is the most direct way to incorporate a substance, and as my camp song shows, it is the most horrific thought that a disgusting substance can arouse. Smelling or touching it is also unappealing. Disgust deters people from eating certain things, or, if it's too late, makes them spit or vomit them out. The facial expression says it all: the nose is wrinkled, constricting the nostrils, and the mouth is opened and the tongue pushed forward as if to squeegee offending material out.

Disgusting things come from animals. They include whole animals, parts of animals (particularly parts of carnivores and scavengers), and body products, especially viscous substances like mucus and pus and, most of all, feces, universally considered disgusting. Decaying animals and their parts are particularly revolting. In contrast, plants are sometimes distasteful, but distaste is different from disgust. When people avoid plant products—say, lima beans or broccoli—it is because they taste bitter or pungent. Unlike disgusting animal products, they are not felt to be unspeakably vile and polluting. Probably the most complicated thought anyone ever had about a disfavored vegetable was Clarence Darrow's: "I don't like spinach, and I'm glad I don't, because if I liked it I'd eat it, and I just hate it." Inorganic and non-nutritive stuff like sand, cloth, and bark are simply avoided, without strong feelings.

Not only are disgusting things always from animals, but things from animals are almost always disgusting. The nondisgusting animal parts are the exception. Of all the parts of all the animals in creation, people eat an infinitesimal fraction, and everything else is untouchable. Many Americans eat only the skeletal muscle of cattle, chickens, swine, and a few fish. Other parts, like guts, brains, kidneys, eyes, and feet, are beyond the pale, and so is any part of any animal not on the list: dogs, pigeons, jellyfish, slugs, toads, insects, and the other millions of animal species. Some Americans are even pickier, and are repulsed by the dark meat of chicken or chicken on the bone. Even adventurous eaters are willing to sample only a small fraction of the animal kingdom. And it is not just pampered Americans who are squeamish about unfamiliar animal parts. Napoleon Chagnon safeguarded his supply of peanut butter and hot dogs from his begging Yanomamø informants by telling them they were the feces and penises of cattle. The Yanomamø, who are hearty eaters of caterpillars and grubs, had no idea what cattle were but lost their appetite and left him to eat in peace.

A disgusting object contaminates everything it touches, no matter how brief the contact or how invisible the effects. The intuition behind not drinking a beverage that has been stirred with a flyswatter or dunked with a sterilized roach is that invisible contaminating bits—children call them cooties—have been left behind. Some objects, such as a new comb or bedpan, are tainted merely because they are designed to touch something disgusting, and others, such as a chocolate dog turd, are tainted by mere resemblance. Rozin observes that the psychology of disgust obeys

the two laws of sympathetic magic—voodoo—found in many traditional cultures: the law of contagion (once in contact, always in contact) and the law of similarity (like produces like).

Though disgust is universal, the list of nondisgusting animals differs from culture to culture, and that implies a learning process. As every parent knows, children younger than two put everything in their mouths, and psychoanalysts have had a field day interpreting their lack of revulsion for feces. Rozin and his colleagues studied the development of disgust by offering children various foods that American adults find disgusting. To the horror of their onlooking parents, sixty-two percent of toddlers ate imitation dog feces ("realistically crafted from peanut butter and odorous cheese"), and thirty-one percent ate a grasshopper.

Rozin suggests that disgust is learned in the middle school-age years, perhaps when children are scolded by their parents or they see the look on their parents' faces when they approach a disgusting object. But I find that unlikely. First, all the subjects older than toddlers behaved virtually the same as the adults did. For example, four-year-olds wouldn't eat imitation feces or drink juice with a grasshopper in it; the only difference between them and the adults was that the children were less sensitive to contamination by brief contact. (Not until the age of eight did the children reject juice briefly dipped with a grasshopper or with imitation dog feces.) Second, children above the age of two are notoriously finicky, and their parents struggle to get them to eat new substances, not to avoid old ones. (The anthropologist Elizabeth Cashdan has documented that children's willingness to try new foods plummets after the third birthday.) Third, if children had to learn what to avoid, then all animals would be palatable except for the few that are proscribed. But as Rozin himself points out, all animals are disgusting except for a few that are permitted. No child has to be taught to revile greasy grimy gopher guts or mutilated monkey meat.

Cashdan has a better idea. The first two years, she proposes, are a sensitive period for learning about food. During those years mothers control children's food intake and children eat whatever they are permitted. Then their tastes spontaneously shrink, and they stomach only the foods they were given during the sensitive period. Those distastes can last to adulthood, though adults occasionally overcome them from a variety of motives: to dine with others, to appear macho or sophisticated, to seek thrills, or to avert starvation when familiar fare is scarce.

What is disgust for? Rozin points out that the human species faces "the omnivore's dilemma." Unlike, say, koalas, who mainly eat eucalyptus leaves and are vulnerable when those become scarce, omnivores choose from a vast menu of potential foods. The downside is that many are poison. Many fish, amphibians, and invertebrates contain potent neurotoxins. Meats that are ordinarily harmless can house parasites like tapeworms, and when they spoil, meats can be downright deadly, because the microorganisms that cause putrefaction release toxins to deter scavengers and thereby keep the meat for themselves. Even in industrialized countries food contamination is a major danger. Until recently, anthrax and trichinosis were serious hazards, and today public health experts recommend draconian sanitary measures so people won't contract salmonella poisoning from their next chicken salad sandwich. In 1996 a world crisis was set off by the discovery that Mad Cow Disease, a pathology found in some British cattle that makes their brains spongy, might do the same to people who eat the cattle.

Rozin ventured that disgust is an adaptation that deterred our ancestors from eating dangerous animal stuff. Feces, carrion, and soft, wet animal parts are home to harmful microorganisms and ought to be kept outside the body. The dynamics of learning about food in childhood fit right in. Which animal parts are safe depends on the local species and their endemic diseases, so particular tastes cannot be innate. Children use their older relatives the way kings used food tasters: if they ate something and lived, it is not poison. Thus very young children are receptive to whatever their parents let them eat, and when they are old enough to forage on their own, they avoid everything else.

But how can one explain the irrational effects of similarity—the revulsion for rubber vomit, chocolate dog turds, and sterilized roaches? The answer is that these items were *crafted* to evoke the same reaction in people that the objects themselves evoke. That is why novelty shops *sell* rubber vomit. The similarity effect merely shows that reassurance by an authority or by one's own beliefs do not disconnect an emotional response. It is no more irrational than other reactions to modern simulacra, such as being engrossed by a movie, aroused by pornography, or terrified on a roller coaster.

What about our feeling that disgusting things contaminate every-

thing they touch? It is a straightforward adaptation to a basic fact about the living world: germs multiply. Microorganisms are fundamentally different from chemical poisons such as those manufactured by plants. The danger of a chemical depends on its dose. Poisonous plants are bitter-tasting because both the plant and the plant-eater have an interest in the plant-eater stopping after the first bite. But there is no safe dose for a microorganism, because they reproduce exponentially. A single, invisible, untastable germ can multiply and quickly saturate a substance of any size. Since germs are, of course, transmittable by contact, it is no surprise that anything that touches a yucky substance is itself forever yucky, even if it looks and tastes the same. Disgust is intuitive microbiology.

Why are insects and other small creatures like worms and toads—what Latin Americans call "animalitos"—so easy to revile? The anthropologist Marvin Harris has shown that cultures avoid animalitos when larger animals are available, and eat them when they are not. The explanation has nothing to do with sanitation, since bugs are safer than meat. It comes from optimal foraging theory, the analysis of how animals ought to—and usually do—allocate their time to maximize the rate of nutrients they consume. Animalitos are small and dispersed, and it takes a lot of catching and preparing to get a pound of protein. A large mammal is hundreds of pounds of meat on the hoof, available all at once. (In 1978 a rumor circulated that McDonald's was extending the meat in Big Macs with earthworms. But if the corporation were as avaricious as the rumor was meant to imply, the rumor could not be true: worm meat is far more expensive than beef.) In most environments it is not only more efficient to eat larger animals, but the small ones should be avoided altogether—the time to gather them would be better spent hunting for a bigger payoff. Animalitos are thus absent from the diets of cultures that have bigger fish to fry, and since, in the minds of eaters, whatever is not permitted is forbidden, those cultures find them disgusting.

What about food taboos? Why, for example, are Hindus forbidden to eat beef? Why are Jews forbidden to eat pork and shellfish and to mix meat with milk? For thousands of years, rabbis have offered ingenious justifications of the Jewish dietary laws. Here are a few listed in the *Encyclopaedia Judaica*:

From Aristeas, first century BC: "The dietary laws are ethical in intent, since abstention from the consumption of blood tames man's instinct for violence by instilling in him a horror of bloodshed. . . . The injunction against the consumption of birds of prey was intended to demonstrate that man should not prey on others."

From Isaac ben Moses Arama: "The reason behind all the dietary prohibitions is not that any harm may be caused to the body, but that these foods defile and pollute the soul and blunt the intellectual powers, thus leading to confused opinions and a lust for perverse and brutish appetites which lead men to destruction, thus defeating the purpose of creation."

From Maimonides: "All the food which the Torah has forbidden us to eat have some bad and damaging effect on the body. . . . The principal reason why the Law forbids swine's flesh is to be found in the circumstances that its habits and its food are very dirty and loathsome. . . . The fat of the intestines is forbidden because it fattens and destroys the abdomen and creates cold and clammy blood. . . . Meat boiled in milk is undoubtedly gross food, and makes a person feel overfull."

From Abraham ibn Ezra: "I believe it is a matter of cruelty to cook a kid in its mother's milk."

From Nahmanides: "Now the reason for specifying fins and scales is that fish which have fins and scales get nearer to the surface of the water and are found more generally in freshwater areas. . . . Those without fins and scales usually live in the lower muddy strata which are exceedingly moist and where there is no heat. They breed in musty swamps and eating them can be injurious to health."

With all due respect to rabbinical wisdom, these arguments can be demolished by any bright twelve-year-old, and as a former temple Sunday School teacher I can attest that they regularly are. Many Jewish adults still believe that pork was banned as a public health measure, to prevent trichinosis. But as Harris points out, if that were true the law would have been a simple advisory against undercooking pork: "Flesh of swine thou shalt not eat until the pink has been cooked from it!"

Harris observes that food taboos often make ecological and economic sense. The Hebrews and the Muslims were desert tribes, and pigs are animals of the forest. They compete with people for water and nutritious foods like nuts, fruits, and vegetables. Kosher animals, in contrast, are

ruminants like sheep, cattle, and goats, which can live off scraggly desert plants. In India, cattle are too precious to slaughter because they are used for milk, manure, and pulling plows. Harris' theory is as ingenious as the rabbis' and far more plausible, though he admits that it can't explain everything. Ancient tribes wandering the parched Judaeen sands were hardly in danger of squandering their resources by herding shrimp and oysters, and it is unclear why the inhabitants of a Polish shtetl or a Brooklyn neighborhood should obsess over the feeding habits of desert ruminants.

Food taboos are obviously an ethnic marker, but by itself that observation explains nothing. Why do people wear ethnic badges to begin with, let alone a costly one like banning a source of nutrients? The social sciences assume without question that people submerge their interests to the group, but on evolutionary grounds that is unlikely (as we shall see later in the chapter). I take a more cynical view.

In any group, the younger, poorer, and disenfranchised members may be tempted to defect to other groups. The powerful, especially parents, have an interest in keeping them in. People everywhere form alliances by eating together, from potlatches and feasts to business lunches and dates. If I can't eat with you, I can't become your friend. Food taboos often prohibit a favorite food of a neighboring tribe; that is true, for example, of many of the Jewish dietary laws. That suggests that they are weapons to keep potential defectors in. First, they make the merest prelude to cooperation with outsiders—breaking bread together—an unmistakable act of defiance. Even better, they exploit the psychology of disgust. Taboo foods are absent during the sensitive period for learning food preferences, and that is enough to make children grow up to find them disgusting. That deters them from becoming intimate with the enemy ("He invited me over, but what will I do if they serve . . . EEEU-UUW!!!"). Indeed, the tactic is self-perpetuating because children grow up into parents who don't feed the disgusting things to *their* children. The practical effects of food taboos have often been noticed. A familiar theme in novels about the immigrant experience is the protagonist's torment over sampling taboo foods. Crossing the line offers a modicum of integration into the new world but provokes open conflict with parents and community. (In *Portnoy's Complaint*, Alex describes his mother as pronouncing *hamburger* as if it were *Hitler*.) But since the elders have no desire for the community to see the taboos in this light, they cloak them in talmudic sophistry and bafflegab.

THE SMELL OF FEAR

Language-lovers know that there is a word for every fear. Are you afraid of wine? Then you have *oenophobia*. Tremulous about train travel? You suffer from *siderodromophobia*. Having misgivings about your mother-in-law is *pentheraphobia*, and being petrified of peanut butter sticking to the roof of your mouth is *arachibutyrophobia*. And then there's Franklin Delano Roosevelt's affliction, the fear of fear itself, or *phobophobia*.

But just as not having a word for an emotion doesn't mean that it doesn't exist, having a word for an emotion doesn't mean that it does exist. Word-watchers, verbivores, and sesquipedalians love a challenge. Their idea of a good time is to find the shortest word that contains all the vowels in alphabetical order or to write a novel without the letter *e*. Yet another joy of lex is finding names for hypothetical fears. That is where these improbable phobias come from. Real people do not tremble at the referent of every euphonious Greek or Latin root. Fears and phobias fall into a short and universal list.

Snakes and spiders are always scary. They are the most common objects of fear and loathing in studies of college students' phobias, and have been so for a long time in our evolutionary history. D. O. Hebb found that chimpanzees born in captivity scream in terror when they first see a snake, and the primatologist Marc Hauser found that his laboratory-bred cotton-top tamarins (a South American monkey) screamed out alarm calls when they saw a piece of plastic tubing on the floor. The reaction of foraging peoples is succinctly put by Irven DeVore: "Hunter-gatherers will not suffer a snake to live." In cultures that revere snakes, people still treat them with great wariness. Even Indiana Jones was afraid of them!

The other common fears are of heights, storms, large carnivores, darkness, blood, strangers, confinement, deep water, social scrutiny, and leaving home alone. The common thread is obvious. These are the situations that put our evolutionary ancestors in danger. Spiders and snakes are often venomous, especially in Africa, and most of the others are obvious hazards to a forager's health, or, in the case of social scrutiny, status. Fear is the emotion that motivated our ancestors to cope with the dangers they were likely to face.

Fear is probably several emotions. Phobias of physical things, of social

scrutiny, and of leaving home respond to different kinds of drugs, suggesting that they are computed by different brain circuits. The psychiatrist Isaac Marks has shown that people react in different ways to different frightening things, each reaction appropriate to the hazard. An animal triggers an urge to flee, but a precipice causes one to freeze. Social threats lead to shyness and gestures of appeasement. People really do faint at the sight of blood, because their blood pressure drops, presumably a response that would minimize the further loss of one's own blood. The best evidence that fears are adaptations and not just bugs in the nervous system is that animals that have evolved on islands without predators lose their fear and are sitting ducks for any invader—hence the expression "dead as a dodo."

Fears in modern city-dwellers protect us from dangers that no longer exist, and fail to protect us from dangers in the world around us. We ought to be afraid of guns, driving fast, driving without a seatbelt, lighter fluid, and hair dryers near bathtubs, not of snakes and spiders. Public safety officials try to strike fear in the hearts of citizens using everything from statistics to shocking photographs, usually to no avail. Parents scream and punish to deter their children from playing with matches or chasing a ball into the street, but when Chicago schoolchildren were asked what they were most afraid of, they cited lions, tigers, and snakes, unlikely hazards in the Windy City.

Of course, fears do change with experience. For decades psychologists thought that animals learn new fears the way Pavlov's dogs learned to salivate to a bell. In a famous experiment, John B. Watson, the founder of behaviorism, came up behind an eleven-month-old boy playing with a tame white rat and suddenly clanged two steel bars together. After a few more clangs, the boy became afraid of the rat and other white furry things, including rabbits, dogs, a sealskin coat, and Santa Claus. The rat, too, can learn to associate danger with a previously neutral stimulus. A rat shocked in a white room will flee it for a black room every time it is dumped there, long after the shocker has been unplugged.

But in fact creatures cannot be conditioned to fear just any old thing. Children are nervous about rats, and rats are nervous about bright rooms, before any conditioning begins, and they easily associate them with danger. Change the white rat to some arbitrary object, like opera glasses, and the child never learns to fear it. Shock the rat in a black room instead of a white one, and that nocturnal creature learns the association more slowly and unlearns it more quickly. The psychologist Mar-

tin Seligman suggests that fears can be easily conditioned only when the animal is evolutionarily prepared to make the association.

Few, if any, human phobias are about neutral objects that were once paired with some trauma. People dread snakes without ever having seen one. After a frightening or painful event, people are more prudent around the cause, but they do not fear it; there are no phobias for electrical outlets, hammers, cars, or air-raid shelters. Television clichés notwithstanding, most survivors of a traumatic event do not get the screaming meemies every time they face a reminder of it. Vietnam veterans resent the stereotype in which they hit the dirt whenever someone drops a glass.

A better way to understand the learning of fears is to think through the evolutionary demands. The world is a dangerous place, but our ancestors could not have spent their lives cowering in caves; there was food to gather and mates to win. They had to calibrate their fears of typical dangers against the actual dangers in the local environment (after all, not *all* spiders are poisonous) and against their own ability to neutralize the danger: their know-how, defensive technology, and safety in numbers.

Marks and the psychiatrist Randolph Nesse argue that phobias are innate fears that have never been unlearned. Fears develop spontaneously in children. In their first year, babies fear strangers and separation, as well they should, for infanticide and predation are serious threats to the tiniest hunter-gatherers. (The film *A Cry in the Dark* shows how easily a predator can snatch an unattended baby. It is an excellent answer to every parent's question of why the infant left alone in a dark bedroom is screaming bloody murder.) Between the ages of three and five, children become fearful of all the standard phobic objects—spiders, the dark, deep water, and so on—and then master them one by one. Most adult phobias are childhood fears that never went away. That is why it is city-dwellers who most fear snakes.

As with the learning of safe foods, the best guides to the local dangers are the people who have survived them. Children fear what they see their parents fear, and often unlearn their fears when they see other children coping. Adults are just as impressionable. In wartime, courage and panic are both contagious, and in some therapies, the phobic watches as an aide plays with a boa constrictor or lets a spider crawl up her arm. Even monkeys watch one another to calibrate their fear. Laboratory-raised rhesus macaques are not afraid of snakes when they first see them, but if they watch a film of another monkey being frightened by a snake, they fear it, too. The monkey in the movie does not instill the fear

so much as awaken it, for if the film shows the monkey recoiling from a flower or a bunny instead of a snake, the viewer develops no fear.

The ability to conquer fear selectively is an important component of the instinct. People in grave danger, such as pilots in combat or Londoners during the blitz, can be remarkably composed. No one knows why some people can keep their heads when all about them are losing theirs, but the main calming agents are predictability, allies within shouting distance, and a sense of competence and control, which the writer Tom Wolfe called *The Right Stuff*. In his book by that name about the test pilots who became Mercury astronauts, Wolfe defined the right stuff as "the ability [of a pilot] to go up in a hurtling piece of machinery and put his hide on the line and then have the moxie, the reflexes, the experience, the coolness, to pull it back in the last yawning moment." That sense of control comes from "pushing the outside of the envelope": testing, in small steps, how high, how fast, how far one can go without bringing on disaster. Pushing the envelope is a powerful motive. Recreation, and the emotion called "exhilaration," come from enduring relatively safe events that look and feel like ancestral dangers. These include most non-competitive sports (diving, climbing, spelunking, and so on) and the genres of books and movies called "thrillers." Winston Churchill once said, "Nothing in life is so exhilarating as to be shot at without result."

THE HAPPINESS TREADMILL

The pursuit of happiness is an inalienable right, says the Declaration of Independence in its list of self-evident truths. The greatest happiness of the greatest number, wrote Jeremy Bentham, is the foundation of morality. To say that everyone wants to be happy sounds trite, almost circular, but it raises a profound question about our makeup. What is this thing that people strive for?

At first happiness might seem like just desserts for biological fitness (more accurately, the states that would have led to fitness in the environment in which we evolved). We are happier when we are healthy, well-fed, comfortable, safe, prosperous, knowledgeable, respected, non-celibate, and loved. Compared to their opposites, these objects of striving are conducive to reproduction. The function of happiness would be to mobilize the mind to seek the keys to Darwinian fitness. When we are

unhappy, we work for the things that make us happy; when we are happy, we keep the status quo.

The problem is, how much fitness is worth striving for? Ice Age people would have been wasting their time if they had fretted about their lack of camping stoves, penicillin, and hunting rifles or if they had striven for them instead of better caves and spears. Even among modern foragers, very different standards of living are attainable in different times and places. Lest the perfect be the enemy of the good, the pursuit of happiness ought to be calibrated by what can be attained through reasonable effort in the current environment.

How do we know what can reasonably be attained? A good source of information is what other people have attained. If they can get it, perhaps so can you. Through the ages, observers of the human condition have pointed out the tragedy: people are happy when they feel better off than their neighbors, unhappy when they feel worse off.

But, O! how bitter a thing it is to look into happiness through another man's eyes!

—William Shakespeare (*As You Like It*, V, ii).

Happiness, n. An agreeable sensation arising from contemplating the misery of others.

—Ambrose Bierce

It is not enough to succeed. Others must fail.

—Gore Vidal

Ven frait zich a hoiker? Ven er zet a gresseren hoiker far zich. (When does a hunchback rejoice? When he sees one with a larger hump.)

—Yiddish saying

Research on the psychology of happiness has borne out the curmudgeons. Kahneman and Tversky give an everyday example. You open your paycheck and are delighted to find you have been given a five percent raise—until you learn that your co-workers have been given a ten percent raise. According to legend, the diva Maria Callas stipulated that any opera house she sang in had to pay her one dollar more than the next highest paid singer in the company.

People today are safer, healthier, better fed, and longer-lived than at any time in history. Yet we don't spend our lives walking on air, and presumably

our ancestors were not chronically glum. It is not reactionary to point out that many of the poor in today's Western nations live in conditions that yesterday's aristocrats could not have dreamed of. People in different classes and countries are often content with their lot until they compare themselves to the more affluent. The amount of violence in a society is more closely related to its inequality than to its poverty. In the second half of the twentieth century, the discontent of the Third World, and later the Second, have been attributed to their glimpses through the mass media of the First.

The other major clue to the attainable is how well off you are now. What you have now is attainable, by definition, and chances are you can do at least a little bit better. Evolutionary theory predicts that a man's reach should exceed his grasp, but not by much. Here we have the second tragedy of happiness: people adapt to their circumstances, good or bad, the way their eyes adapt to sun or darkness. From that neutral point, improvement is happiness, loss is misery. Again, the sages said it first. The narrator of E. A. Robinson's poem (and later Simon and Garfunkel's song) envies the factory owner, Richard Cory, who "glittered when he walked."

So on we worked, and waited for the light,
And went without the meat, and cursed the bread;
And Richard Cory, one calm summer night,
Went home and put a bullet through his head.

The futility of striving has led many dark souls to deny that happiness is possible. For the show-business personality Oscar Levant, "Happiness is not something you experience, it's something you remember." Freud said that the goal of psychotherapy was "to transform hysterical misery into common unhappiness." A colleague, consulting with me by email about a troubled graduate student, wrote, "sometimes i wish i was young then i remember that wasn't so great either."

But here the curmudgeons are only partly right. People do come to feel the same across an astonishing range of good and bad fortunes. But the baseline that people adapt to, on average, is not misery but satisfaction. (The exact baseline differs from person to person and is largely inherited.) The psychologists David Myers and Ed Diener have found that about eighty percent of people in the industrialized world report that they are at least "fairly satisfied with life," and about thirty percent say they are "very happy." (As far as we can tell, the reports are sincere.) The percentages are the same for all ages, for both sexes, for blacks and

whites, and over four decades of economic growth. As Myers and Diener remark, "Compared with 1957, Americans have twice as many cars per person—plus microwave ovens, color TVs, VCRs, air conditioners, answering machines, and \$12 billion worth of new brand-name athletic shoes a year. So, are Americans happier than they were in 1957? They are not."

Within an industrialized country, money buys only a little happiness: the correlation between wealth and satisfaction is positive but small. Lottery winners, after their jolt of happiness has subsided, return to their former emotional state. On the brighter side, so do people who have suffered terrible losses, such as paraplegics and survivors of the Holocaust.

These findings do not necessarily contradict the singer Sophie Tucker when she said, "I have been poor and I have been rich. Rich is better." In India and Bangladesh, wealth predicts happiness much better than it does in the West. Among twenty-four Western European and American nations, the higher the gross national product per capita, the happier the citizens (though there are many explanations). Myers and Diener point out that wealth is like health: not having it makes you miserable, but having it does not guarantee happiness.

The tragedy of happiness has a third act. There are twice as many negative emotions (fear, grief, anxiety, and so on) as positive ones, and losses are more keenly felt than equivalent gains. The tennis star Jimmy Connors once summed up the human condition: "I hate to lose more than I like to win." The asymmetry has been confirmed in the lab by showing that people will take a bigger gamble to avoid a sure loss than to improve on a sure gain, and by showing that people's mood plummets more when imagining a loss in their lives (for example, in course grades, or in relationships with the opposite sex) than it rises when imagining an equivalent gain. The psychologist Timothy Ketelaar notes that happiness tracks the effects of resources on biological fitness. As things get better, increases in fitness show diminishing returns: more food is better, but only up to a point. But as things get worse, decreases in fitness can take you out of the game: not enough food, and you're dead. There are many ways to become infinitely worse off (from an infection, starvation, getting eaten, a fall, ad infinitum) and not many ways to become vastly better off. That makes prospective losses more worthy of attention than gains; there are more things that make us unhappy than things that make us happy.

Donald Campbell, an early evolutionary psychologist who studied the

psychology of pleasure, described humans as being on a "hedonic treadmill," where gains in well-being leave us no happier in the long run. Indeed, the study of happiness often sounds like a sermon for traditional values. The numbers show that it is not the rich, privileged, robust, or good-looking who are happy; it is those who have spouses, friends, religion, and challenging, meaningful work. The findings can be overstated, because they apply to averages, not individuals, and because cause and effect are hard to tease apart: being married might make you happy, but being happy might help you get and stay married. But Campbell echoed millennia of wise men and women when he summed up the research: "The direct pursuit of happiness is a recipe for an unhappy life."

THE SIRENS' SONG

When we say that someone is led by emotion rather than reason, we often mean that the person sacrifices long-term interests for short-term gratification. Losing one's temper, surrendering to a seducer, blowing one's paycheck, and turning tail at the dentist's door are examples. What makes us so short-sighted?⁴

The ability to defer a reward is called self-control or delay of gratification. Social scientists often treat it as a sign of intelligence, of the ability to anticipate the future and plan accordingly. But discounting the future, as economists call it, is part of the logic of choice for any agent that lives longer than an instant. Going for the quick reward instead of a distant payoff is often the rational strategy.

Which is better, a dollar now or a dollar a year from now? (Assume there is no inflation.) A dollar now, you might say, because you can invest it and have more than a dollar in a year. Unfortunately, the explanation is circular: the reason that interest exists in the first place is to pay people to give up the dollar that they would rather have now than a year from now. But economists point out that even if the explanation is misplaced, the answer is right: now really *is* better. First, a dollar now is available if a pressing need or opportunity arises in less than a year. Second, if you forgo the dollar now, you have no guarantee that you will get it back a year from now. Third, you might die within a year and never get to enjoy it. It is rational, therefore, to discount the future: to consume a resource now unless investing it brings a high enough return. The interest rate you

should demand depends on how important the money is to you now, how likely you are to get it back, and how long you expect to live.

The struggle to reproduce is a kind of economy, and all organisms, even plants, must "decide" whether to use resources now or save them for the future. Some of these decisions are made by the body. We grow frail with age because our genes discount the future and build strong young bodies at the expense of weak old ones. The exchange pays off over the generations because an accident may cause the body to die before it gets old, in which case any sacrifice of vigor for longevity would have gone to waste. But most decisions about the future are made by the mind. At every moment we choose, consciously or unconsciously, between good things now and better things later.

Sometimes the rational decision is "now," particularly when, as the sayings go, life is short or there is no tomorrow. The logic is laid bare in firing-squad jokes. The condemned man is offered the ceremonial last cigarette and responds, "No thanks, I'm trying to quit." We laugh because we know it is pointless for him to delay gratification. Another old joke makes it clear why playing it safe is not always called for. Murray and Esther, a middle-aged Jewish couple, are touring South America. One day Murray inadvertently photographs a secret military installation, and soldiers hustle the couple off to prison. For three weeks they are tortured in an effort to get them to name their contacts in the liberation movement. Finally they are hauled in front of a military court, charged with espionage, and sentenced to death by firing squad. The next morning they are lined up in front of the wall and the sergeant asks them if they have any last requests. Esther wants to know if she can call her daughter in Chicago. The sergeant says that's not possible, and turns to Murray. "This is crazy," Murray shouts, "we're not spies!" and he spits in the sergeant's face. "Murray!" Esther cries. "Please! Don't make trouble!"

Most of the time we are pretty sure that we will not die in minutes. But we all die sometime, and we all risk forgoing the opportunity to enjoy something if we defer it too long. In our ancestors' nomadic lifestyle, without an ability to accumulate possessions or to count on long-lived social institutions like depositors' insurance, the payoffs for consumption must have been even higher. But even if they were not, *some* urge to indulge now had to have been built into our emotions. Most likely, we evolved a mechanism to estimate our longevity and the opportunities and risks posed by different choices (eating now or later, setting up camp or pushing on) and to tune the emotions accordingly.

The political scientist James Q. Wilson and the psychologist Richard Herrnstein have pointed out that many criminals act as if they discount the future steeply. A crime is a gamble whose payoff is immediate and whose possible cost comes later. They attributed the discounting to low intelligence. The psychologists Martin Daly and Margo Wilson have a different explanation. In the American inner cities, life expectancy for young males is low, and they know it. (In *Hoop Dreams*, the documentary about aspiring basketball players in a Chicago ghetto, there is an arresting scene in which the mother of one of the boys rejoices that he is alive on his eighteenth birthday.) Moreover, the social order and long-term ownership rights which would guarantee that investments are repaid are tenuous. These are precisely the circumstances in which steeply discounting the future—taking risks, consuming rather than investing—is adaptive.

More puzzling is *myopic* discounting: the tendency in all of us to prefer a large late reward to a small early one, but then to flip our preference as time passes and both rewards draw nearer. A familiar example is deciding before dinner to skip dessert (a small early reward) in order to lose weight (a large late one), but succumbing to temptation when the waiter takes the dessert orders. Myopic discounting is easy to produce in the lab: give people (or pigeons, for that matter) two buttons, one delivering a small reward now, the other delivering a larger reward later, and the subject will flip from choosing the large reward to choosing the small reward as the small one becomes imminent. The weakness of the will is an unsolved problem in economics and psychology alike. The economist Thomas Schelling asks a question about the "rational consumer" that can also be posed of the adapted mind:

How should we conceptualize this rational consumer whom all of us know and who some of us are, who in self-disgust grinds his cigarettes down the disposal swearing that this time he means never again to risk orphaning his children with lung cancer and is on the street three hours later looking for a store that's still open to buy cigarettes; who eats a high-calorie lunch knowing that he will regret it, does regret it, cannot understand how he lost control, resolves to compensate with a low-calorie dinner, eats a high-calorie dinner knowing he will regret it, and does regret it; who sits glued to the TV knowing that again tomorrow he'll wake early in a cold sweat unprepared for that morning meeting on which so much of his career depends; who spoils the trip to Disneyland by losing his temper when his children do what he knew they were going to do when he resolved not to lose his temper when they did it?

Schelling notes the strange ways in which we defeat our self-defeating behavior: putting the alarm clock across the room so we won't turn it off and fall back to sleep, authorizing our employers to put part of each paycheck away for retirement, placing tempting snacks out of reach, setting our watches five minutes ahead. Odysseus had his crewmates plug their ears with wax and tie him to the mast so he could hear the Sirens' alluring song and not steer the ship toward them and onto the rocks.

Though myopic discounting remains unexplained, Schelling captures something important about its psychology when he roots the paradox of self-control in the modularity of the mind. He observes that "people behave sometimes as if they had two selves, one who wants clean lungs and long life and another who adores tobacco, or one who wants a lean body and another who wants dessert, or one who yearns to improve himself by reading Adam Smith on self-command . . . and another who would rather watch an old movie on television. The two are in continual contest for control." When the spirit is willing but the flesh is weak, such as in pondering a diet-busting dessert, we can feel two very different kinds of motives fighting within us, one responding to sights and smells, the other to doctors' advice. What about when the rewards are of the same kind, like a dollar today versus two dollars tomorrow? Perhaps an imminent reward engages a circuit for dealing with sure things and a distant one a circuit for betting on an uncertain future. One outranks the other, as if the whole person was designed to believe that a bird in the hand is worth two in the bush. In the modern environment, with its reliable knowledge of the future, that often leads to irrational choices. But our ancestors might have done well to distinguish between what is definitely enjoyable now and what is conjectured or rumored to be more enjoyable tomorrow. Even today, the delay of gratification is sometimes punished because of the frailty of human knowledge. Retirement funds go bankrupt, governments break promises, and doctors announce that everything they said was bad for you is good for you and vice versa.

I AND THOU

Our most ardent emotions are evoked not by landscapes, spiders, roaches, or dessert, but by other people. Some emotions, such as anger, make us want to harm people; others, such as love, sympathy, and grati-

tude, make us want to help them. To understand these emotions, we first have to understand why organisms should be designed to help or to hurt one another.

Having seen nature documentaries, you may believe that wolves weed out the old and weak deer to keep the herd healthy, that lemmings commit suicide to prevent the population from starving, or that stags ram into each other for the right to breed so that the fittest individuals may perpetuate the species. The underlying assumption—that animals act for the good of the ecosystem, the population, or the species—seems to follow from Darwin's theory. If in the past there were ten populations of lemmings, nine with selfish lemmings who ate their groups into starvation and one in which some died so that others might live, the tenth group would survive and today's lemmings should be willing to make the ultimate sacrifice. The belief is widespread. Every psychologist who has written about the function of the social emotions has talked about their benefit to the group.

When people say that animals act for the good of the group, they seem not to realize that the assumption is in fact a radical departure from Darwinism and almost certainly wrong. Darwin wrote, "Natural selection will never produce in a being any structure more injurious than beneficial to that being, for natural selection acts solely by and for the good of each." Natural selection could select groups with selfless members only if each group could enforce a pact guaranteeing that all their members stayed selfless. But without enforcement, nothing could prevent a mutant or immigrant lemming from thinking, in effect, "To heck with this! I'll let everyone *else* jump off the cliff, and then enjoy the food they leave behind." The selfish lemming would reap the rewards of the others' selflessness without paying any costs himself. With that advantage, his descendants would quickly take over the population, even if the population as a whole was worse off. And that is the fate of any tendency toward sacrifice. Natural selection is the cumulative effect of the relative successes of different replicators. That means that it selects for the replicators that replicate best, namely, the selfish ones.

The inescapable fact that adaptations benefit the replicator was first articulated by the biologist George Williams and later amplified by Richard Dawkins in *The Selfish Gene*. Almost all evolutionary biologists now accept the point, though there are debates over other issues. Selection among groups is possible on paper, but most biologists doubt that the special circumstances that let it happen are ever found in the real

world. Selection among branches of the tree of life is possible, but that has nothing to do with whether organisms are designed for unselfishness. Animals just don't care what happens to their group, species, or ecosystem. Wolves catch the old and weak deer because they are the easiest to catch. Hungry lemmings set out for better feeding grounds and sometimes fall or drown by accident, not suicide. Stags fight because each wants to breed, and one concedes when defeat is inevitable, or as part of a strategy that works on average against others playing the same strategy. Males who fight are wasteful to the group—indeed, males *in general* are wasteful to the group when they make up half of it, because a few studs could sire the next generation without eating half the food.

Biologists often describe these acts as self-interested behavior, but what causes behavior is the activity of the brain, especially the circuitry for emotions and other feelings. Animals behave selfishly because of how their emotion circuits are wired. My full stomach, my warmth, my orgasms, feel better to me than yours do, and I want mine, and will seek mine, more than yours. Of course, one animal cannot directly feel what's in another one's stomach, but it could feel it indirectly by observing the second animal's behavior. So it is an interesting psychological fact that animals usually don't experience other animals' observable well-being as their own pleasure. It is an even more interesting fact that they sometimes do.

Earlier I said that natural selection selects selfish replicators. If organisms were replicators, all organisms should be selfish. But organisms do not replicate. Your parents did not replicate when they had you, because you are not identical to either of them. The blueprint that made you—your set of genes—is not the same as the blueprint that made them. Their genes were shuffled, randomly sampled to make sperm and eggs, and combined with each other's during fertilization to create a new combination of genes and a new organism unlike them. The only things that actually replicated were the genes and fragments of genes whose copies made it into you, some of which you will in turn pass down to your children, and so on. In fact, even if your mother had cloned herself, she would not have replicated; only her genes would have. That is because any changes she underwent in her lifetime—losing a finger, acquiring a

tattoo, having her nose pierced—were not passed on to you. The only change you could have inherited was a mutation of one of the genes in the egg that was to become you. Genes, not bodies, replicate, and that means that genes, not bodies, should be selfish.

DNA, of course, has no feelings; "selfish" means "acting in ways that make one's own replication more likely." The way for a gene to do that in an animal with a brain is to wire the brain so that the animal's pleasures and pains cause it to act in ways that lead to more copies of the gene. Often that means causing an animal to enjoy the states that make it survive and reproduce. A full belly is satisfying because full bellies keep animals alive and moving and reproducing, leading to more copies of the genes that build brains that make full bellies feel satisfying.

By building a brain that makes eating fun, a gene helps to spread copies of itself lying in the animal's gonads. The actual DNA that helps build a brain, of course, doesn't itself get passed into the egg or sperm; only the copies of the gene inside the gonads do. But here is an important twist. The genes in an animal's gonads are not the *only* extant copies of the brain-building genes; they are merely the most convenient ones for the brain-building gene to help replicate. *Any* copy capable of replicating, anywhere in the world, is a legitimate target, if it can be identified and if steps can be taken to help it replicate. A gene that worked to replicate copies of itself inside some *other* animal's gonads could do as well as a gene that worked to replicate copies of itself inside *its own* animal's gonads. As far as the gene is concerned, a copy is a copy; which animal houses it is irrelevant. To a brain-building gene, the only thing special about that animal's gonads is the *certainty* that copies of the gene will be found in those gonads (the certainty comes from the fact that the cells in an animal's body are genetic clones). That is why the brain-building genes make animals enjoy their own well-being so much. If a gene could build a brain that could tell when copies of itself were sitting in *another* animal's gonads, it would make the brain enjoy the *other* animal's well-being, and make it act in ways that increased that other animal's well-being.

When does a copy of a gene in one animal also sit inside another? When the animals are related. In most animals there is a one-in-two chance that any gene in a parent will have a copy lying inside its offspring, because offspring get half their genes from each parent. There is also a one-in-two chance that a copy is lying inside a full sibling, because full siblings inherit their genes from the same pair of parents. There is a

one-in-eight chance that a copy is lying inside a first cousin, and so on. A gene that built a brain that made its owner help its relatives would indirectly help to replicate itself. The biologist William Hamilton noted that if the benefit to the relative, multiplied by the probability that a gene is shared, exceeds the cost to the animal, that gene would spread in the population. Hamilton developed and formalized an idea that had been entertained by several other biologists as well, most famously in a wise-crack by the biologist J. B. S. Haldane when he was asked if he would lay down his life for his brother. "No," he said, "but for two brothers or eight cousins."

When an animal behaves to benefit another animal at a cost to itself, biologists call it altruism. When altruism evolves because the altruist is related to the beneficiary so the altruism-causing gene benefits itself, they call it kin selection. But when we look into the psychology of the animal doing the behaving, we can give the phenomenon another name: love.

The essence of love is feeling pleasure in another's well-being and pain in its harm. These feelings motivate acts that benefit the loved one, like nurturing, feeding, and protecting. We now understand why many animals, including humans, love their children, parents, grandparents, grandchildren, siblings, aunts, uncles, nephews, nieces, and cousins: people helping relatives equals genes helping themselves. The sacrifices made for love are modulated by the degree of relatedness: people make more sacrifices for their children than for their nephews and nieces. They are modulated by the expected reproductive life of the beneficiary: parents sacrifice more for children, who have a longer life ahead of them, than children sacrifice for parents. And they are modulated by the beneficiary's own feelings of love. People love their grandmothers not because their grandmothers are expected to reproduce, but because their grandmothers love *them*, and love the rest of their family. That is, you help people who enjoy helping you and helping your relatives. That is also why men and women fall in love. The other parent of my child has as much of a genetic stake in the child as I do, so what is good for her is good for me.

Many people think that the theory of the selfish gene says that "animals try to spread their genes." That misstates the facts and it misstates the theory. Animals, including most people, know nothing about genetics and care even less. People love their children not because they want to spread their genes (consciously or unconsciously) but because they can't help it. That love makes them try to keep their children warm, fed, and

safe. What is selfish is not the real motives of the person but the metaphorical motives of the genes that built the person. Genes "try" to spread *themselves* by wiring animals' brains so the animals love their kin and try to keep warm, fed, and safe.

The confusion comes from thinking of people's genes as their true self, and the motives of their genes as their deepest, truest, unconscious motives. From there it's easy to draw the cynical and incorrect moral that all love is hypocritical. That confuses the real motives of the person with the metaphorical motives of the genes. Genes are not puppetmasters; they acted as the recipe for making the brain and body and then they got out of the way. They live in a parallel universe, scattered among bodies, with their own agendas.

Most discussions of the biology of altruism are really not about the biology of altruism. It's easy to see why nature documentaries, with their laudable conservationist ethic, disseminate the agitprop that animals act in the interests of the group. One subtext is, Don't hate the wolf that just ate Bambi; he's acting for the greater good. The other is, Protecting the environment is nature's way; we humans had better shape up. The opposing theory of the selfish gene has been bitterly attacked out of the fear that it vindicates the philosophy of Gordon Gekko in *Wall Street*: greed is good, greed works. Then there are those who believe in selfish genes but urge us to face up to the sad truth: at heart, Mother Teresa is really selfish.

I think moralistic science is bad for morals and bad for science. Surely paving Yosemite is unwise, Gordon Gekko is bad, and Mother Teresa is good regardless of what came out in the latest biology journals. But I suppose it is only human to feel *afrisson* when learning about what made us what we are. So I offer a more hopeful way of reflecting on the selfish gene.

The body is the ultimate barrier to empathy. Your toothache simply does not hurt me the way it hurts you. But genes are not imprisoned in bodies; the same gene lives in the bodies of many family members at once. The dispersed copies of a gene call to one another by endowing bodies with emotions. Love, compassion, and empathy are invisible fibers that connect genes in different bodies. They are the closest we will ever come to feeling someone else's toothache. When a parent wishes

she could take the place of a child about to undergo surgery, it is not the species or the group or her body that wants her to have that most unselfish emotion; it is her selfish genes.

Animals are nice not just to their relatives. The biologist Robert Trivers developed a suggestion from George Williams on how another kind of altruism could evolve (where altruism, again, is defined as behavior that benefits another organism at a cost to the behaver). Dawkins explains it with a hypothetical example. Imagine a species of bird that suffers from a disease-carrying tick and must spend a good deal of time removing them with its beak. It can reach every part of its body but the top of its head. Every bird would benefit if some other bird groomed its head. If the birds in a group all responded to the sight of a head presented to them by grooming it, the group would prosper. But what would happen if a mutant presented its head for grooming but never groomed anyone else? These freeloaders would be parasite-free, *and* could use the time they saved not grooming others to look for food. With that advantage they would eventually dominate the population, even if it made the group more vulnerable to extinction. The psychologist Roger Brown explains, "One can imagine a pathetic final act in which all birds on stage present to one another heads that none will groom."

But say a different, grudge-bearing mutant arose. This mutant groomed strangers, groomed birds that in the past had groomed it, but refused to groom birds that had refused to groom it. Once a few of them had gained a toehold, these grudgers could prosper, because they would groom one another and not pay the costs of grooming the cheaters. And once they were established, neither indiscriminate groomers nor cheaters could drive them out, though in some circumstances cheaters could lurk as a minority.

The example is hypothetical, illustrating how altruism among non-kin—what Trivers called reciprocal altruism—can evolve. It is easy to confuse the thought experiment with a real observation; Brown remarks, "When I have used the example in teaching, it has sometimes come back to me on exams as a real bird, often as 'Skinner's pigeons,' sometimes the black-headed gull, and once the robin." Some species do practice reciprocal altruism, but not many, because it evolves only under special condi-

tions. An animal must be able to grant a large benefit to another at a small cost to itself, and the roles must commonly reverse. The animals must devote part of their brains to recognizing each other as individuals (see Chapter 2), and, if repayment comes long after the favor, to remembering who helped them and who refused, and to deciding how to grant and withhold favors accordingly.

Humans are, of course, a brainy species, and are zoologically unusual in how often they help unrelated individuals (Chapter 3). Our lifestyles and our minds are particularly adapted to the demands of reciprocal altruism. People have food, tools, help, and information to trade. With language, information is an ideal trade good because its cost to the giver—a few seconds of breath—is minuscule compared with the benefit to the recipient. Humans are obsessed with individuals; remember the Blick twins from Chapter 2, one of whom bit a police officer but neither of whom could be punished because each benefited from reasonable doubt that he and not his twin did the deed. And the human mind is equipped with goal-setting demons that regulate the doling out of favors; as with kin-directed altruism, reciprocal altruism is behaviorist shorthand for a set of thoughts and emotions. Trivers and the biologist Richard Alexander have shown how the demands of reciprocal altruism are probably the source of many human emotions. Collectively they make up a large part of the moral sense.

The minimal equipment is a cheater-detector and a tit-for-tat strategy that begrudges a gross cheater further help. A gross cheater is one who refuses to reciprocate at all, or who returns so little that the altruist gets back less than the cost of the initial favor. Recall from Chapter 5 that Cosmides has shown that people do reason unusually well about cheaters. But the real intrigue begins with Trivers' observation that there is a more subtle way to cheat. A subtle cheater reciprocates enough to make it worth the altruist's while, but returns less than he is capable of giving, or less than the altruist would give if the situation were reversed. That puts the altruist in an awkward position. In one sense she is being ripped off. But if she insists on equity, the subtle cheater could break off the relationship altogether. Since half a loaf is better than none, the altruist is trapped. She does have one kind of leverage, though. If there are *other* trading partners in the group who don't cheat at all, or who cheat subtly but less stingily, she can give them her business instead.

The game has become more complicated. Selection favors cheating

when the altruist will not find out or when she will not break off her altruism if she does find out. That leads to better cheater-detectors, which leads to more subtle cheating, which leads to detectors for more subtle cheating, which leads to tactics to get away with subtle cheating without being detected by the subtle-cheater-detectors, and so on. Each detector must trigger an emotion demon that sets up the appropriate goal—continuing to reciprocate, breaking off the relationship, and so on.

Here is how Trivers reverse-engineered the moralistic emotions as strategies in the reciprocity game. (His assumptions about the causes and consequences of each emotion are well supported by the literature in experimental social psychology and by studies of other cultures, though they are hardly necessary, as real-life examples no doubt will flood into mind.)

Liking is the emotion that initiates and maintains an altruistic partnership. It is, roughly, a willingness to offer someone a favor, and is directed to those who appear willing to offer favors back. We like people who are nice to us, and we are nice to people whom we like.

Anger protects a person whose niceness has left her vulnerable to being cheated. When the exploitation is discovered, the person classifies the offending act as unjust and experiences indignation and a desire to respond with moralistic aggression: punishing the cheater by severing the relationship and sometimes by hurting him. Many psychologists have remarked that anger has moral overtones; almost all anger is righteous anger. Furious people feel they are aggrieved and must redress an injustice.

Gratitude calibrates the desire to reciprocate according to the costs and benefits of the original act. We are grateful to people when their favor helps us a lot and has cost them a lot.

Sympathy, the desire to help those in need, may be an emotion for earning gratitude. If people are most grateful when they most need the favor, a person in need is an opportunity to make an altruistic act go farthest.

Guilt can rack a cheater who is in danger of being found out. H. L. Mencken defined *conscience* as "the inner voice which warns us that someone might be looking." If the victim responds by cutting off all future aid, the cheater will have paid dearly. He has an interest! in preventing the rupture by making up for the misdeed and keeping it from happening again. People feel guilty about private transgressions because

they may become public; confessing a sin before it is discovered is evidence of sincerity and gives the victim better grounds to maintain the relationship. *Shame*, the reaction to a transgression after it has been discovered, evokes a public display of contrition, no doubt for the same reason.

Lily Tomlin said, "I try to be cynical, but it's hard to keep up." Trivers notes that once these emotions evolved, people had an incentive to mimic them to take advantage of other people's reactions to the real thing. Sham generosity and friendship may induce genuine altruism in return. Sham moral anger when no real cheating took place may nonetheless win reparations. Sham guilt may convince a wronged party that the cheater has reformed his ways, even if cheating is about to resume. Feigning dire straits may evoke genuine sympathy. Sham sympathy which gives the appearance of helping may elicit real gratitude. Sham gratitude may mislead an altruist into expecting a favor to be reciprocated. Trivers notes that none of this hypocrisy need be conscious; indeed, as we shall see, it is most effective when it is not.

The next round in this evolutionary contest is, of course, developing an ability to discriminate between real emotions and sham emotions. We get the evolution of *trust* and *distrust*. When we see someone going through the motions of generosity, guilt, sympathy, or gratitude rather than showing signs of the genuine emotion, we lose the desire to cooperate. For example, if a cheater makes amends in a calculating manner rather than out of credible guilt, he may cheat again when circumstances allow him to get away with it. The search for signs of trustworthiness makes us into mind readers, alert for any twitch or inconsistency that betrays a sham emotion. Since hypocrisy is easiest to expose when people compare notes, the search for trustworthiness makes us avid consumers of gossip. In turn, our reputation becomes our most valuable possession, and we are motivated to protect (and inflate) it with conspicuous displays of generosity, sympathy, and integrity and to take umbrage when it is impugned.

Are you keeping up? The ability to guard against sham emotions can in turn be used as a weapon against real emotions. One can protect one's own cheating by imputing false motives to someone else—by saying that a person really isn't aggrieved, friendly, grateful, guilty, and so on, when she really is. No wonder Trivers was the first to propose that the expansion of the human brain was driven by a cognitive arms race, fueled by the emotions needed to regulate reciprocal altruism.

Like kin selection, reciprocal altruism has been condemned as painting, even condoning, a bleak picture of human motives. Is sympathy nothing but a cheap way to buy gratitude? Is niceness just a business tactic? Not at all. Go ahead and think the worst about the sham emotions. But the reason the real ones are felt is not that they are hoped to help the feeler; it is that they in fact helped the feeler's ancestors. And it's not just that you shouldn't visit the iniquities of the fathers upon the children; the fathers may never have been iniquitous to begin with. The first mutants who felt sympathy and gratitude may have prospered not by their own calculation but because the feelings made it worth their neighbors' while to cooperate with them. The emotions themselves may have been kind and heartfelt in every generation; indeed, once sham-emotion-detectors evolved, they would be most effective when they *are* kind and heartfelt. Of course, the genes are metaphorically selfish in endowing people with beneficent emotions, but who cares about the moral worth of deoxyribonucleic acid?

Many people still resist the idea that the moral emotions are designed by natural selection to further the long-term interests of individuals and ultimately their genes. Wouldn't it be better for everyone if we were built to enjoy what was best for the group? Companies wouldn't pollute, public service unions wouldn't strike, citizens would recycle bottles and take the bus, and those teenagers would stop ruining a quiet Sunday afternoon with their jet-skis.

Once again I think it is unwise to confuse how the mind works with how it would be nice for the mind to work. But perhaps some comfort may be taken in a different way of looking at things. Perhaps we should *rejoice* that people's emotions aren't designed for the good of the group. Often the best way to benefit one's group is to displace, subjugate, or annihilate the group next door. Ants in a colony are closely related, and each is a paragon of unselfishness. That's why ants are one of the few kinds of animal that wage war and take slaves. When human leaders have manipulated or coerced people into submerging their interests into the group's, the outcomes are some of history's worst atrocities. In *Love and Death*, Woody Allen's pacifist character is urged to defend the czar and Mother Russia with the dubious call to duty that under French rule he would have to eat croissants and rich food with heavy sauces. People's

desire for a comfortable life for themselves, their family, and their friends may have braked the ambitions of many an emperor.

THE DOOMSDAY MACHINE

It is 1962, and you are the president of the United States. You have just learned that the Soviet Union has dropped an atomic bomb on New York. You know they will not attack again. In front of you is the phone to the Pentagon, the proverbial button, with which you can retaliate by bombing Moscow.

You are about to press the button. The nation's policy is to retaliate in kind against a nuclear attack. The policy was designed to deter attackers; if you don't follow through, the deterrent would have been a sham.

On the other hand, you are thinking, the damage has been done. Killing millions of Russians will not bring millions of dead Americans back to life. The bomb will add radioactive fallout to the atmosphere, harming your own citizens. And you will go down in history as one of the worst mass murderers of all time. Retaliation now would be sheer spite.

But then, it is precisely this line of thinking that emboldened the Soviets to attack. They *knew* that once the bomb fell you would have nothing to gain and much to lose by retaliating. They thought they were calling your bluff. So you had better retaliate to show them it wasn't a bluff.

But then again, what's the point of proving *now* that you weren't bluffing *then*? The present cannot affect the past. The fact remains that if you push the button, you will snuff out millions of lives for no reason.

But wait—the Soviets knew you would think it is pointless to prove you weren't bluffing after they tried to call your bluff. That's why they called your bluff. The very fact that you are thinking this way brought on the catastrophe—so you shouldn't think this way.

But not thinking this way *now* is too late . . .

You curse your freedom. Your predicament is that you have the choice to retaliate, and since retaliating is not in your interests, you may decide not to do it, exactly as the Soviets anticipated. If only you didn't *have* the choice! If only your missiles had been wired to a reliable nuclear-fireball-detector and went off automatically. The Soviets would not have dared to

attack, because they would have known retaliation was certain.

This train of reasoning was taken to its logical conclusion in the novel and film *Dr. Strangelove*. A deranged American officer has ordered a nuclear bomber to attack the Soviet Union, and it cannot be recalled. The president and his advisors meet in the war room with the Soviet ambassador to persuade him, and by telephone the Soviet leader, that the imminent attack is an accident and that the Soviets should not retaliate. They learn it is too late. The Soviets had installed the Doomsday Machine: a network of underground nuclear bombs that is set off automatically if the country is attacked or if anyone tries to disarm it. The fallout will destroy all human and animal life on earth. They installed the machine because it was cheaper than pinpoint missiles and bombers, and because they feared the United States might be building one and wanted to prevent a Doomsday gap. President Muffley (played by Peter Sellers) confers with the country's top nuclear strategist, the brilliant Dr. Strangelove (played by Peter Sellers):

"But," Muffley said, "is it really possible for it to be triggered automatically and at the same time impossible to untrigger?"

. . . Doctor Strangelove said quickly, "But precisely. Mister President, it is not only possible, it is essential. That is the whole idea of this machine. Deterrence is the art of producing in the enemy the fear to attack. And so because of the automated and irrevocable decision-making process which rules out human meddling, the Doomsday Machine is terrifying, simple to understand, and completely credible and convincing." . . .

President Muffley said, "But this is fantastic, Doctor Strangelove. How can it be triggered automatically?"

Strangelove said, "Sir, it is remarkably simple to do that. When you merely wish to bury bombs there is no limit to the size. . . . After they are buried they are connected to a gigantic complex of computers. A specific and closely defined set of circumstances under which the bombs are to be exploded is programmed into the tape memory banks. . . ." Strangelove turned so he looked directly at [the Soviet Ambassador]. "There is only one thing I don't understand, Mister Ambassador. The whole point of the Doomsday Machine is lost if you keep it a secret. Why didn't you tell the world?"

[The ambassador] turned away. He said quietly but distinctly, "It was to be announced at the Party Congress on Monday. As you know, the Premier loves surprises."

The German-accented, leather-gloved, wheelchair-bound Dr. Strangelove, with his disconcerting tic of giving the Nazi salute, is one of cinema's all-time eeriest characters. He was meant to symbolize a kind of intellectual who until recently was prominent in the public's imagination: the nuclear strategist, paid to think the unthinkable. These men, who included Henry Kissinger (on whom Sellers based his portrayal), Herman Kahn, John von Neumann, and Edward Teller, were stereotyped as amoral nerds who cheerfully filled blackboards with equations about megadeaths and mutual assured destruction. Perhaps the scariest thing about them was their paradoxical conclusions—for example, that safety in the nuclear age comes from exposing one's cities and protecting one's missiles.

But the unsettling paradoxes of nuclear strategy apply to *any* conflict between parties whose interests are partly competing and partly shared. Common sense says that victory goes to the side with the most intelligence, self-interest, coolness, options, power, and clear lines of communication. Common sense is wrong. Each of these assets can be a liability in contests of strategy (as opposed to contests of chance, skill, or strength), where behavior is calculated by predicting what the other guy will do in response. Thomas Schelling has shown that the paradoxes are ubiquitous in social life. We shall see that they offer great insight into the emotions, particularly the headstrong passions that convinced the Romantics that emotion and reason were opposites. But first let's put the emotions aside and just examine the logic of conflicts of strategy.

Take bargaining. When two people haggle over a car or a house, a bargain is struck when one side makes the final concession. Why does he concede? Because he is sure she will not. The reason she won't concede is that she thinks he will concede. She thinks he will because she thinks he thinks she thinks he will. And so on. There always is a range of prices that the buyer and seller would both accept. Even if a particular price within that range is not the best price for one party, it is preferable to canceling the deal outright. Each side is vulnerable to being forced to settle for the worst acceptable price because the other side realizes that he or she would have no choice if the alternative was to reach no agreement at all. But when both parties can guess the range, *any* price within the range is a point from which at least one party would have been willing to back off, and the other party knows it.

Schelling points out that the trick to coming out ahead is "a voluntary but irreversible sacrifice of freedom of choice." How do you persuade

someone that you will not pay more than \$16,000 for a car that is really worth \$20,000 to you? You can make a public, enforceable \$5,000 bet with a third party that you won't pay more than \$16,000. As long as \$16,000 gives the dealer a profit, he has no choice but to accept. Persuasion would be futile; it's against your interests to compromise. By tying your own hands, you improve your bargaining position. The example is fanciful, but real ones abound. The dealer appoints a salesperson who is not authorized to sell at less than a certain price even if he says he wants to. A homebuyer cannot get a mortgage if the bank's appraiser says he paid too much. The homebuyer exploits that powerlessness to get a better price from the seller.

Not only can power be a liability in conflicts of strategy, communication can be, too. When you are haggling from a pay phone with a friend about where to meet for dinner, you can simply announce that you will be at Ming's at six-thirty and hang up. The friend has to accede if she wants to meet you at all.

Paradoxical tactics also enter into the logic of promises. A promise can secure a favor only when the beneficiary of the promise has good reason to believe it will be carried out. The promiser is thus in a *better* position when the beneficiary knows that the promiser is *bound* by his promise. The law gives companies the right to sue and the right to be sued. The right to be sued? What kind of "right" is that? It is a right that confers the power to make a promise: to enter into contracts, borrow money, and engage in business with someone who might be harmed as a result. Similarly, the law that empowers banks to foreclose on a mortgage makes it worth the bank's while to grant the mortgage, and so, paradoxically, benefits the *borrower*. In some societies, Schelling notes, eunuchs got the best jobs because of what they could not do. How does a hostage persuade his kidnapper not to kill him to prevent him from identifying the kidnapper in court? One option is to deliberately blind himself. A better one is to confess to a shameful secret that the kidnapper can use as blackmail. If he has no shameful secret, he can create one by having the kidnapper photograph him in some unspeakably degrading act.

Threats, and defenses against threats, are the arena in which Dr. Strangelove really comes into his own. There are boring threats, in which the threatener has an interest in carrying out the threat—for example, when a homeowner threatens a burglar that she will call the police. The fun begins when carrying out the threat is costly to the threatener, so its value is only as a deterrent. Again, freedom is costly; the threat is credible

only when the threatener has no choice but to carry it out and the target knows it. Otherwise, the target can threaten the threatener right back by refusing to comply. The Doomsday Machine is an obvious example, though the secrecy defeated its purpose. A hijacker who threatens to blow up a plane if anyone tries to disarm him will have a better chance of seeing Cuba if he wears explosives that go off with the slightest jostling. A good way to win the teenagers' game of chicken, in which two cars approach each other at high speed and the first driver to swerve loses face, is to conspicuously remove your own steering wheel and throw it away.

With threats, as with promises, communication can be a liability. The kidnapper remains incommunicado after making the ransom demand so he cannot be persuaded to give up the hostage for a smaller ransom or a safe escape. Rationality is also a liability. Schelling points out that "if a man knocks at the back door and says that he will stab himself unless you give him \$10, he is more likely to get the \$10 if his eyes are bloodshot." Terrorists, kidnappers, hijackers, and dictators of small countries have an interest in appearing mentally unbalanced. An absence of self-interest is also an advantage. Suicide bombers are almost impossible to stop.

To defend yourself *against* threats, make it impossible for the threatener to make you an offer you can't refuse. Again, freedom, information, and rationality are handicaps. "Driver does not know combination to safe," says the sticker on the delivery truck. A man who is worried that his daughter may be kidnapped can give away his fortune, leave town and remain incommunicado, lobby for a law that makes it a crime to pay ransom, or break the hand with which he signs checks. An invading army may burn bridges behind it to make retreat impossible. A college president tells protesters he has no influence on the town police, and genuinely wants no influence. A racketeer cannot sell protection if the customer makes sure he is not at home when the racketeer comes around.

Because an expensive threat works both ways, it can lead to a cycle of self-incapacitation. Protesters attempt to block the construction of a nuclear power plant by lying down on the railroad tracks leading to the site. The engineer, being reasonable, has no choice but to stop the train. The railroad company counters by telling the engineer to set the throttle so that the train moves very slowly and then to jump out of the train and walk beside it. The protesters must scramble. Next time the protesters handcuff themselves to the tracks; the engineer does not dare leave the train. But the protesters must be certain the engineer sees them in

enough time to stop. The company assigns the next train to a nearsighted engineer.

in these examples, many of them from Schelling, the paradoxical power comes from a physical constraint like handcuffs or an institutional constraint like the police. But strong passions can do the same thing. Say a bargainer publicly announces that he will not pay more than \$16,000 for the car, and everyone knows he could not tolerate the shame of going back on his word. The unavoidable shame is as effective as the enforceable bet, and he will get the car at his price. If Mother Teresa offered to sell you her car, you would not insist on a guarantee because presumably she is constitutionally incapable of cheating you. The hothead who can figuratively explode at any moment enjoys the same tactical advantage as the hijacker who can literally explode at any moment. In *The Maltese Falcon*, Sam Spade (Humphrey Bogart) dares the henchmen of Kasper Gutman (Sidney Greenstreet) to kill him, knowing that they need him to retrieve the falcon. Gutman replies, "That's an attitude, sir, that calls for the most delicate judgment on both sides, because as you know, sir, in the heat of action men are likely to forget where their best interests lie, and let their emotions carry them away." In *The Godfather*, Vito Corleone tells the heads of the other crime families, "I'm a superstitious man. And if some unlucky accident should befall my son, if my son is struck by a bolt of lightning, I will blame some of the people here."

Dr. Strangelove meets The Godfather. Is passion a doomsday machine? People consumed by pride, love, or rage have lost control. They may be irrational. They may act against their interests. They may be deaf to appeals. (The man running amok calls to mind a doomsday machine that has been set off.) But though this be madness, yet there is method in it. Precisely these sacrifices of will and reason are effective tactics in the countless bargains, promises, and threats that make up our social relations.

The theory stands the Romantic model on its head. The passions are no vestige of an animal past, no wellspring of creativity no enemy of the intellect. The intellect is designed to relinquish control to the passions so that they may serve as guarantors of its offers, promises, and threats against suspicions that they are lowballs, double-crosses, and bluffs. The apparent firewall between passion and reason is not an ineluctable part of the archi-

texture of the brain; it has been programmed in deliberately, because only if the passions are in control can they be credible guarantors.

The doomsday-machine theory has been proposed independently by Schelling, Trivers, Daly and Wilson, the economist Jack Hirshleifer, and the economist Robert Frank. Righteous anger, and the attendant thirst for redress or vengeance, is a credible deterrent if it is uncontrollable and unresponsive to the deterrer's costs. Such compulsions, though useful in the long run, can drive people to fight far out of proportion to the stakes. In 1982 Argentina annexed the British colony of the Falklands, desolate islands with virtually no economic or strategic importance. In earlier decades it might have made sense for Britain to defend them as an immediate deterrent to anyone with designs on the rest of its empire, but at that point there was no empire left to defend. Frank points out that for what they spent to reclaim the islands, Britain could have given each Falklander a Scottish castle and a lifetime pension. But most Britons were proud that they stood up to the Argentinians. The same sense of fairness makes us sue expensively for small amounts or seek a refund for a defective product despite red tape that costs us more in lost wages than the product was worth.

The lust for revenge is a particularly terrifying emotion. All over the world, relatives of the slain fantasize day and night about the bittersweet moment when they might avenge a life with a life and find peace at last. The emotion strikes us as primitive and dreadful because we have contracted the government to settle our scores for us. But in many societies an irresistible thirst for vengeance is one's only protection against deadly raids. Individuals may differ in the resolve with which they will suffer costs to carry out vengeance. Since that resolve is an effective deterrent only if it is advertised, it is accompanied by the emotion traditionally referred to as honor: the desire to publicly avenge even minor trespasses and insults. The hair-trigger of honor and revenge can be tuned to the degree of threat in the environment. Honor and vengeance are raised to godly virtues in societies that lie beyond the reach of law enforcement, such as remote horticulturalists and herders, the pioneers of the Wild West, street gangs, organized crime families, and entire nation-states when dealing with one another (in which case the emotion is called "patriotism"). But even within a modern state society where it serves no purpose, the emotion of vengeance cannot easily be turned off. Most legal theories, even from the highest-minded philosophers, acknowledge that retribution is one of the legitimate goals of criminal punishment,

over and above the goals of deterring potential criminals and incapacitating, deterring, and rehabilitating the offender. Enraged crime victims, long disenfranchised from the American legal system, have recently pressed for a say in plea-bargaining and sentencing decisions.

As Strangelove explained, the whole point of a doomsday machine is lost if you keep it a secret. That principle may explain one of the longest-standing puzzles of the emotions: why we advertise them on our face.

Darwin himself never argued that facial expressions were naturally selected adaptations. In fact, his theory was downright Lamarckian. Animals have to move their faces for practical reasons: they bare the teeth to bite, widen the eyes for a panoramic view, and pull back the ears to protect them in a fight. These measures turned into habits that the animal performed when it merely anticipated an event. The habits were then passed to their offspring. It may seem strange that Darwin was no Darwinian in one of his most famous books, but remember that Darwin was fighting on two fronts. He had to explain adaptations to satisfy his fellow biologists, but he also made much of pointless features and animal vestiges in humans to combat creationists, who argued that functional design was a sign of God's handiwork. If God had really designed humans from scratch, Darwin asked, why would he have installed features that are useless to us but similar to features that are useful to animals?

Many psychologists still can't understand why broadcasting one's emotional state might be beneficial. Wouldn't the proverbial smell of fear just egg on one's enemies? One psychologist has tried to revive an old idea that facial muscles are tourniquets that send more blood to the parts of the brain that have to cope with the current challenge. Aside from being hydraulically improbable, the theory cannot explain why we are more expressive when there are other people around.

But if the passionate emotions are guarantors of threats and promises, advertising is their reason for being. But here a problem arises. Remember that real emotions create a niche for sham emotions. Why whip yourself into a rage when you can *simulate* a rage, deter your enemies, and not pay the price of pursuing dangerous vengeance if it fails? Let *others* be doomsday machines, and you can reap the benefits of the terror they sow. Of course, when counterfeit facial expressions begin to drive out

the real ones, people call each other's bluffs, and the facial expressions, real and fake, become worthless.

Facial expressions are useful only if they are hard to fake. As a matter of fact, they *are* hard to fake. People don't really believe that the grinning flight attendant is happy to see them. That is because a social smile is formed with a different configuration of muscles from the genuine smile of pleasure. A social smile is executed by circuits in the cerebral cortex that are under voluntary control; a smile of pleasure is executed by circuits in the limbic system and other brain systems and is involuntary. Anger, fear, and sadness, too, recruit muscles that can't be controlled voluntarily, and the genuine expressions are hard to fake, though we can pantomime an approximation. Actors must simulate facial expressions for a living, but many cannot avoid a mannered look. Some great actors, like Laurence Olivier, are highly coordinated athletes who have doggedly learned to control every muscle. Others learn method acting, inspired by Konstantin Stanislavsky, in which actors make themselves *feel* an emotion by remembering or imagining a charged experience, and the expression pops on the face reflexively.

The explanation is incomplete, because it raises another question: *why* did we never evolve the ability to control our expressions? You can't just say that it would hurt everyone if counterfeit expressions were circulated. True enough, but in a world of honest emoters the faker would prosper, so fakers should always drive out emoters. I don't know the answer, but there are obvious places to look. Zoologists worry about the same problem: how can honest animal signals, like cries, gestures, and advertisements of health, evolve in a world of would-be fakers? One answer is that honest signals can evolve if they are too expensive to fake. For example, only a healthy peacock can afford a splendid tail, so healthy peacocks bear the burden of a cumbersome tail as a display of conspicuous consumption that only they can afford. When the healthiest peacocks display, the less healthy ones have no choice but to follow, because if they hide their health altogether the peahens will assume the worst, namely that they are at death's door.

Is there anything about emotional expressions that would make it inherently costly to put them under voluntary control? Here is a guess. In designing the rest of the human, natural selection had good engineering reasons to segregate the voluntary, cognitive systems from the systems that control housekeeping and physical-plant functions such as the

regulation of heartbeat, breathing rate, blood circulation, sweat, tears, and saliva. None of your conscious beliefs are pertinent to how fast your heart ought to beat, so there's no point in letting you control it. In fact, it would be downright dangerous, since you might forget to pump when you got distracted, or you might try out your own harebrained ideas on what the best pulse rate should be.

Now, say selection handcuffed each emotion to a physiological control circuit, and the activity of the circuit was visible to an observer as flushing, blushing, blanching, sweating, trembling, quavering, croaking, weeping, and the facial reflexes Darwin discussed. An observer would have good reason to believe that the emotion was genuine, since a person could not fake it unless he had voluntary control of his heart and other organs. Just as the Soviets would have wanted to show everyone the wiring of the Doomsday Machine *to prove* that it was automatic and irreversible and their description of it no bluff, people might have an interest in showing everyone that an emotion is holding their body hostage and their angry words are no bluff. If so, it would explain why emotions are so intimately tied to the body, a fact that puzzled William James and a century of psychologists after him.

The handcuffing may have been easy for natural selection, because the major human emotions seem to have grown out of evolutionary precursors (anger from fighting, fear from fleeing, and so on), each of which engaged a suite of involuntary physiological responses. (This might be the grain of truth in the Romantic and triune-brain theories: modern emotions may *exploit* the involuntariness of older reflexes, even if they did not inherit it by default.) And once the handcuffs were in place for honest emoters, everyone else would have had little choice but to don them too, like the unhealthy peacocks forced to muster tails. A chronic poker face would suggest the worst: that the emotions a person declares in word and deed are shams.

This theory is unproven, but no one can deny the phenomenon. People are vigilant for sham emotions and put the most faith in involuntary physiological giveaways. That underlies an irony of the telecommunications age. Long-distance phone service, electronic mail, faxes, and videoconferencing should have made the face-to-face business meeting obsolete. But meetings continue to be a major expense for corporations and support entire industries like hotels, airlines, and rental cars. Why do we insist on doing business in the flesh? Because we do not trust someone until we see what makes him sweat.

FOOLS FOR LOVE

Why does romantic love leave us bewitched, bothered, and bewildered? Could it be another paradoxical tactic like handcuffing oneself to railroad tracks? Quite possibly. Offering to spend your life and raise children with someone is the most important promise you'll ever make, and a promise is most credible when the promiser can't back out. Here is how the economist Robert Frank has reverse-engineered mad love.

Unsentimental social scientists and veterans of the singles scene agree that dating is a marketplace. People differ in their value as potential marriage partners. Almost everyone agrees that Mr. or Ms. Right should be good-looking, smart, kind, stable, funny, and rich. People shop for the most desirable person who will accept them, and that is why most marriages pair a bride and a groom of approximately equal desirability. Mate-shopping, however, is only part of the psychology of romance; it explains the statistics of mate choice, but not the final pick.

Somewhere in this world of five billion people there lives the best-looking, richest, smartest, funniest, kindest person who would settle for you. But your dreamboat is a needle in a haystack, and you may die single if you insist on waiting for him or her to show up. Staying single has costs, such as loneliness, childlessness, and playing the dating game with all its awkward drinks and dinners (and sometimes breakfasts). At some point it pays to set up house with the best person you have found so far.

But that calculation leaves your partner vulnerable. The laws of probability say that someday you will meet a more desirable person, and if you are always going for the best you can get, on that day you will dump your partner. But your partner has invested money, time, childrearing, and forgone opportunities in the relationship. If your partner was the most desirable person in the world, he or she would have nothing to worry about, because you would never want to desert. But failing that, the partner would have been foolish to enter the relationship.

Frank compares the marriage market with the rental market. Landlords desire the best of all tenants but settle for the best they can find, and renters want the best of all apartments but settle for the best they can find. Each invests in the apartment (the landlord may paint it the tenant's favorite color; the tenant may install permanent decorations), so

each would be harmed if the other suddenly terminated the agreement. If the tenant could leave for a better flat, the landlord would have to bear the costs of an unrented unit and the search for a new tenant; he would have to charge a high rent to cover that risk, and would be loath to paint. If the landlord could evict the tenant for a better one, the tenant would have to search for a new home; she would be willing to pay only a low rent, and would not bother to keep the apartment in good shape, if she had to expose herself to that risk. If the best tenant were renting the best apartment, the worries would be moot; neither would want to end the arrangement. But since both have to compromise, they protect themselves by signing a lease that is expensive for either to break. By agreeing to restrict his own freedom to evict, the landlord can charge a higher rent. By agreeing to restrict her own freedom to leave, the tenant can demand a lower rent. Lack of choice works to each one's advantage.

Marriage laws work a bit like leases, but our ancestors had to find some way to commit themselves before the laws existed. How can you be sure that a prospective partner won't leave the minute it is rational to do so—say, when a 10-out-of-10 moves in next door? One answer is, don't accept a partner who wanted you for rational reasons to begin with; look for a partner who is committed to staying with you because you are you. Committed by what? Committed by an emotion. An emotion that the person did not decide to have, and so cannot decide not to have. An emotion that was not triggered by your objective mate-value and so will not be alienated by someone with greater mate-value. An emotion that is guaranteed not to be a sham because it has physiological costs like tachycardia, insomnia, and anorexia. An emotion like romantic love.

"People who are sensible about love are incapable of it," wrote Douglas Yates. Even when courted by the perfect suitor, people are unable to will themselves to fall in love, often to the bewilderment of the matchmaker, the suitor, and the person himself or herself. Instead it is a glance, a laugh, a manner that steals the heart. Remember from Chapter 2 that spouses of one twin are not attracted to the other; we fall in love with the individual, not with the individual's qualities. The upside is that when Cupid does strike, the lovestruck one is all the more credible in the eyes of the object of desire. Murmuring that your lover's looks, earning power, and IQ meet your minimal standards would probably kill the romantic mood, even though the statement is statistically true. The way to a person's heart is to declare the opposite—that you're in love because you can't help it. Tipper Gore's Parents' Music Resource Center notwith-

standing, the sneering, body-pierced, guitar-smashing rock musician is typically not singing about drugs, sex, or Satan. He is singing about love. He is courting a woman by calling attention to the irrationality, uncontrollability, and physiological costs of his desire. I want you so bad, it's driving me mad, Can't eat, can't sleep, Heart beats like a big bass drum, You're the only one, Don't know why I love you like I do, You drive me crazy, Can't stop lovin'you, Ain't nobody can do it to me the way you can, I like the way you walk, I like the way you talk, et cetera, et cetera.

Of course, one can well imagine a woman not being swept off her feet by these proclamations. (Or a man, if it is a woman doing the declaring.) They set off a warning light in the other component of courtship, smart shopping. Groucho Marx said that he would not belong to any club that would have him as a member. Usually people do not want any suitor who wants them too badly too early, because it shows that the suitor is desperate (so they should wait for someone better), and because it shows that the suitor's ardor is too easily triggered (hence too easily triggerable by someone else). The contradiction of courtship—flaunt your desire while playing hard to get—comes from the two parts of romantic love: setting a minimal standard for candidates in the mate market, and capriciously committing body and soul to one of them.

THE SOCIETY OF FEELINGS

Mental life often feels like a parliament within. Thoughts and feelings vie for control as if each were an agent with strategies for taking over the whole person, you. Might our mental agents use paradoxical tactics with one another—handcuffs, doomsday machines, unbreakable contracts with third parties? The analogy is imperfect because natural selection designs people to compete but does not design organs, including mental agents, to compete; the interests of the whole person are paramount. But the whole person has many goals, like food, sex, and safety, and that requires a division of labor among mental agents with different priorities and kinds of expertise. The agents are bound by an entente that benefits the whole person over a lifetime, but over the short term the agents may outwit one another with devious tactics.

Self-control is unmistakably a tactical battle between parts of the mind. Schelling observes that the tactics people use to control them-

selves are interchangeable with the tactics they use to control; others. How do you prevent your child from scratching his hives in his sleep? Put mittens on him. How do you prevent yourself from scratching your hives in your sleep? Put mittens on yourself. If Odysseus had not plugged his shipmates' ears, they would have done it on their own. The self that wants a trim body outwits the self that wants dessert by throwing out the brownies at the opportune moment when it is in control.

So we do seem to use paradoxical tactics against ourselves. The agent in control at one time makes a voluntary but irreversible sacrifice of freedom of choice for the whole body, and gets its way in the long run. That is the bright spot in this whole depressing discussion of selfish genes and doomsday machines. Social life is not always the equivalent of global thermonuclear war because the part of us with the longest view of the future, when in control of the body, can voluntarily sacrifice freedom of choice for the body at other times. We sign contracts, submit to laws, and hitch our reputations to public declarations of loyalty to friends and mates. These are not tactics to defeat someone else, but tactics to defeat the darker parts of ourselves.

One more speculation on the battle inside the head. No one knows what, if anything, grief is for. Obviously the loss of a loved one is unpleasant, but why should it be devastating? Why the debilitating pain that stops people from eating, sleeping, resisting diseases, and getting on with life? Jane Goodall describes a young chimp, Flint, who after the death of his beloved mother became depressed and died himself as if of a broken heart.

Some have suggested that grief is an enforced interlude for reassessment. Life will never be the same, so one must take time to plan how to cope with a world that has been turned upside down. Perhaps grief also gives people time to contemplate how a lapse of theirs may have allowed the death and how they might be more careful in the future. There may be an element of truth to the suggestion. Bereaved people find that they ache all over again every time they discover another habit to unlearn, like setting out an extra plate or buying groceries for two. And blaming oneself is a common symptom. But the pain of grief makes planning harder, not easier, and is too extreme and long-lasting to be useful as a strategy session.

William James wrote, "It takes a mind debauched by learning to carry the process of making the natural seem strange so far as to ask for the 'why' of any instinctive human act." Though legitimate to a scientist, the question "Why do we grieve?" is preposterous to common sense. If

you didn't grieve when someone died, could you really have loved him when he was alive? It's logically possible but seems psychologically impossible; grief is the other side of love. And there may lie the answer. Perhaps grief is an internal doomsday machine, pointless once it goes off, useful only as a deterrent. What parents have not lain awake contemplating the horror of losing a child? Or worried themselves sick with awful images when a child is late or lost? These thoughts are powerful reminders to protect and cherish a loved one in the face of myriad other demands on one's time and thoughts. Like all deterrents, grief would be effective only if it is certain and terrible.

KIDDING OURSELVES

The playwright Jerome K. Jerome once said, "It is always the best policy to tell the truth, unless, of course, you are an exceptionally good liar." It's hard to be a good liar, even when it comes to your own intentions, which only you can verify. Intentions come from emotions, and emotions have evolved displays on the face and body. Unless you are a master of the Stanislavsky method, you will have trouble faking them; in fact, they probably evolved *because* they were hard to fake. Worse, lying is stressful, and anxiety has its own telltale markers. They are the rationale for polygraphs, the so-called lie detectors, and humans evolved to be lie detectors, too. Then there is the annoying fact that some propositions logically entail others. Since *some* of the things you say will be true, you are always in danger of exposing your own lies. As the Yiddish saying goes, a liar must have a good memory.

Trivers, pursuing his theory of the emotions to its logical conclusion, notes that in a world of walking lie detectors the best strategy is to believe your own lies. You can't leak your hidden intentions if you don't think that they *are* your intentions. According to his theory of self-deception, the conscious mind sometimes hides the truth from itself the better to hide it from others. But the truth is useful, so it should be registered somewhere in the mind, walled off from the parts that interact with other people. There is an obvious similarity to Freud's theory of the unconscious and the defense mechanisms of the ego (such as repression, projection, denial, and rationalization), though the explanation is completely different. George Orwell stated it in *1984*: "The secret of ruler-

ship is to combine a belief in one's own infallibility with a power to learn from past mistakes."

The neuroscientist Michael Gazzaniga has shown that the brain blithely weaves false explanations about its motives. Split-brain patients have had their cerebral hemispheres surgically disconnected as a treatment for epilepsy. Language circuitry is in the left hemisphere, and the left half of the visual field is registered in the isolated right hemisphere, so the part of the split-brain person that can talk is unaware of the left half of his world. The right hemisphere is still active, though, and can carry out simple commands presented in the left visual field, like "Walk" or "Laugh." When the patient (actually, the patient's left hemisphere) is asked why he walked out (which we know was a response to the command presented to the right hemisphere), he ingenuously replies, "To get a Coke." When asked why he is laughing, he says, "You guys come up and test us every month. What a way to make a living!"

Our confabulations, not coincidentally, present us in the best light. Literally hundreds of experiments in social psychology say so. The humorist Garrison Keillor describes the fictitious community of Lake Wobegon, "where the women are strong, the men are good-looking, and all the children are above average." Indeed, most people claim they are above average in any positive trait you name: leadership, sophistication, athletic prowess, managerial ability, even driving skill. They rationalize the boast by searching for an *aspect* of the trait that they might in fact be good at. The slow drivers say they are above average in safety, the fast ones that they are above average in reflexes.

More generally, we delude ourselves about how benevolent and how effective we are, a combination that social psychologists call *benefectance*. When subjects play games that are rigged by the experimenter, they attribute their successes to their own skill and their failures to the luck of the draw. When they are fooled in a fake experiment into thinking they have delivered shocks to another subject, they derogate the victim, implying that he deserved the punishment. Everyone has heard of "reducing cognitive dissonance," in which people invent a new opinion to resolve a contradiction in their minds. For example, a person will recall enjoying a boring task if he had agreed to recommend it to others for paltry pay. (If the person had been enticed to recommend the task for generous pay, he accurately recalls that the task was boring.) As originally conceived of by the psychologist Leon Festinger, cognitive dissonance is an unsettled feeling that arises from an inconsistency in one's beliefs.

But that's not right: there is no contradiction between the proposition "The task is boring" and the proposition "I was pressured into lying that the task was fun." Another social psychologist, Eliot Aronson, nailed it down: people doctor their beliefs only to eliminate a contradiction with the proposition "I am nice and in control." Cognitive dissonance is always triggered by blatant evidence that you are not as beneficent and effective as you would like people to think. The urge to reduce it is the urge to get your self-serving story straight.

Sometimes we have glimpses of our own self-deception. When does a negative remark sting, cut deep, hit a nerve? When some part of us knows it is true. If every part knew it was true, the remark would not sting; it would be old news. If no part thought it was true, the remark would roll off; we could dismiss it as false. Trivers recounts an experience that is all too familiar (at least to me). One of his papers drew a published critique, which struck him at the time as vicious and unprincipled, full of innuendo and slander. Rereading the article years later, he was surprised to find that the wording was gentler, the doubts more reasonable, the attitude less biased than he had remembered. Many others have made such discoveries; they are almost the definition of "wisdom."

If there were a verb meaning "to believe falsely," it would not have any significant first person, present indicative.

—Ludwig Wittgenstein

There's one way to find out if a man is honest: ask him; if he says yes, you know he's crooked.

—Mark Twain

Our enemies' opinion of us comes closer to the truth than our own.

—François La Rochefoucauld

Oh wad some power the giftie gie us
To see oursels as ithers see us!

—Robert Burns

No one can examine the emotions without seeing in them the source of much human tragedy. I don't think we should blame the animals; it's

clear enough how natural selection engineered our instincts to suit our needs. We shouldn't blame selfish genes, either. They endow us with selfish motives, but they just as surely endow us with the capacity for love and a sense of justice. What we should appreciate and fear is the cunning designs of the emotions themselves. Many of their specs are not for gladness and understanding: think of the happiness treadmill, the Sirens' song, the sham emotions, the doomsday machines, the caprice of romance, the pointless punishment of grief. But self-deception is perhaps the crudest motive of all, for it makes us feel right when we are wrong and emboldens us to fight when we ought to surrender. Trivers writes,

Consider an argument between two closely bound people, say, husband and wife. Both parties believe that one is an altruist—of long standing, relatively pure in motive, and much abused—while the other is characterized by a pattern of selfishness spread over hundreds of incidents. They only disagree over who is altruistic and who selfish. It is noteworthy that the argument may appear to burst forth spontaneously, with little or no preview, yet as it rolls along, two whole landscapes of information processing appear to lie already organized, waiting only for the lightning of anger to show themselves.

In cartoons and movies, the villains are mustache-twirling degenerates, cackling with glee at their badness. In real life, villains are convinced of their rectitude. Many biographers of evil men start out assuming that their subjects are cynical opportunists and reluctantly discover that they are ideologues and moralists. If Hitler was an actor, concluded one, he was an actor who believed in the part.

Still, thanks to the complexity of our minds, we need not be perpetual dupes of our own chicanery. The mind has many parts, some designed for virtue, some designed for reason, some clever enough to outwit the parts that are neither. One self may deceive another, but every now and then a third self sees the truth.

FAMILY VALUES

Come on, people now, smile on your brother! Everybody get together, try to love one another right now. This is the dawning of the Age of Aquarius: harmony and understanding, sympathy and trust abounding; no more falsehoods or derisions, golden living dreams of visions, mystic crystal revelation, and the mind's true liberation. Imagine no possessions; I wonder if you can. No need for greed or hunger, a brotherhood of man. Imagine all the people sharing all the world. You may say I'm a dreamer, but I'm not the only one. I hope someday you'll join us, and the world will be as one.

Incredible as it may seem, many of us used to believe this treacle. A leading idea of the 1960s and 70s was that mistrust, jealousy, competitiveness, greed, and manipulation were social institutions due for reform. Some people thought they were unnecessary evils, like slavery or the denial of the vote to women. Others thought they were hidebound traditions whose inefficiency had gone unnoticed, as with the genius who figured out that toll bridges could charge a dollar to the traffic going one way instead of fifty cents to the traffic going both ways.

These sentiments came not just from rock musicians but from America's distinguished social critics. In his 1970 book *The Greening of America*, the Yale law professor Charles Reich heralded a nonviolent revolution being led by the college-age generation. The youth of America had evolved a new consciousness, he said. It was less guilty and anxious, nonjudgmental, noncompetitive, nonmaterialistic, affectionate, honest, unmanipulative, unaggressive, communal, and unconcerned with status

and careers. The new consciousness, emerging like flowers through the pavement, was expressed in their music, communes, hitchhiking, drugs, moon-gazing, peace salute, and even their clothing. Bell-bottoms, he said, "give the ankles a special freedom as if to invite dancing right on the street." The new consciousness promised "a higher reason, a more human community, and a new and liberated individual. Its ultimate creation will be a new and enduring wholeness and beauty—a renewed relationship of man to himself, to other men, to society, to nature, and to the land."

Greening sold a million copies in a few months. It was serialized in the *New Yorker* and discussed in a dozen articles in the *New York Times* and in a volume of essays by the leading intellectuals of the day. John Kenneth Galbraith gave it a positive review (though with a caveat expressed in his title: "Who's Minding the Store?"). The book recently came out in a twenty-fifth anniversary edition.

Reich wrote his book in the Yale dining halls, and based it on his conversations with the students there. Those students, of course, were among the most privileged individuals in the history of humanity. With Mom and Dad paying the bills, everyone around them coming from the upper classes, and Ivy League credentials about to launch them into the expanding economy of the 1960s, it was easy to believe that all you need is love. After graduation day, Reich's generation became the Gucci-wearing, Beemer-driving, condo-owning, gourmet-baby-breeding urban professionals of the 1980s and 90s. Universal harmony was a style as ephemeral as the bell-bottoms, a status symbol that distanced them from rednecks, jocks, and the less hip preppies. As the post-60s rock musician Elvis Costello asked, "Was it a millionaire who said 'Imagine no possessions'?"

The Woodstock Nation was not the first Utopian dream to be shattered. The free-love communes of nineteenth-century America collapsed from sexual jealousy and the resentment of both sexes over the leaders' habit of accumulating young mistresses. The socialist Utopias of the twentieth century became repressive empires led by men who collected Cadillacs and concubines. In anthropology, one South Sea island paradise after another has turned out to be nasty and brutish. Margaret Mead said that nonchalant sex made the Samoans satisfied and free of crime; it turned out that the boys tutored one another in rape techniques. She called the Arapesh "gentle"; they were headhunters. She said that the Tshambuli reversed our sex roles, the men wearing curls and

makeup. In fact the men beat their wives, exterminated neighboring tribes, and treated homicide as a milestone in a young man's life which entitled him to wear the face paint that Mead thought was so effeminate.

In *Human Universals*, the anthropologist Donald Brown has assembled the traits that as far as we know are found in all human cultures. They include prestige and status, inequality of power and wealth, property, inheritance, reciprocity, punishment, sexual modesty, sexual regulations, sexual jealousy, a male preference for young women as sexual partners, a division of labor by sex (including more child care by women and greater public political dominance by men), hostility to other groups, and conflict within the group, including violence, rape, and murder. The list should come as no surprise to anyone familiar with history, current events, or literature. There are a small number of plots in the world's fiction and drama, and the scholar Georges Polti claims to have listed them all. More than eighty percent are defined by adversaries (often murderous), by tragedies of kinship or love, or both. In the real world, our life stories are largely stories of conflict: the hurts, guilts, and rivalries inflicted by parents, siblings, children, spouses, lovers, friends, and competitors.

This chapter is about the psychology of social relations. The Age of Aquarius notwithstanding, that means it is largely about inborn motives that put us into conflict with one another. Given that our brains were shaped by natural selection, it could hardly be otherwise. Natural selection is driven by the competition among genes to be represented in the next generation. Reproduction leads to a geometric increase in descendants, and on a finite planet not every organism alive in one generation can have descendants several generations hence. Therefore organisms reproduce, to some extent, at one another's expense. If one organism eats a fish, that fish is no longer available to be eaten by another organism. If one organism mates with a second one, it denies an opportunity at parenthood to a third. Everyone alive today is a descendant of millions of generations of ancestors who lived under these constraints but reproduced nonetheless. That means that all people today owe their existence to having winners as ancestors, and everyone today is designed, at least in some circumstances, to compete.

That does *not* mean that people (or any other animals) house an aggressive urge that must be discharged, an unconscious death wish, a rapacious sex drive, a territorial imperative, a thirst for blood, or the other

ruthless instincts that are often mistakenly equated with Darwinism. In *The Godfather*, Sollozzo says to Tom Hagen, "I don't like violence, Tom. I'm a businessman. Blood is a big expense." Even in the harshest competition, an intelligent organism must be a strategist, assessing whether its goals might best be served by retreat, conciliation, or living and letting live. As I explained in Chapter 5, it is genes, not organisms, that must compete or die; sometimes the genes' best strategy is to design organisms that cooperate, and yes, even smile on their brother and love one another. Natural selection does not forbid cooperation and generosity; it just makes them difficult engineering problems, like stereoscopic vision. The difficulty of building an organism to see in stereo has not prevented natural selection from installing stereo vision in humans, but we would never have come to understand stereo if we thought it just came free with having two eyes and failed to look for the sophisticated neural programs that accomplish it. Similarly, the difficulty of building an organism to cooperate and be generous has not prevented natural selection from installing cooperation and generosity in humans, but we will never understand these capacities if we think they just come free with living in groups. The on-board computers of social organisms, especially of humans, should run sophisticated programs that assess the opportunities and risks at hand and compete or cooperate accordingly.

The conflict of interest among the members of a species also does not call for a conservative political agenda, as journalists and social scientists often fear. Some worry that if our motives put us into conflict with others, exploitation and violence would be morally correct; since they are deplorable, conflict had better not be part of our nature. The reasoning, of course, is fallacious: nothing says that nature has to be nice, and what people want to do is not necessarily what they ought to do. Others worry that if conflicting motives are inevitable, it would be futile to try to reduce violence and exploitation; our current social arrangements would be the best one can hope for. But that does not follow either. Among modern Western societies, homicide rates vary from 0.5 per million persons per year in Iceland in the first half of the twentieth century, to 10 in most European countries at present, to 25 in Canada, to 100 in the United States and Brazil. There is plenty of room for practical measures that could reduce the murder rate before we are faced with the academic question of whether it can ever be reduced to zero. Moreover, there are ways to reduce conflict other than to dream of a golden future of indiscriminate love. People in all societies not only perpetrate violence but

deplore it. And people everywhere take steps to reduce violent conflict, such as sanctions, redress, censure, mediation, ostracism, and law.

I hope this discussion strikes you as trite, so I can get on with the content of the chapter. My goal is not to convince you that people don't always want the best for one another, but to try to explain when and why that should be true. But sometimes the trite has to be stated. The observation that conflict is part of the human condition, banal though it is, contradicts fashionable beliefs. One is expressed in the gluey metaphor of social relations as attachment, bonding, and cohesion. Another is the assumption that we unthinkingly play out the roles society assigns to us, and that social reform is a matter of rewriting the roles. I suspect that if you pressed many academics and social critics you would find views no less Utopian than those of Charles Reich.

If the mind is an organ of computation engineered by natural selection, our social motives should be strategies that are tailored to the tournaments we play in. People should have distinct kinds of thoughts and feelings about kin and non-kin, and about parents, children, siblings, dates, spouses, acquaintances, friends, rivals, allies, and enemies. Let's explore them in turn.

KITH AND KIN

Smile on your brother, sang the Youngbloods; a brotherhood of man, sang John Lennon. When we talk of beneficence, we use the language of kinship. Our father who art in heaven; the fatherhood of God; church fathers; Father Christmas; father figure; patriotism. The mother country; the mother church; Mother Superior; motherhood and apple pie; maternal. Blood brothers; black brothers; brothers-in-arms; brotherly love; temple brotherhoods; brethren; fraternities; Brother, can you spare a dime? Sisterhood is powerful; sister cities; soul sisters; sisters of mercy; sororities. The family of man; crime families; one big happy family.

The kinship metaphors have a simple message: treat certain people as kindly as you treat your blood relatives. We all understand the presupposition. The love of kin comes naturally; the love of non-kin does not. That is the fundamental fact of the social world, steering everything from how we grow up to the rise and fall of empires and religions. The explanation is straightforward. Relatives share genes to a greater extent than

nonrelatives, so if a gene makes an organism benefit a relative (say, by feeding or protecting it), it has a good chance of benefiting a copy of itself. With that advantage, genes for helping relatives will increase in a population over the generations. The vast majority of altruistic acts in the animal kingdom benefit the actor's kin. The most extreme examples of kin-directed altruism are found among social insects like ants and bees, in which the workers give their all to the colony. They are permanently sterile and defend the colony with kamikaze tactics like blowing up to spray noxious chemicals on an invader or stinging it with a barbed stinger that pulls the insect's body apart when dislodged. Such dedication comes largely from an unusual genetic system which makes them more closely related to their sisters than they would be to their offspring. By defending the colony they help their mothers make sisters instead of making offspring of their own.

Genes can't call to one another or pull the strings of behavior directly. In humans, "kin altruism" and "benefiting one's genes" are shorthand for two collections of psychological machinery, one cognitive, one emotional.

Humans are equipped with a desire and an ability to learn their family tree. Genealogy is a special kind of knowledge. First, the relationships are digital. You're either someone's mother or you aren't. You might be eighty percent sure that Bill is John's father, but that is not the same as thinking that Bill is eighty percent of a father to John. We speak of half-brothers, but everyone knows the expression is shorthand for having the same mother and different fathers or vice versa. Second, kinship is a relation. No one is a father or a sister, period; they have to be the father or the sister *of* someone. Third, kinship is topological. Everyone is a node in a web whose links are defined by parenthood, generation, and gender. Kinship terms are logical expressions that are read off the geometry and labeling of the web: a "parallel cousin," for example, is one's father's brother's child or one's mother's sister's child. Fourth, kinship is self-contained. Age, place of birth, acquaintanceship, status, occupation, zodiac sign, and all the other categories in which we place people lie in a different plane from the categories of kinship and need not be consulted when we calculate kinship.

Homo sapiens is obsessed with kinship. All over the world, when people are asked to talk about themselves, they begin with their parentage and family ties, and in many societies, especially foraging groups, people rattle off endless genealogies. For adoptees, childhood refugees, or

descendants of slaves, curiosity about biological kin can drive a lifelong quest. (Entrepreneurs hope to exploit this motive when they send out those computer-generated postcards that offer to trace Steven Pinker's ancestors and find the **Pinker** family seal and coat of arms.) Of course, people ordinarily do not test each other's DNA; they assess kinship by indirect means. Many animals do it by smell. Humans do it with several kinds of information: who grows up together, who resembles whom, how people interact, what reliable sources say, and what can be logically deduced from other kin relationships.

Once we know how we are related to other people, the other component of the psychology of kinship kicks in. We feel a measure of solidarity, sympathy, tolerance, and trust toward our relatives, added on to whatever other feelings we may have for them. ("Home," according to the poem by Robert Frost, is "something you somehow haven't to deserve.") The added good will one feels toward kin is doled out according to a feeling that reflects the probability that the kind act will help a relative propagate copies of one's genes. That in turn depends on the nearness of the relative to oneself in the family tree, the confidence one has in that nearness, and the impact of the kindness on the relative's prospects of reproducing (which depends on age and need). So parents love their children above all others, cousins love each other but not as much as siblings do, and so on. Of course, no one crunches genetic and actuarial data and then decides how much to love. Rather, the mental programs for familial love were calibrated in the course of evolution so that love *correlated with* the probability in the ancestral environment that a loving act would benefit copies of genes for loving acts.

You might think this is just the banal observation that blood is thicker than water. But in today's intellectual climate, the observation is a shocking, radical thesis. A Martian who wanted to learn about human interactions from a textbook in social psychology would have no inkling that humans behave any differently to their relatives than to strangers. Some anthropologists have argued that our sense of kinship has nothing to do with biological relatedness. The conventional wisdom of Marxists, academic feminists, and café intellectuals embraces some astonishing claims: that the nuclear family of husband, wife, and children is a historical aberration unknown in centuries past and in the non-Western world; that in primitive tribes marriage is uncommon and people are indiscriminately promiscuous and free of jealousy; that throughout history the bride and groom had no say in their marriage; that romantic love was

invented by the troubadours of medieval Provence and consisted of the adulterous love of a knight for a married lady; that children used to be thought of as miniature adults; that in olden times children died so often that mothers were unaffected by the loss; that concern for one's children is a recent invention. These beliefs are false. Blood really *is* thicker than water, and no aspect of human existence is untouched by that part of our psychology.

Families are important in all societies, and their core is a mother and her biological children. All societies have marriage. A man and a woman enter a publicly acknowledged alliance whose primary goal is children; the man has a "right" of exclusive sexual access to the woman; and they both are obligated to invest in their children. The details vary, often according to the patterns of blood relationships in the society. Generally, when men can be confident that they are the fathers of their wives' children, nuclear families form, usually near the husband's extended kin. In the smaller number of societies where men are not so confident (for example, when they are away for long stretches of military service or farm labor), families live near the mother's kin, and children's principal male benefactors are their closest blood relatives, their maternal uncles. Even then, biological fatherhood is recognized and valued. Both sides of the extended family take an interest in the marriage and the children, and the children feel solidarity with both sides, even when the official rules of descent recognize only one side (as in our own surnames, which are reckoned according to the father's family).

Women fare better when they stay near their relatives and the men move around, because they are surrounded by fathers, brothers, and uncles who can come to their aid in disputes with their husbands. The dynamic was vividly enacted in *The Godfather* when the son of Marlon Brando's character, Sonny Corleone, nearly murdered his sister's husband when he found out that the husband had battered her. Life imitated art two decades later when the real-life son of Brando, Christian Brando, did murder his sister's boyfriend when he found out that the boyfriend had battered her. When a woman has to leave home to live near her husband's family, he can brutalize her with impunity. In many societies, marriages between cousins are encouraged, and the marriages

are relatively harmonious because the usual bickering between husband and wife is mitigated by their sympathy for each other as blood relatives.

These days it's impolite to talk about parental love having anything to do with biological relatedness because it sounds like a slur on the many parents with adopted children and stepchildren. Of course couples love their adopted children; if they weren't unusually committed to simulating a natural family experience they would not have adopted to start with. But stepfamilies are different. The stepparent has shopped for a spouse, not a child; the child is a cost that comes as part of the deal. Stepparents have a poor reputation; even Webster's unabridged dictionary defines *stepmother*, in one of its two definitions, as "one that fails to give proper care or attention." The psychologists Martin Daly and Margo Wilson comment:

The negative characterization of stepparents is by no means peculiar to our culture. The folklorist who consults Stith Thompson's massive *Motif-Index of Folk Literature* will encounter such pithy synopses as "Evil stepmother orders stepdaughter to be killed" (Irish myth), and "Evil stepmother works stepdaughter to death in absence of merchant husband" (India). For convenience, Thompson divided stepfather tales into two categories: "cruel stepfathers" and "lustful stepfathers." From Eskimos to Indonesians, through dozens of tales, the stepparent is a villain in every piece.

Daly and Wilson note that many social scientists assume that the difficulties plaguing step relationships are *caused* by "the myth of the cruel stepparent." But why, they ask, should stepparents in so many cultures be targets of the same slander? Their own explanation is more direct.

The ubiquity of Cinderella stories . . . is surely a reflection of certain basic, recurring tensions in human society. Women must often have been forsaken with dependent children throughout human history, and both fathers and mothers were often prematurely widowed. If the survivor wished to forge a new marital career, then the fate of the children became problematic. [Among the Tikopia and the Yanomamø, the husband] demands the death of his new wife's prior children. Other solutions have included leaving the children with postmenopausal matrilineal relatives, and the *levirate*, a widespread custom by which a widow and her children are inherited by the dead man's brother or other near relative. In the absence of such arrangements, children were obliged to tag along as stepchildren under the care of nonrelatives with no particular

benevolent interest in their welfare. They surely had genuine cause for alarm.

In one study of emotionally healthy middle-class families in the United States, only half of the stepfathers and a quarter of the stepmothers claimed to have "parental feeling" toward their stepchildren, and fewer still claimed to "love" them. The enormous pop-psychology literature on reconstituted families is dominated by one theme: coping with antagonisms. Many professionals now advise warring families to give up the ideal of duplicating a biological family. Daly and Wilson found that stepparenthood is the strongest risk factor for child abuse ever identified. In the case of the worst abuse, homicide, a stepparent is forty to a hundred times more likely than a biological parent to kill a young child, even when confounding factors—poverty, the mother's age, the traits of people who tend to remarry—are taken into account.

Stepparents are surely no more cruel than anyone else. Parenthood is unique among human relationships in its one-sidedness. Parents give; children take. For obvious evolutionary reasons, people are wired to want to make these sacrifices for their own children but not for anyone else. Worse, as we shall see, children are wired to demand these sacrifices of the adults charged with their care, and that can make them downright annoying to people other than their parents and close kin. The writer Nancy Mitford said, "I love children, especially when they cry, for then someone takes them away." But if you are married to the children's parent, no one ever takes them away. The indifference, even antagonism, of stepparents to stepchildren is simply the standard reaction of a human to another human. It is the endless patience and generosity of a biological parent that is special. This point should not diminish our appreciation of the many benevolent stepparents; if anything, it should enhance it, for they are especially kind and self-sacrificing people.

It is often said that you are more likely to be killed by a relative in the home than by a mugger in the street. That sounds suspicious to anyone who knows about evolutionary theory, and it turns out to be false.

Homicide statistics are an important kind of evidence for theories of human relationships. As Daly and Wilson explain, "Killing one's antago-

nist is the ultimate conflict resolution technique, and our ancestors discovered it long before they were people." Homicides cannot be written off as the product of a diseased mind or a sick society. In most cases a killing is unplanned and undesired; it is the disastrous climax of an escalating battle in which brinkmanship has been carried too far. For every killing there must be countless arguments that cool down and countless threats that are not carried out. That makes homicide an excellent assay for conflict and its causes. Unlike lesser conflicts which can only be discovered through reports that the participants can fudge, a homicide leaves a missing person or a dead body, which are hard to ignore, and homicides are meticulously investigated and documented.

People sometimes do murder their relatives. There are infanticides, filicides, parricides, matricides, fratricides, siblicides, uxoricides, familicides, and several unnamed kinds of kin-killing. In a typical data set from an American city, a quarter of the homicides are committed by strangers, a half by acquaintances, and a quarter by "relatives." But most of the relatives are not blood kin. They are spouses, in-laws, and step relations. Only two to six percent of homicide victims are done in by their blood relatives. In fact, that is surely an overestimate. People see their blood relatives more often than they see other people, so relatives are more often within striking distance. When one focuses on people who live together, so that the opportunities for interacting are held constant, one finds that the risk of being killed by a nonrelative is at least eleven times greater than the risk of being killed by a blood relative, and probably much higher than that.

The de-escalation of conflicts among blood relatives is part of a larger pattern of kin solidarity called nepotism. In everyday usage the word refers to bestowing favors on relatives (literally, "nephews") as a perquisite of a job or social rank. Institutional nepotism is officially illicit in our society, though it is widely practiced, and in most societies people are surprised to hear that we consider it a vice. In many countries a newly appointed official openly fires all the civil servants under him and replaces them with relatives. Relatives are natural allies, and before the invention of agriculture and cities, societies were organized around clans of them. One of the fundamental questions of anthropology is how foraging people divide themselves into bands or villages, typically with about fifty members though varying with the time and place. Napoleon Chagnon amassed meticulous genealogies that link thousands of members of the Yanomamö, the foraging and horticultural people of the

Amazon rainforest whom he has studied for thirty years. He showed how kinship is the cement that keeps villages together. Close kin fight each other less often and come to each other's aid in fights more often. A village fissions when its population grows, the villagers become less related to one another, and they increasingly get on each other's nerves. A fight erupts, loyalties divide along blood lines, and one party storms off with his closer kin to form a new village.

A spouse is the most familiar example of *fictive* kin: genetically unrelated people who are called kin and claim the emotions ordinarily directed at kin. The biologist Richard Alexander has pointed out that if spouses are faithful, if each acts on behalf of the union's children rather than other blood relatives, and if the marriage lasts the lifetime of both, the genetic interests of a couple are identical. Their genes are tied up in the same package, their children, and what is good for one spouse is good for the other. Under these idealized conditions, marital love should be stronger than any other kind.

In reality, people's blood kin do claim some of their loyalties, and no one can ever be certain that a spouse is one hundred percent faithful, much less that the spouse will never desert or die. In a simpleminded species, the strength of spousal love might be set at some optimum medium level reflecting the overall probability of nepotism, infidelity, desertion, and widowhood. But humans are sensitive to the particulars of their marriages and fine-tune their emotions accordingly. It is no surprise to a biologist that in-laws, infidelity, and stepchildren are the major causes of marital strife.

Because a couple's genes are in the same boat, and each spouse shares genes with his or her kin, the kin have an interest—in both senses of the word—in their marriage. If your son marries my daughter, our genetic fortunes are partly linked in our common grandchildren, and to that extent what is good for you is good for me. Marriages make in-laws into natural allies, and that is one reason why in all cultures marriages are alliances between clans, not just between spouses. The other reason is that when parents have power over their adult children, as they had in all cultures until recently, the children are excellent trade goods. Since my children don't want to marry each other, you have something I need:

a spouse for my child. Thus dowries and bride-prices are ubiquitous in human cultures, though goods like status and allegiance in conflicts with third parties are also factored into the deal. Like all business transactions, the successful sale or trade of an offspring proves the good faith of the parties and makes them more likely to trust each other in the future. So in-laws are both genetic partners and business partners.

For future-minded parents, in-laws should be chosen carefully. Not only should parents assess the assets and trustworthiness of prospective in-laws, but they should size up whether the dollop of good will that comes free with a common genetic interest in the grandchildren would be put to the best use. It might be wasted on an already secure ally or an implacable foe, but could make all the difference for a clan whose sympathies are somewhere in between. Strategic matchmaking is one outcome of the psychology of kinship; another is rules about who can marry whom. In many cultures people are encouraged to marry their cross cousins and forbidden to marry their parallel cousins. A cross cousin is the child of your mother's brother or of your father's sister; a parallel cousin is a child of your mother's sister or of your father's brother. Why the distinction? Consider the most common arrangement, in which daughters are traded among clans of related males, and imagine yourself contemplating marriage with various cousins (it doesn't matter whether you are male or female). If you marry your cross cousin, you are consummating an exchange with a proven trading partner: a clan with which your own family (presided over by your paternal grandfather) has traded a bride in the past (your mother or your aunt). If you marry your parallel cousin, either you are marrying within the clan (if your father and the father of your betrothed are brothers) and bringing in no external goods, or you are marrying someone from a clan of strangers (if your mother and the mother of your betrothed are sisters).

These intrigues have spawned two of the modern myths of kinship: that in traditional societies, people have no voice in whom they marry, and that kinship has nothing to do with genetic relatedness. The grain of truth in the first myth is that parents everywhere wield as much power as they can to influence whom their children marry. Children do not, however, passively accept their parents' choice. People everywhere have powerful emotions about whom they want to marry—that is, romantic love—and engagements are often fierce battles of wills between parents and children. Even when parents have the final say, the children lobby day and night to make their feelings known, and the feelings almost always enter

into the decision. The plot of Sholem Aleichem's *Tevye's Daughters* (adapted into the musical *Fiddler on the Roof*) unfolds on this battlefield, and similar plots are found across the world. When children elope, it is a catastrophe for their parents. The business deal or strategic opportunity of a lifetime may have just been frittered away. Worse, if the parents had pledged the child years before—which often happens, because children are born at different times and the second half of an exchange must wait until a child reaches marriageable age—the parents are now in default and at the mercy of the loan sharks. Or the parents may have mortgaged themselves to the eyeballs to buy a spouse for the departed child. Defaults on marriage agreements are a leading cause of feuding and warfare in traditional societies. With the stakes so high, it is no wonder that the parents' generation always teaches that romantic love is frivolous or does not exist at all. The intellectuals who conclude that romantic love is a recent invention of medieval troubadours or of Hollywood scriptwriters have taken this establishment propaganda at face value.

Those who take fictive kin as evidence that kinship has nothing to do with biology have also bought an official doctrine. A big problem with marriage rules, like the one mandating marriage between cross cousins, is that the age and sex mixture of a group fluctuates, so sometimes there will be no eligible partners for a child. As with all rules, the challenge is to work around them without making them a farce. An obvious solution is to redefine who is related to whom. An eligible bachelor might be called a cross cousin even if the genealogical diagram says otherwise, saving a daughter from spinsterhood without setting the precedent that other children can marry whom they please. But deep down no one is fooled by these face-saving measures. A similar hypocrisy applies to other fictive kin. With kin emotions being so powerful, manipulators try to tap them for solidarity among non-kin by *calling* the non-kin kin. The tactic has been rediscovered again and again, from tribal chiefs to modern preachers and sappy rock musicians. But even in tribes where fictive kin labels are publicly treated with the utmost seriousness, if you press someone in private he will acknowledge that so-and-so is not *really* his brother or cousin. And when people show their true colors in a dispute, the colors go with blood relatives, not fictive ones. Many modern parents tell their children to address family friends as Uncle and Aunt. When I was a child, my friends and I used to refer to them as our fake uncles and fake aunts. Children are even more adamant in resisting the ubiquitous pressure to call their new stepparents Mom and Dad.

For millennia, kin emotions have shaped even the largest societies. The reach of parental love can extend over generations via gifts and inheritance. Parental love causes the fundamental paradox of politics: no society can be simultaneously fair, free, and equal. If it is fair, people who work harder can accumulate more. If it is free, people will give their wealth to their children. But then it cannot be equal, for some people will inherit wealth they did not earn. Ever since Plato called attention to these tradeoffs in *The Republic*, most political ideologies can be defined by the stance they take on which of these ideals should yield.

Another surprising consequence of kin solidarity is that the family is a subversive organization. That conclusion flies in the face of the right-wing view that the church and state have always been steadfast upholders of the family and of the left-wing view that the family is a bourgeois, patriarchal institution designed to suppress women, weaken class solidarity, and manufacture docile consumers. The journalist Ferdinand Mount has documented how every political and religious movement in history has sought to undermine the family. The reasons are obvious. Not only is the family a rival coalition competing for a person's loyalties, but it is a rival with an unfair advantage: relatives innately care for one another more than comrades do. They bestow nepotistic benefits, forgive the daily frictions that strain other organizations, and stop at nothing to avenge wrongs against a member. Leninism, Nazism, and other totalitarian ideologies always demand a new loyalty "higher" than, and contrary to, family ties. So have religions from early Christianity to the Moonies ("We're your family now!"). In Matthew 10:34-37, Jesus says:

Think not that I am come to send peace on earth: I came not to send peace, but a sword. For I am come to set a man at variance against his father, and the daughter against her mother, and the daughter in law against her mother in law. And a man's foes shall be they of his own household. He that loveth father or mother more than me is not worthy of me: and he that loveth son or daughter more than me is not worthy of me.

When Jesus said "Suffer the little children to come unto me," he was saying that they should not go unto their parents.

Successful religions and states eventually realize they have to coexist with families, but they do what they can to contain them, particularly the

most threatening ones. The anthropologist Nancy Thornhill has found that the incest laws of most cultures are not created to deal with the problem of brother-sister marriages; brothers and sisters don't want to marry to begin with. Although brother-sister incest may be included in the prohibition and may help to legitimize it, the real targets of the laws are marriages that threaten the interests of the lawmakers. The rules ban marriages among more distant relatives like cousins, and are promulgated by the rulers of stratified societies to prevent wealth and power from accumulating in families, which could be future rivals. The anthropologist Laura Betzig has shown that the medieval church's rules on sex and marriage were also weapons against familial dynasties. In feudal Europe, parents did not bequeath their estates in equal parts to all of their children. Plots of land could not be subdivided every generation or they would become uselessly small, and a title can fall on only one heir. The custom of primogeniture arose, in which everything went to the oldest son and the other sons hit the road to seek their fortunes, often joining armies or the church. The church filled up with disinherited younger sons, who then manipulated marriage rules to make it harder for owners and title-holders to bear legitimate heirs. If they died without sons, the properties and titles passed back to the disinherited brothers or the church they served. According to their laws, a man could not divorce a childless wife, remarry while she was alive, adopt an heir, bear an heir with a woman closer than a seventh cousin, or have sex on various special days that added up to more than half the year. The story of Henry VIII reminds us that much of European history revolves around battles between powerful individuals trying to leverage family feelings for political gain—marrying strategically, striving for heirs—and other powerful individuals trying to foil them.

PARENTS AND CHILDREN

For an organism designed by natural selection, leaving descendants is the reason for being and the goal of all toil and struggle. The love of a parent for a child should be vast, and so it is. But it should not be boundless. Robert Trivers discovered a subtle but profound implication of genetics for the psychology of the family.

In most sexual species, parents bequeath fifty percent of their genes to each offspring. One strategy for maximizing the number of genes in

the next generation is to pump out as many babies as possible as quickly as possible. That is what most organisms do. Baby organisms, however, are more vulnerable than adults because they are smaller and less experienced, and in most species the majority never make it to adulthood. All organisms therefore face a "choice" of allocating their time, calories, and risk to caring for an existing offspring and upping its odds of survival, or cranking out new offspring and letting them all fend for themselves. Depending on details of the species' ecosystem and body plan, either strategy can be genetically profitable. Birds and mammals have opted to care for their offspring, mammals by the extreme step of evolving organs that siphon nutrients from their own bodies and package them for their offspring as milk. Birds and mammals invest calories, time, risk, and bodily wear and tear on their offspring, and are repaid in increases in the offspring's life expectancy.

In theory, a parent could go to the other extreme and care for its first-born all its life—say, by suckling it until the parent died of old age. But that would make little sense because at some point the calories being turned into milk could better be invested in bearing and suckling a new offspring. As the first-born grows, each additional pint of milk is less and less crucial to its survival, and it becomes better and better equipped to find its own food. A younger offspring becomes a better investment, and the parent should wean the older one.

A parent should transfer investment from an older child to a younger one when the benefit to the younger exceeds the cost to the older. The reckoning is based on the fact that the two children are equally related to the parent. But these calculations are from the parent's point of view; the first child sees it differently. He shares fifty percent of his genes with his younger sibling, but he shares *one hundred percent* of his genes with *himself*. As far as he is concerned, the parent should continue to invest in him until the benefit to a younger sibling is greater than *twice* the cost to him. The genetic interests of the parent and the child diverge. Each child should want more parental care than the parent is willing to give, because parents want to invest in all of their offspring equally (relative to their needs), whereas each child wants more of the investment for himself. The tension is called parent-offspring conflict. In essence it is sibling rivalry: siblings compete among themselves for their parents' investment, whereas the parents would be happiest if each accepted a share proportional to his or her needs. But sibling rivalry can be played out with parents, too. In evolutionary terms, the only reason a parent

withholds investment from one offspring is to save it for future ones. An offspring's conflict with its parents is really a rivalry with unborn Siblings.

A tangible example is weaning conflict. The calories a mother converts to milk are not available to grow a new offspring, so nursing suppresses ovulation. At some point mammalian mothers wean their young so their bodies can prepare for bearing a subsequent offspring; When they do, the young mammal puts up a holy stink, hounding the mother for access to the teat for weeks or months before acquiescing.

When I mentioned the theory of parent-offspring conflict to console a colleague whose two-year-old son had become a pest after the birth of a younger brother, he snapped, "All you're saying is that people are selfish!" Sleepless for weeks, he could be forgiven for missing the point. Clearly, parents aren't selfish; parents are the least selfish entities in the known universe. But they aren't infinitely selfless either, or every whine and tantrum would be music to their ears. And the theory predicts that children aren't completely selfish, either. If they were, they would murder each newborn sibling to free up all the parents' investment for themselves and would demand to be breast-fed all their lives. The reason they don't is that they are *partly* related to their present and future siblings. A gene that made a child murder his newborn sister would have a fifty percent chance of destroying a copy of itself, and in most species that cost outweighs the benefit of having one's mother's milk all to oneself. (In some species, like spotted hyenas and some birds of prey, the costs don't outweigh the benefits, and siblings do murder one another.) A gene that made a fifteen-year-old want to nurse would foreclose an opportunity for his mother to manufacture new copies of that gene inside viable siblings. Either cost would exceed twice the benefit, so most organisms have their siblings' interests at heart, though discounted relative to their own. The point of the theory is not that children want to take or that parents don't want to give; it's that children want to take *more* than what their parents want to give.

Parent-offspring conflict begins in the womb. A woman with an unborn child seems like a vision of harmony and nurturance, but beneath the glow a mighty battle goes on inside her. The fetus tries to mine the mother's body for nutrients at the expense of her ability to bear¹ future

children. The mother is a conservationist, trying to keep her body in reserve for posterity. The human placenta is a tissue of the fetus that invades the mother's body and taps into her bloodstream. Through it the fetus secretes a hormone that ties up maternal insulin, increasing the levels of blood sugar which it can then skim off. But the resulting diabetes compromises the mother's health, and over evolutionary time she has fought back by secreting more insulin, which prompted the fetus to secrete more of the hormone that ties up insulin, and so on, until the hormones reached a thousand times their usual concentration. The biologist David Haig, who first noticed prenatal parent-offspring conflict, remarks that the raised hormone levels are like raised voices: a sign of conflict. In a similar tug-of-war, the fetus increases the mother's blood pressure, forcing more nutrients its way at the expense of her health.

The battle continues once the baby is born. The first decision of motherhood is whether to let the newborn die. Infanticide has been practiced in all the world's cultures. In ours, "killing babies" is a synonym for depravity, one of the most shocking crimes imaginable. One might think it is a form of Darwinian suicide and proof that other cultures' values are incommensurable with ours. Daly and Wilson show that it is neither.

Parents of all species face the choice of whether to continue to invest in a newborn. Parental investment is a precious resource, and if a newborn is likely to die there is no point in throwing good money after bad by fledging or suckling it. The time and calories would be better spent on its littermates or clutchmates, in starting over with new offspring, or in waiting until the circumstances are better. Thus most animals let their runtish or sickly offspring die. Similar calculations enter into human infanticide. In foraging peoples, women have their first child in their late teens, nurse them on demand for four infertile years, and see many die before adulthood. If a woman is lucky, she might raise two or three children to maturity. (The large broods of our grandparents are historical aberrations resulting from agriculture, which provided substitutes for mother's milk.) To raise even a small number of children to adulthood, a woman has to make hard choices. Women in the world's cultures let infants die in circumstances in which the odds of survival are low: when the infant is deformed, a twin, fatherless, or fathered by a man who isn't the woman's husband, and when the mother is young (and so has opportunities to try again), lacks social support, had the infant soon after another child, is overburdened with older offspring, or is otherwise in

desperate straits, such as from a famine. Infanticide in the modern West is similar. The statistics show that the mothers who let their infants die are young, poor, and unwed. There are many explanations, but the parallel with the rest of the world is unlikely to be a coincidence.

Infanticidal mothers are not heartless, and even when infant mortality is common, people never treat young life casually. Mothers experience infanticide as an unavoidable tragedy. They grieve for the child and remember it with pain all their lives. In many cultures people try to distance their emotions from a newborn until they are assured it will survive. They may not touch, name, or grant legal personhood to a baby until a danger period is over, much like our own customs of the christening and the bris (the circumcision of eight-day-old Jewish boys).

The emotions of new mothers, which would drive the decision to keep a baby or let it die, may have been shaped by these actuarial facts. Postpartum depression has been written off as a hormonal delirium, but as with all explanations of complex emotions, one must *ask why* the brain is wired so as to let hormones have their effects. In most of human evolutionary history, a new mother had good reason to pause and take stock. She faced a decision between a definite tragedy now and a chance of an even greater tragedy years hence, and the choice was not to be taken lightly. Even today, the typical rumination of a depressed new mother—how will I cope with this burden?—is a genuine issue. The depression is most severe in the circumstances that lead mothers elsewhere in the world to commit infanticide, such as poverty, marital conflict, and single motherhood.

The emotional response called "bonding" is also surely more sophisticated than the stereotype in which a woman is smitten with a lifelong attachment to her baby if she interacts with it in a critical window after birth, like the victims of Puck in *A Midsummer Night's Dream* who became infatuated with the first person they saw upon awakening. Mothers appear to proceed from a cool assessment of the infant and their current prospects, to an appreciation of the infant as a uniquely wonderful individual after about a week, to a gradual deepening of love over the next few years.

The infant is an interested party, and fights for its interests with the only weapon at its disposal: cuteness. Newborns are precociously responsive to their mothers; they smile, make eye contact, perk up to her speech, even mimic her facial expressions. These advertisements of a functioning nervous system could melt a mother's heart and tip the balance in a close

decision of whether to keep the baby. The ethologist Konrad Lorenz pointed out that the geometry of babies—a large head, a bulbous cranium, large eyes low in the face, pudgy cheeks, and short limbs—elicits tenderness and affection. The geometry comes from the baby-assembly process. The head end grows fastest in the womb, and the other end catches up after birth; babies grow into their brain and their eyes. Lorenz showed that animals with that geometry, such as ducks and rabbits, strike people as cute. In his essay "A Biological Homage to Mickey Mouse," Stephen Jay Gould showed that cartoonists exploit the geometry to make their characters more appealing. It's conceivable that the genes exploit it too, exaggerating the juvenile features of a newborn, particularly those that signal good health, to make it look cuter to its mother.

Once a child is allowed to live, the battle between the generations continues. How could an offspring hold its own in the battle? As Trivers notes, babies cannot fling their mothers to the ground and nurse at will; they have to use psychological tactics. A baby has to manipulate its parents' genuine concern for its welfare to induce them to give more than they would otherwise be willing to give. Since parents can learn to ignore cries of "wolf," the tactics have to be more insidious. An infant knows its own condition better than a parent does, because the infant's brain is connected to sensors throughout its body. Both the parent and the infant have an interest in the parent's responding to the infant's needs, such as by feeding it when it is hungry and cuddling it when it is cold. That gives the infant an opening to elicit more care than the parent wants to give. The baby can cry when it is not so cold or hungry, or withhold a smile until it gets its way. The baby need not literally be faking. Since parents should evolve to recognize sham crying, the baby's most effective tactic might be to feel genuinely miserable, even when there is no biological need. Self-deception may begin early.

The child can also resort to extortion by howling at night or throwing a tantrum in public, situations in which the parents are averse to letting the noise continue and are apt to capitulate. Worse, the parents' interest in their children's welfare allows the children to hold themselves hostage, say, by thrashing about in a violent tantrum or refusing to do something both parties know the child would enjoy. Thomas Schelling notes that children are in an excellent position to use paradoxical tactics (Chapter 6). They can cover their ears, scream, avoid their parents' gaze, or regress, all of which prevent them from registering or understanding their parents' threats. We get the evolution of the brat.

The theory of parent-offspring conflict is an alternative to two popular ideas. One is Freud's Oedipal complex, the hypothesis that boys have an unconscious wish to have sex with their mothers and kill their fathers, and therefore fear that their fathers will castrate them. (Similarly, in the Electra complex, little girls want to have sex with their fathers.) There is indeed a fact to be explained. In all cultures, young children are sometimes possessive of their mothers and cool to the mother's consort. Parent-offspring conflict offers a straightforward explanation. Daddy's interest in Mommy takes her attention away from me—and, even worse, threatens to create a baby brother or sister. Children may well have evolved tactics for delaying that sad day by diminishing their mothers' interest in sex and keeping their fathers away from her. It would be a straightforward extension of weaning conflict. The theory explains why so-called Oedipal feelings are as common in girls as in boys, and avoids the preposterous idea that little boys want to copulate with their mothers.

Daly and Wilson, who proposed the alternative, believe that Freud's mistake was to run together two different kinds of parent-offspring conflict. Young children are in conflict with their father over access to their mother, but it is not a sexual rivalry. And older children may have a sexual conflict with their parents, especially their fathers, but it is not a rivalry over the mother. In many societies fathers compete with their sons for sexual partners, explicitly or implicitly. In polygynous societies, where a man can have several wives, they might literally compete for the same women. And in most societies, polygynous or monogamous, a father must subsidize his son's quest for a wife at the expense of his other children or his own aspirations. The son may be impatient for the father to begin diverting resources to him; a still-robust father is a roadblock to his career. Filicides and parricides in most of the world are touched off by such competition.

Parents also arrange marriages, which is a polite way of saying that they sell or trade their children. Here again interests can conflict. Parents may hammer out a package deal in which one child gets a catch and another gets a loser. In polygynous societies a father may trade his daughters for wives for himself. Whether a daughter is traded for a daughter-in-law or for a wife, her value can hinge on her virginity: men don't want to marry a woman who might be carrying another man's child.

(Effective birth control is recent and still far from universal.) Therefore fathers take an interest in their daughters' sexuality, a mimic of the Electra complex but without either party desiring the other. In many societies men take horrifying measures to guarantee a daughter's "purity." They may lock her up, cloak her from head to toe, and extirpate her interest in sex by the horrible custom known by the euphemism "female circumcision" (it is a circumcision in the same sense that Lorena Bobbitt performed a bris). When the measures fail, they may execute an unchaste daughter to preserve what they call, ironically, the family's "honor." (In 1977 a Saudi princess was publicly stoned to death for bringing dishonor to her grandfather, the brother of the king, by having an indiscreet affair in London.) Parent-daughter conflict is a special case of conflict over the "ownership" of women's sexuality, a topic to which we will return.

The other popular theory subverted by parent-offspring conflict is the biology-culture distinction, in which babies are a bundle of uncivilized instincts and parents socialize them into competent, well-adjusted members of society. Personality, in this conventional wisdom, is shaped in the formative years by the parenting process. Parents and children both want the children to prosper in the social milieu, and since children are in no position to shape themselves, socialization represents a confluence of their interests.

Trivers reasoned that, according to the theory of parent-offspring conflict, parents should *not* necessarily have their children's interests at heart when they try to socialize them. Just as parents often act against a child's interests, they may try to *train* the child to act against its *own* interests. Parents want each child to act more altruistically to its siblings than the child wants to. That is because it pays the parents for a child to be altruistic when the benefit to a sibling exceeds the cost to the child, but it pays the *child* to be altruistic only when the benefit exceeds *twice* the cost. For more distant kin such as half-siblings and cousins, the difference between the parents' interests and the child's interests is even greater, because the parent is more closely related to the half-sibling or cousin than the child is. Similarly, parents may try to persuade children that staying home to help at the nest, allowing themselves to be sold in

marriage, and other outcomes that are good for the parent (and hence the child's unborn siblings) are in fact good for the child. As in all arenas of conflict, parents may resort to deception and, since children are no fools, self-deception. So even if children acquiesce to a parent's rewards, punishments, examples, and exhortations for the time being because they are smaller and have no choice, they should not, according to the theory, allow their personalities to be shaped by these tactics.

Trivers went out on a limb with that prediction. The idea that parents shape their children is so ingrained that most people don't even realize it is a testable hypothesis and not a self-evident truth. The hypothesis has now been tested, and the outcome is one of the most surprising in the history of psychology.

Personalities differ in at least five major ways: whether a person is sociable or retiring (extroversion-introversion), whether a person worries constantly or is calm and self-satisfied (neuroticism-stability), whether the person is courteous and trusting or rude and suspicious (agreeableness-antagonism), whether a person is careful or careless (conscientiousness-undirectedness), and whether a person is daring or conforming (openness-nonopenness). Where do these traits come from? If they are genetic, identical twins should share them, even if they were separated at birth, and biological siblings should share them more than adoptive siblings do. If they are a product of socialization by parents, adoptive siblings should share them, and twins and biological siblings should share them more when they grow up in the same home than when they grow up in different homes. Dozens of studies have tested these kinds of predictions on thousands of people in many countries. The studies have looked not only at these personality traits but at actual outcomes in life such as divorce and alcoholism. The results are clear and replicable, and they contain two shockers.

One result has become well known. Much of the variation in personality—about fifty percent—has genetic causes. Identical twins separated at birth are alike; biological siblings raised together are more alike than adopted siblings. That means that the other fifty percent must come from the parents and the home, right? Wrong! Being brought up in one home versus another accounts, at most, for *percent* of the differences among people in personality. Identical twins separated at birth are not only similar; they are virtually as similar as identical twins raised together. Adoptive siblings in the same home are not just different; they are about as different as two children plucked from the population

at random. The biggest influence that parents have on their children is at the moment of conception.

(I hasten to add that parents are unimportant only when it comes to *differences* among them and differences among their grown children. Anything that *all* normal parents do that affects all children is not measured in these studies. Young children surely need the love, protection, and tutelage of a sane parent. As the psychologist Judith Harris has put it, the studies imply only that children would turn into the same kinds of adults if you left them in their homes and social milieus but switched all the parents around.)

No one knows where the other forty-five percent of the variation comes from. Perhaps personality is shaped by unique events impinging on the growing brain: how the fetus lay in the womb, how much maternal blood it diverted, how it was squeezed during birth, whether it was dropped on its head or caught certain viruses in the early years. Perhaps personality is shaped by unique experiences, like being chased by a dog or receiving an act of kindness from a teacher. Perhaps the traits of parents and the traits of children interact in complicated ways, so that two children growing up with the same parents really have different environments. One kind of parent may reward a rambunctious child and punish a placid one; another kind of parent may do the opposite. There is no good evidence for these scenarios, and I think two others are more plausible, both of which see personality as an adaptation rooted in the divergence of interests between parents and offspring. One is the child's battle plan for competing with its siblings, which I will discuss in the following section. The other is the child's battle plan for competing in its peer group.

Judith Harris has amassed evidence that children everywhere are socialized by their peer group, not by their parents. At all ages children join various play groups, circles, gangs, packs, cliques, and salons, and they jockey for status within them. Each is a culture that absorbs some customs from the outside and generates many of its own. Children's cultural heritage—the rules of Ringolevio, the melody and lyrics of the nyah-nyah song, the belief that if you kill someone you legally have to pay for his gravestone—is passed from child to child, sometimes for thousands of years. As children grow up they graduate from group to group and eventually join adult groups. Prestige at one level gives one a leg up at the next; most significantly, the leaders of young adolescent cliques are the first to date. At all ages children are driven to figure out what it takes to succeed among their peers and to give these strategies precedence over

anything their parents foist on them. Weary parents know they are no match for a child's peers, and rightly obsess over the best neighborhood in which to bring their children up. Many successful people immigrated to this country as children and were not handicapped in the least by culturally inept parents who never learned the language or customs. As a researcher of language development I have always been struck by the way in which children rapidly pick up the language (especially the accent) of their peers, though they spend more time with their parents.

Why aren't children putty in parents' hands? Like Trivers and Harris, I suspect it is because children's genetic interests overlap only partly with their parents'. Children take their calories and protection from their parents, because their parents are the only ones willing to provide them, but they get their information from the best sources they can find and forge their strategies for dealing with life themselves. Their own parents may not be the wisest and most knowledgeable adults around, and worse, the rules at home are often stacked against the children in favor of their born and unborn siblings. And as far as reproduction is concerned, the home is a dead end. The child will have to compete for mates, and before that for the status necessary to find and keep them, in other arenas, which play by different rules. The child had better master them.

The conflict of interest between parents and offspring is unacknowledged in our public discourse about children. In most times and places, the advantage has been to the parents, and they have wielded their power as cruel tyrants. This century has seen the tables turn. Child-welfare experts flood the bookstores with parenting manuals and the government with policy advice. All politicians paint themselves as friends of children and their opponents as enemies. Childrearing manuals used to advise mothers on how to make it through the day. With Dr. Spock, the spotlight fell on the child and the mother became a nonperson, there only to create mental health in the child and to take the blame if the child turned out bad.

The child-welfare revolution was one of the great liberation movements of all time, but like all realignments of power, it can go too far. Feminist social critics have argued that mothers' interests have been erased by the child-care gurus. In discussing her book *The Myths of Motherhood*, Shari Thurer notes:

The most pervasive myth is the denial of maternal ambivalence: that mothers really both love and hate their children. There's a real silence about the ambivalent feelings; . . . it's tantamount to being a bad mother. [In my clinical practice], anger and rage are normal. Children are endlessly demanding, and they'll just suck you dry. Women shouldn't have to feel that they are supposed to meet all of the child's needs. But the myth is that mother love is natural and operative at all times.

Even the advocates of mothers' rights often feel they must frame their arguments in terms of the interests of the child (an overburdened mother is a bad mother) rather than in terms of the interests of the mother (an overburdened mother is unhappy).

More conservative social critics have also begun to notice that parents' and children's interests can diverge. Barbara Dafoe Whitehead has reviewed data showing that sex education does not succeed in its advertised function of reducing teenage pregnancies. Today's teens know all about sex and its hazards, but the girls end up pregnant anyway, quite possibly because they don't mind the idea of having babies. If the teens' parents do mind, they may have to enforce their interests by controlling the teenagers (with chaperones and curfews), not just by educating them.

I mention these debates not to take a side but to call attention to the long reach of parent-offspring conflict. Evolutionary thinking is often put down as a "reductionistic approach" that aims to redefine all social and political issues as technical problems of biology. The criticism has it backwards. The evolution-free discourse that has prevailed for decades has treated childrearing as a technological problem of determining which practices grow the best children. Trivers' insight is that decisions about childrearing are inherently about how to allocate a scarce resource—the parents' time and effort—to which several parties have a legitimate claim. As such, childrearing will always be partly a question of ethics and politics, not just of psychology and biology.

BROTHERS AND SISTERS-

Ever since Cain slew Abel, siblings have been entangled by many emotions. As people of the same generation who know each other well, they react to each other as individuals: they may like or dislike one another, compete if they are of the same sex, or feel sexual attraction if they are

not. As close kin, they feel a big extra dose of affection and solidarity. But though they share fifty percent of their genes with each other, each sibling shares one hundred percent of its genes with itself, so brotherly or sisterly love has its limits. Being offspring of the same parents, siblings are rivals for their parents' investment, from weaning to the reading of the will. And though genetic overlap makes a pair of siblings natural allies, it also makes them unnatural parents, and that genetic alchemy tempers their sexual feelings.

If people gave birth to a single litter of interchangeable w-tuplets, parent-offspring conflict would be a raw struggle among the siblings, each demanding more than its share. But all children are different, if for no other reason than that they are born at different times. Parents may not want to invest one nth of their energy in each of their n children, but may, like shrewd portfolio managers, try to pick winners and losers and invest accordingly. The investment decisions are not conscious forecasts of the number of grandchildren expected from each child, but emotional responses that were tuned by natural selection to have outcomes that maximized that number in the environment in which we evolved. Though enlightened parents try mightily never to play favorites, they don't always succeed. In one study, fully two-thirds of British and American mothers confessed to loving one of their children more.

How do parents make Sophie's Choice and sacrifice a child when circumstances demand it? Evolutionary theory predicts that the main criterion should be age. Childhood is a minefield, and the older a child gets, the luckier a parent is to have it alive and the more irreplaceable the child is as an expected source of grandchildren, right up until sexual maturity. (From then on, the reproductive years begin to be used up and the child's expected number of offspring declines.) For example, the actuarial tables show that a four-year-old in a foraging society will, on average, give a parent 1.4 times as many grandchildren as a newborn, an eight-year-old 1.5 times as many, and a twelve-year-old 1.7 times as many. So if parents already have a child when an infant arrives and cannot feed them both, they should sacrifice the infant. In no human society do parents sacrifice an older child when a younger one is born. In our society, the chance that a parent will kill a child drops steadily with the child's age, especially during the vulnerable first year. When parents are asked to imagine the loss of a child, they say they would grieve more for older children, up until the teenage years. The rise and fall of anticipated

grief correlates almost perfectly with the life expectancies of hunter-gatherer children.

On the other hand, a younger child, being more helpless, has more use for a parent's daily ministrations. Parents report more tender feelings for their younger offspring, even though they seem to value the older ones more. The calculations begin to change when parents get older and a new child is likely to be their last one. There is nothing to save for, and the baby of the family is likely to be indulged. Parents also favor children that one might call, in a cold-hearted way, better investments: more vigorous, better looking, more talented.

Given that parents are apt to play favorites, offspring should be selected to manipulate their parents' investment decisions in their favor. Children are exquisitely sensitive to favoritism, right through adulthood and after the parents' deaths. They should calculate how to make the best of the hand that nature dealt them and of the dynamics of the poker game they were born into. The historian Frank Sulloway has argued that the elusive nongenetic component of personality is a set of strategies to compete with siblings for parental investment, and that is why children in the same family are so different. Each child develops in a different family ecology and forms a different plan for getting out of childhood alive. (The idea is an alternative to Harris' proposal that personality is a strategy for coping in peer groups, though both could be right.)

A first-born child has been spotted several advantages. The first-born, merely by having survived to its present age, is more precious to the parents, and of course is bigger, stronger, and wiser and will be so for as long as the younger one is a child. Having ruled the roost for a year or more, the first-born sees the newcomer as a usurper. Thus he (or she) should identify with his parents, who have aligned their interests with his, and should resist changes to the status quo, which has always served him well. He should also learn how best to wield the power that fate has granted him. In sum, a first-born should be a conservative and a bully. Second-born children have to cope in a world that contains this obsequious martinet. Since they cannot get their way with thuggery and toadyism, they must cultivate the opposite strategies. They should become appeasers and cooperators. And with less at stake in the status quo, they should be receptive to change. (These dynamics depend, too, on the innate components of the personalities of the siblings and on their sex, size, and spacing; your mileage may vary.)

Later-borns have to be flexible for another reason. Parents invest in

the children who show the most promise of success in the world. The first-born has staked a claim in whatever personal and technical skills she is best at. There's no point in a later-born competing on that turf; any success would have to come at the expense of the older and more experienced sibling, and he (or she) would be forcing his parents to pick a winner, with daunting odds against him. Instead, he should find a different niche in which to excel. That gives his parents an opportunity to diversify their investments, because he complements his older sibling's skills in competition outside the family. Siblings in a family exaggerate their differences for the same reason that species in an ecosystem evolve into different forms: each niche supports a single occupant.

Family therapists have discussed these dynamics for decades, but is there any hard evidence? Sulloway analyzed data on 120,000 people from 196 adequately controlled studies of birth order and personality. As he predicted, first-borns are less open (more conforming, traditional, and closely identified with parents), more conscientious (more responsible, achievement-oriented, serious, and organized), more antagonistic (less agreeable, approachable, popular, and easygoing), and more neurotic (less well-adjusted, more anxious). They are also more extroverted (more assertive, more leaderly), though the evidence is cloudy because they are more serious, which makes them seem more introverted.

Family politics affects not only what people say in paper-and-pencil tests but how they act in the world when playing for high stakes. Sulloway analyzed biographical data from 3,894 scientists who had voiced opinions on radical scientific revolutions (such as the Copernican revolution and Darwinism), 893 members of the French National Convention during the Terror of 1793-1794, more than seven hundred protagonists in the Protestant Reformation, and the leaders of sixty-two American reform movements such as the abolition of slavery. In each of these shake-ups, later-borns were more likely to support the revolution, first-borns were more likely to be reactionary. The effects are not by-products of family size, family attitudes, social class, or other confounding factors. When evolutionary theory was first proposed and still incendiary, later-borns were *ten times* as likely to support it as first-borns. Other alleged causes of radicalism, such as nationality and social class, have only minor effects. (Darwin himself, for example, was upper-class but later-born.) Later-born scientists are also less specialized, trying their hands in a greater number of scientific fields.

If personality is an adaptation, why should people carry the strategies

that served them in the rumpus room right into adulthood? One possibility is that siblings never completely escape the orbit of their parents, but compete all their lives. That is certainly true in traditional societies, including foraging groups. Another is that tactics like assertiveness and conservatism are skills like any other. As a young person invests more and more in honing them, she becomes increasingly loath to retrace the learning curve to cultivate new strategies for dealing with people.

The discovery that children brought up in the same family are no more similar than they would be if they had been brought up on different planets shows how poorly we understand the development of personality. All we know is that cherished ideas about the influence of parents are wrong. The most promising hypotheses, I suspect, will come from recognizing that childhood is a jungle and that the first problem children face in life is how to hold their own among siblings and peers.

The relationship between a brother and a sister has an added twist: one is male, one is female, and those are the ingredients of a sexual relationship. People have sex with and marry those with whom they interact the most—their co-workers, the girl or boy next door—and the people most like themselves—those of the same class, religion, race, and appearance. The forces of sexual attraction should pull siblings together like magnets. Even if familiarity breeds some contempt and only a tiny fraction of siblings hit it off, there should be millions of brothers and sisters wanting to have sex and get married. There are virtually none. Not in our society, not in any well-studied human society, not in most animals in the wild. (Prepubertal children sometimes engage in sexual play; I'm talking about real intercourse between mature siblings.)

Do brothers and sisters avoid copulating because their parents discourage it? Almost certainly not. Parents try to socialize their children to be more affectionate with each other ("Go ahead—kiss your sister!"), not less. And if they did discourage sex, it would be the only case in all of human experience in which a sexual prohibition worked. Teenage brothers and sisters do not sneak off for trysts in parks and the back seats of cars.

The incest taboo—a public prohibition against sex or marriage between close relatives—has been an obsession of anthropology for a

century, but it does not explain what keeps siblings apart. Avoiding incest is universal; taboos against incest are not. And most incest taboos are not about sex within the nuclear family. Some are about sex with fictive kin and merely enforce sexual jealousy. For example, polygynous men may pass laws to keep their sons away from their junior wives, officially the sons' "stepmothers." As we have seen, most taboos prohibit marriage (not sex) between more distant kin, such as cousins, and are ploys that rulers use to prevent wealth from accumulating in rival families. Sometimes sex among family members falls under the umbrella of more general codes against incest, but nowhere is it the target.

Brothers and sisters simply don't find each other appealing as sexual partners. That is an understatement: the thought makes them acutely uncomfortable or fills them with disgust. (People who grew up without siblings of the opposite sex do not understand the emotion.) Freud claimed that the strong emotion is itself proof of an unconscious desire, especially when a male claims revulsion at the thought of coitus with his mother. By that reasoning we may conclude that people have an unconscious desire to eat dog feces and to stick needles in their eyes.

Repugnance at sex with a sibling is so robust in humans and other long-lived, mobile vertebrates that it is a good candidate for an adaptation. The function would be to avoid the costs of inbreeding: a reduction in the fitness of offspring. There is a grain of biological truth behind the folklore that incest "thickens the blood" and the stereotypes of defective hillbillies and royal twits. Harmful mutations steadily drip into the gene pool. Some are dominant, cripple their bearers, and are soon selected out. But most are recessive and do no harm until they build up in the population and meet up with copies of themselves when two carriers mate. Since close relatives share genes, if they mate they run a much higher risk that two copies of a harmful recessive gene will match up in their offspring. Since all of us carry the equivalent of one to two lethal recessive genes, when a brother and sister mate they are quite likely to have a compromised offspring, both in theory and in the studies that have measured the risks. The same is true for mother-son and father-daughter matings (and, to a lesser extent, to matings between more distant kin). It stands to reason that humans (and many other animals) have evolved an emotion that makes the thought of sex with a family member a turnoff.

Incest avoidance showcases the complicated software engineering behind our emotions for other people. We feel stronger bonds of affec-

tion to family members than to acquaintances or strangers. We clearly perceive the sexual attractiveness of family members, and even take pleasure in looking at them. But the affection and appreciation of beauty don't translate into a desire to copulate, though if the same emotions had been elicited by a nonrelative, the urge might be irresistible. The way a single bit of knowledge can turn lust into horror has been used to great dramatic effect in the dozens of plots that Polti classifies as "Involuntary crimes of love," of which Sophocles' *Oedipus Rex* is the most famous.

Incest avoidance has two twists. One is that different couplings within the family have different genetic costs and benefits, both for the participants and for the bystanders. We might expect sexual repugnance to be adjusted accordingly. For both males and females, the benefit of having a child with an immediate family member is that the child contains seventy-five percent of each parent's genes, instead of the usual fifty percent (the extra twenty-five percent comes from the genes shared by the parents by virtue of their being related which are then passed on to the child). The costs are the risk of a deformed child and the forgone opportunity to have a child with someone else. The forgone opportunities, however, differ for males and females. Also, children are always sure who their mothers are but are not always sure who their fathers are. For both these reasons, incest has to be costed out separately for each of the possible couplings in a family.

Neither a mother nor a son has any advantage in the mother coupling with the son as opposed to with the boy's father that could offset the genetic risks. And since men are generally not attracted to women old enough to be their mothers, the net result is that mother-son incest virtually never happens.

For incest between fathers and daughters and between brothers and sisters, the calculations come out differently depending on whose point of view we take. A hypothetical ancestral girl made pregnant by a brother or father would be precluded from having a child with a nonrelative for the nine months of pregnancy, and were she to keep the baby, for another two to four years of nursing. She wastes a precious opportunity for reproduction on a child that may be deformed. Incest should be thoroughly repugnant. But a male who impregnates his sister or daughter could be adding to the number of offspring he sires, because her pregnancy does not foreclose his impregnating someone else. There is a risk that the child will be deformed, but if it isn't, the child is a sheer bonus (more accurately, the extra dose of his genes in that child are the bonus).

Incest repugnance might be weaker, making him more likely to cross the line. It is a special case of the lower costs of reproduction for males and their less discriminating sexual desire, to which we will return.

A father, moreover, can never be certain that a daughter is his, so the genetic cost to him could be zero. That could weaken the suppression of desire even further compared to her brother, who is certain to be related to his sister because they share a mother. For stepfathers and stepbrothers, there is no genetic cost at all. It is no surprise, then, that between half and three-quarters of all reported incest cases are between stepfathers and stepdaughters, most of them initiated by the stepfather. Most of the rest are between fathers and daughters, and virtually all are coerced by the father. Some are between girls and other older male relatives, also mostly coerced. A mother gets no genetic benefit from a mating between her husband and her daughter (compared with a mating between her daughter and a son-in-law), but suffers the cost of defective grandchildren, so her interests are aligned with her daughter's and she should be a force opposing incest. Incestuous exploitation of girls might be even more common if their mothers were not around. These battles are driven by strong emotions, but the emotions are not an alternative to the genetic analysis; the analysis explains why they exist. And of course, in science as in detective work, to try to figure out the motive for a crime is not to excuse the crime.

People cannot directly sense their genetic overlap with another person; as with the rest of perception, the brain must combine information from the senses with assumptions about the world to make an intelligent guess. Chapter 4 showed that when the world violates the assumptions, we fall prey to an illusion, and that is exactly what happens in the perception of kinship. The nineteenth-century anthropologist Edward Westermarck conjectured that growing up in intimate closeness with a person in the early years is the key information the brain uses to put the person in the category "sibling." Similarly, when an adult raises a child the adult should perceive the child as "son" or "daughter" and the child should perceive the adult as "mother" or "father." The classifications then negate sexual desire.

These algorithms presuppose a world in which children who are raised together are biological siblings and vice versa. That is certainly true of foraging peoples. A mother's children grow up with her and usually with their father, too. When the assumption is false, people should be the victim of a kinship illusion. If they grow up with a person who is

not a relative, they should be sexually indifferent or repelled. If they do not grow up with a person who is a relative, they should fail to be repelled. Being told in so many words that a date is really your brother or sister may be enough to kill the romantic mood, but an unconscious imprinting mechanism at work during a critical period in early childhood is surely even more powerful.

Both kinds of illusions have been documented. The Israeli communal villages called *kibbutzim* were founded early in the twentieth century by Utopian planners determined to break down the nuclear family. Boys and girls of the same age shared living quarters from shortly after birth through adolescence and were raised together by nurses and teachers. When they became sexually mature, the children who had grown up together very rarely married or even had sex, though marriages were not discouraged. In some parts of China, brides used to move into their in-laws' homes, giving rise to frictions that you can well imagine. Parents hit on the brilliant idea of adopting a bride for their son when she was still a child, guaranteeing that she would forever be under her mother-in-law's thumb. What they did not realize was that the arrangement mimicked the psychological cues to siblinghood. When the couple grew up, they found each other unsexy, and compared with conventional couples, their marriages were unhappy, unfaithful, unfecund, and short. In parts of Lebanon, paternal parallel cousins grow up together as if they were siblings. Parents pressure the cousins into marrying, but the couples are sexually apathetic, relatively childless, and prone to divorce. Unconventional childrearing arrangements have been found to have the same outcome on all continents, and various alternative explanations can be ruled out.

Conversely, people who *do* commit incest often have not grown up together. A study of sibling incest offenders in Chicago found that the only ones who had contemplated marriage were those who had been raised apart. Fathers who sexually abuse their daughters tend to have spent less time with them when they were small. Stepfathers who have had as much contact with their young stepdaughters as biological fathers do are no more likely to abuse them. There are anecdotes that adoptees who seek out their biological parents and siblings often find themselves sexually attracted to them, though I know of no controlled studies.

The Westermarck effect explains the most famous incest offender of all: Oedipus. Laius, king of Thebes, was warned by an oracle that his son would slay him. When Jocasta, his wife, bore a son, he tied the baby up

and left him exposed on a mountain. Oedipus was found and raised by a shepherd and then adopted by the king of Corinth and brought up as his son. On a visit to Delphi, Oedipus learned that he was fated to kill his father and marry his mother, so he left Corinth vowing never to return. On his way toward Thebes, he encountered Laius and killed him in a quarrel. When he then outwitted the Sphinx, his reward was the throne of Thebes and the hand of its widowed queen, Jocasta—the biological mother he did not grow up with. They had four children before he got the bad news.

But the ultimate triumph of the Westermarck theory has been pointed out by John Tooby. The idea that boys want to sleep with their mothers strikes most men as the silliest thing they have ever heard. Obviously it did not seem so to Freud, who wrote that as a boy he once had an erotic reaction to watching his mother dressing. But Freud had a wet-nurse, and may not have experienced the early intimacy that would have tipped off his perceptual system that Mrs. Freud was his mother. The Westermarck theory has out-Freuded Freud.

MEN AND WOMEN

Men and women. Women and men. It will never work.

—EKICAJONG

Sometimes, of course, it does work. A man and a woman can fall in love, and the key ingredient is an expression of commitment, as we saw in Chapter 6. A man and a woman need each other's DNA and hence can enjoy sex. A man and a woman have a common interest in their children, and their enduring love has evolved to protect that interest. And a husband and wife can be each other's best friends, and can enjoy the lifelong dependability and trust that underlies the logic of friendship (more on this later). These emotions are rooted in the fact that if a man and woman are monogamous, together for life, and not nepotistic toward their own families, their genetic interests are identical.

Unfortunately, that is a big "if." Even the happiest couples can fight like cats and dogs, and today fifty percent of marriages in the United States end in divorce. George Bernard Shaw wrote, "When we want to read of the deeds that are done for love, whither do we turn? To the murder column." Conflict between men and women, sometimes deadly,

is universal, and it suggests that sex is not a bonding force in human affairs but a divisive one. Once again, that banality must be stated because the conventional wisdom denies it. One of the Utopian ideals of the 1960s, reiterated ever since by sex gurus like Dr. Ruth, is the intensely erotic, mutually enjoyable, guilt-free, emotionally open, life-long monogamous pair-bond. The alternative from the counterculture was the intensely erotic, mutually enjoyable, guilt-free, emotionally open, round-robin orgy. Both were attributed to our hominid ancestors, to earlier stages of civilization, or to primitive tribes still out there somewhere. Both are as mythical as the Garden of Eden.

The battle between the sexes is not just a skirmish in the war between unrelated individuals but is fought in a different theater, for reasons first explained by Donald Symons. "With respect to human sexuality," he wrote, "there is a female human nature and a male human nature, and these natures are extraordinarily different. . . . Men and women differ in their sexual natures because throughout the immensely long hunting and gathering phase of human evolutionary history the sexual desires and dispositions that were adaptive for either sex were for the other tickets to reproductive oblivion."

Many people deny that there are any interesting differences between the sexes. At my own institution, students taking Psychology of Gender used to be taught that the only well-established difference between men and women is that men like women and women like men. Symons' two human natures are dismissed as "gender stereotypes," as if that were proof that they are false. The belief that spiders spin webs and pigs don't is also a stereotype, but is no less true for that. As we shall see, some stereotypes about sexual feelings have been verified beyond a reasonable doubt. In fact, researchers in sex differences have found that many gender stereotypes *underestimate* the documented differences between the sexes.

Why is there sex to begin with? Lord Chesterfield noted of sex that "the pleasure is momentary, the position ridiculous, and the expense damnable." Biologically speaking, the costs are damnable indeed, so why do almost all complex organisms reproduce sexually? Why don't women give virgin birth to daughters who are clones of themselves instead of

wasting half their pregnancies on sons who lack the machinery to make grandchildren and are nothing but sperm donors? Why do people and other organisms swap out half their genes for the genes of another member of the species, generating variety in their offspring for variety's sake? It's not to evolve faster, because organisms are selected for fitness in the present. It's not to adapt to environmental change, because a random change in an already adapted organism is more likely to be for the worse than for the better, there being vastly more ways to be badly adapted than to be well adapted. The best theory, proposed by John Tooby, William Hamilton, and others, and now supported by several kinds of evidence, is that sex is a defense against parasites and pathogens (disease-causing microorganisms).

From a germ's point of view, you are a big yummy mound of cheesecake, there for the eating. Your body takes a different view, and has evolved a battery of defenses, from your skin to your immune system, to keep them out or do them in. An evolutionary arms race goes on between hosts and pathogens, though a better analogy might be an escalating contest between lockpickers and locksmiths. Germs are small, and they evolve diabolical tricks for infiltrating and hijacking the machinery of the cells, for skimming off its raw materials, and for passing themselves off as the body's own tissues to escape the surveillance of the immune system. The body responds with better security systems, but the germs have a built-in advantage: there are more of them and they can breed millions of times faster, which makes them evolve faster. They can evolve substantially within the lifetime of a host. Whatever molecular locks the body has evolved, the pathogens can evolve keys to open them.

Now, if an organism is asexual, once the pathogens crack the safe of its body they also have cracked the safes of its children and siblings. Sexual reproduction is a way of changing the locks once a generation. By swapping half the genes out for a different half, an organism gives its offspring a head start in the race against the local germs. Its molecular locks have a different combination of pins, so the germs have to start evolving new keys from scratch. A malevolent pathogen is the one thing in the world that rewards change for change's sake.

Sex poses a second puzzle. Why do we come in *two* sexes? Why do we make one big egg and lots of little sperm, instead of two equal blobs that coalesce like mercury? It is because the cell that is to become the baby cannot be just a bag of genes; it needs the metabolic machinery of the

rest of a cell. Some of that machinery, the mitochondria, has its own genes, the famous mitochondrial DNA which is so useful in dating evolutionary splits. Like all genes, the ones in mitochondria are selected to replicate ruthlessly. And that is why a cell formed by fusing two equal cells faces trouble. The mitochondria of one parent and the mitochondria of the other parent wage a ferocious war for survival inside it. Mitochondria from each parent will murder their counterparts from the other, leaving the fused cell dangerously underpowered. The genes for the rest of the cell (the ones in the nucleus) suffer from the crippling of the cell, so they evolve a way of heading off the internecine warfare. In each pair of parents, one "agrees" to unilateral disarmament. It contributes a cell that provides no metabolic machinery, just naked DNA for the new nucleus. The species reproduces by fusing a big cell that contains a half-set of genes plus all the necessary machinery with a small cell that contains a half-set of genes and nothing else. The big cell is called an egg and the small cell is called a sperm.

Once an organism has taken that first step, the specialization of its sex cells can only escalate. A sperm is small and cheap, so the organism might as well make many of them, and give them outboard motors to get to the egg quickly and an organ to launch them on their way. The egg is big and precious, so the organism had better give it a head start by packing it with food and a protective cover. That makes it more expensive still, so to protect the investment the organism evolves organs that let the fertilized egg grow inside the body and absorb even more food, and that release the new offspring only when it is large enough to survive. These structures are called male and female reproductive organs. A few animals, hermaphrodites, put both kinds of organs in every individual, but most specialize further and divide up into two kinds, each allocating all their reproductive tissue to one kind of organ or the other. They are called males and females.

Trivers has worked out how all the prominent differences between males and females stem from the difference in the minimum size of their investment in offspring. Investment, remember, is anything a parent does that increases the chance of survival of an offspring while decreasing the parent's ability to produce other viable offspring. The investment can be energy, nutrients, time, or risk. The female, by definition, begins with a bigger investment—the larger sex cell—and in most species commits herself to even more. The male contributes a puny package of genes and usually leaves it at that. Since every offspring requires one of each,

the female's contribution is the limiting step on how many offspring can be produced: at most, one offspring for each egg she creates and nurtures. Two cascades of consequences flow from this difference.

First, a single male can fertilize several females, which forces other males to go mateless. That sets up a competition among males for access to females. A male may beat up other males to prevent them from getting to a female, or compete for the resources necessary to mate, or court a female to get her to choose him. Males therefore vary in reproductive success. A winner can beget many offspring, a loser will beget none.

Second, the reproductive success of males depends on how many females they mate with, but the reproductive success of females does not depend on how many males they mate with. That makes females more discriminating. Males woo females and mate with any female that lets them. Females scrutinize males and mate only with the best ones: the ones with the best genes, the ones most willing and able to feed and protect her offspring, or the ones that the other females tend to prefer.

Male competition and female choice are ubiquitous in the animal kingdom. Darwin called attention to these two spectacles, which he dubbed sexual selection, but was puzzled as to why it should be males that compete and females that choose rather than the other way around. The theory of parental investment solves the puzzle. The greater-investing sex chooses, the lesser-investing sex competes. Relative investment, then, is the cause of sex differences. Everything else—testosterone, estrogen, penises, vaginas, Y chromosomes, X chromosomes—is secondary. Males compete and females choose only because the slightly bigger investment in an egg that *defines* being female tends to get multiplied by the rest of the animal's reproductive habits. In a few species, the whole animal reverses the initial difference in investment between egg and sperm, and in those cases *females* should compete and *males* should choose. Sure enough, these exceptions prove the rule. In some fishes, the male broods the young in a pouch. In some birds, the male sits on the egg and feeds the young. In those species, the females are aggressive and try to court the males, who select partners carefully.

In a typical mammal, though, the female does almost all the investing. Mammals have opted for a body plan in which the female carries the fetus inside her, nourishes it with her blood, and nurses and protects it after it is born until the offspring has grown big enough to fend for itself. The male contributes a few seconds of copulation and a sperm cell

weighing one ten-trillionth of a gram. Not surprisingly, male mammals compete for opportunities to have sex with female mammals. The details depend on the rest of the animal's way of life. Females live alone or in groups, in small groups or large ones, in stable groups or temporary ones, using sensible criteria like where the food is, where it's safest, where they can easily bear and raise young, and whether they need strength in numbers. Males go where the females are. Female elephant seals, for example, congregate on beach strips which a male can easily patrol. A single male can monopolize the group, and males fight bloody battles for this jackpot. Bigger fighters are better fighters, so the males have evolved to be four times the size of the females.

Apes have a wide variety of sexual arrangements. That means, by the way, that there is no such thing as an "ape legacy" that humans are doomed to live by. Gorillas live on the fringes of forests in small groups of one male and several females, and the males fight each other for control over females, the males evolving to be twice the females' size. Gibbon females are solitary and widely dispersed, and the male finds a female's territory and acts as a faithful consort. Since other males are off in other territories, they fight no more than females do and are no bigger. Orangutan females are solitary but close enough together that a male can monopolize two or more of their ranges, and the males are about 1.7 times the size of the females. Chimps live in large, unstable groups that no male could dominate. Groups of males live with the females, and the males compete for dominance, which confers more opportunities to copulate. The males are about 1.3 times as large as the females. With lots of males around, a female has an incentive to mate with many of them so that a male can never be sure that an infant is not his and hence will not murder the infant to make its mother available to bear his own offspring. Bonobo (pygmy chimp) females are almost indiscriminately promiscuous, and the males fight less and are about the same size as females. They compete in a different way: inside the females' bodies.

Sperm can survive in the vagina for several days, so a promiscuous female can have several males' sperm competing inside her for a chance at fertilizing the egg. The more sperm a male produces, the greater the chance that one of his will get there first. That explains why chimpanzees have enormous testicles for their body size. Bigger testes make more sperm, which have a better chance inside promiscuous females. A gorilla is four times the weight of a chimpanzee, but his testicles are four times smaller. The females in his harem have no chance to copulate with

any other male, so his sperm do not have to compete. Gibbons, who are monogamous, have small testicles, too.

In almost all primates (indeed, in almost all mammals), the males are deadbeat dads, contributing nothing to their offspring but DNA. Other species are more fatherly. Most birds, many fishes and insects, and social carnivores such as wolves have males that protect or feed their offspring. The evolution of male parental investment is helped along by several things. One is external fertilization, found in most fishes, where the female drops her eggs and the male fertilizes them in the water. The male is guaranteed that the fertilized eggs carry his genes, and since they have been released while the young are undeveloped, he has an opportunity to help. But in most mammals the cards are stacked against doting fatherhood. The egg is tucked away inside the mother, where some other male can fertilize it, so a male is never certain an offspring is his. He faces the danger of wasting his investment on another male's genes. Also, the embryo does most of its growing inside the mother, where the father can't get at it to help directly. And a father can easily desert and try to mate with another female, whereas the female is left holding the bag and cannot get rid of the fetus or offspring without having to go through the long process of nurturing an embryo all over again to get back to where she started. Fatherhood is also promoted when a species' lifestyle makes the benefits exceed the costs: when the offspring would be vulnerable without him, when he can easily provision them with concentrated food like meat, and when the young are easy to defend.

When males become devoted fathers, the rules of the mating game change. A female may choose a mate based on his ability and his willingness to invest in their offspring, insofar as she can judge. Females, not just males, compete for mates, though the prizes are different: males compete for fertile females willing to copulate, females compete for flush males willing to invest. Polygamy is no longer a matter of one male beating up all the others, or the females all wanting to be inseminated by the fiercest or prettiest male. When males invest *more* than females, as we have seen, the species may be polyandrous, with tough females keeping harems of males. (The mammals' body plan has foreclosed that option.) When one male has much more to invest than others (because, say, he controls a better territory), females may be better off sharing him—polygyny—than each having her own mate, because a fraction of a big resource may be better than the entirety of a small one. When males'

contributions are more equal, the undivided attention of one becomes valuable, and the species settles on monogamy.

Many birds appear to be monogamous. In *Manhattan*, Woody Allen says to Diane Keaton, "I think people should mate for life, like pigeons or Catholics." The movie came out before ornithologists began to submit birds to DNA testing, which revealed, to their shock, that pigeons are not so faithful either. In some species of birds, a third of the offspring contain the DNA of a male other than the female's consort. The male bird is adulterous because he tries to raise the offspring of one female and mate with others, hoping that her offspring will survive on their own, or best of all, be raised by a cuckolded consort. The female bird is adulterous because she has a chance of getting the best of both worlds: the genes of the fittest male and the investment of the most willing male. The victim of cuckoldry is worse off than if he had failed to breed at all, because he has devoted his worldly efforts to the genes of a competitor. So in species whose males invest, the male's jealousy is directed not only at rival males but at the female. He may guard her, follow her around, copulate repeatedly, and avoid females that show signs of having recently mated.

The human mating system is not like any other animal's. But that does not mean it escapes the laws governing mating systems, which have been documented in hundreds of species. Any gene predisposing a male to be cuckolded, or a female to receive less paternal help than her neighbors, would quickly be tossed from the gene pool. Any gene that allowed a male to impregnate all the females, or a female to bear the most indulged offspring of the best male, would quickly take over. These selection pressures are not small. For human sexuality to be "socially constructed" and independent of biology, as the popular academic view has it, not only must it have miraculously escaped these powerful pressures, but it must have withstood equally powerful pressures of a different kind. If a person played out a socially constructed role, other people could shape the role to prosper at his or her expense. Powerful men could brainwash the others to enjoy being celibate or cuckolded, leaving the women for them. Any willingness to accept socially constructed gender roles would be selected out, and genes for resisting the roles would take over.

What kind of animal is *Homo sapiens*? We are mammals, so a woman's minimum parental investment is much larger than a man's. She contributes nine months of pregnancy and (in a natural environment) two to four years of nursing. He contributes a few minutes of sex and a teaspoon of semen. Men are about 1.15 times as large as women, which tells us that they have competed in our evolutionary history, with some men mating with several women and some men mating with none. Unlike gibbons, who are isolated, monogamous, and relatively sexless, and gorillas, who are clustered, harem-forming, and relatively sexless, we are gregarious, with men and women living together in large groups and constantly facing opportunities to couple. Men have smaller testicles for their body size than chimpanzees but bigger ones than gorillas and gibbons, suggesting that ancestral women were not wantonly promiscuous but were not always monogamous either. Children are born helpless and remain dependent on adults for a large chunk of the human lifespan, presumably because knowledge and skills are so important to the human way of life. So children need parental investment, and men, because they get meat from hunting and other resources, have something to invest. Men far exceed the minimum investment that their anatomy would let them get away with: they feed, protect, and teach their children. That should make cuckoldry a concern to men, and a man's willingness and ability to invest in children a concern to women. Because men and women live together in large groups, like chimps, but the males invest in their offspring, like birds, we developed marriage, in which a man and woman form a reproductive alliance that is meant to limit demands from third parties for sexual access and parental investment.

These facts of life have never changed, but others have. Until recently, men hunted and women gathered. Women were married soon after puberty. There was no contraception, no institutionalized adoption by nonrelatives, and no artificial insemination. Sex meant reproduction and vice versa. There was no food from domesticated plants or animals, so there was no baby formula; all children were breast-fed. There was also no paid day care, and no househusbands; babies and toddlers hung around with their mothers and other women. These conditions persisted through ninety-nine percent of our evolutionary history and have shaped our sexuality. Our sexual thoughts and feelings are adapted to a world in which sex led to babies, whether or not we want to make babies now. And they are adapted to a world in which children were a mother's problem more than a father's. When I use terms like "should," "best," and

"optimal," they will be a shorthand for the strategies that would have led to reproductive success in that world. I will not be referring to what is morally right, attainable in the modern world, or conducive to happiness, which are different matters altogether.

The first question of strategy is how many partners to want. Remember that when the minimum investment in offspring is greater for females, a male can have more offspring if he mates with many females, but a female does not have more offspring if she mates with many males—one per conception is enough. Suppose a foraging man with one wife can expect two to five children with her. A premarital or extramarital liaison that conceives a child would increase his reproductive output by twenty to fifty percent. Of course, if the child starves or is killed because the father isn't around, the father is genetically no better off. The optimal liaison, then, is with a married woman whose husband would bring up the child. In foraging societies, fertile women are almost always married, so sex with a woman is usually sex with a married woman. Even if she is not, more fatherless children live than die, so a liaison with an unmarried partner can increase reproduction, too. None of this math applies to women. A part of the male mind, then, should want a variety of sexual partners for the sheer sake of having a variety of sexual partners.

Do you think that the only difference between men and women is that men like women and women like men? Any bartender or grandmother you ask would say that men are more likely to have a wandering eye, but perhaps that is just an old-fashioned stereotype. The psychologist David Buss has looked for the stereotype in the people most likely to refute it—men and women in elite liberal American universities a generation after the feminist revolution, in the heyday of politically correct sensibilities. The methods are refreshingly direct.

Confidential questionnaires asked a series of questions. How strongly are you seeking a spouse? The answers were on average identical for men and women. How strongly are you seeking a one-night stand? The women said, Not very strongly; the men said, Pretty strongly. How many sexual partners would you like to have in the next month? In the next two years? In your lifetime? Women said that in the next month eight-tenths of a sexual partner would be just about right. They wanted

one in the next two years, and four or five over their lifetimes. Men wanted two sex partners within the month, eight in the next two years, and eighteen over their lifetimes. Would you consider having sex with a desirable partner that you had known for five years? For two years? For a month? For a week? Women said "probably yes" for a man they had known for a year or more, "neutral" for one they had known for six months, and "definitely not" for someone they had known a week or less. Men said "probably yes" as long as they had known the woman for a week. How short a time would a man have to know a woman before he would definitely *not* have sex with her? Buss never found out; his scale did not go down past "one hour." When Buss presented these findings at a university and explained them in terms of parental investment and sexual selection, a young woman raised her hand and said, "Professor Buss, I have a simpler explanation of your data." Yes, he said, what is it? "Men are slime."

Are men really slime, or are they just trying to look like slime? Perhaps in questionnaires men try to exaggerate their studliness but women want to avoid looking easy. The psychologists R. D. Clark and Elaine Hatfield hired attractive men and women to approach strangers of the opposite sex on a college campus and say to them, "I have been noticing you around campus. I find you very attractive," and then ask one of three questions: (a) "Would you go out with me tonight?" (b) "Would you come over to my apartment tonight?" (c) "Would you go to bed with me tonight?" Half the women consented to a date. Half the men consented to a date. Six percent of the women consented to go to the stooge's apartment. Sixty-nine percent of the men consented to go to the stooge's apartment. None of the women consented to sex. Seventy-five percent of the men consented to sex. Of the remaining twenty-five percent, many were apologetic, asking for a rain check or explaining that they couldn't because their fiancée was in town. The results have been replicated in several states. When the studies were conducted, contraception was widely available and safe-sex practices were heavily publicized, so the results cannot be dismissed simply because women might be more cautious about pregnancy and sexually transmitted diseases.

An awakening of male sexual desire by a new partner is known as the Coolidge effect, after a famous anecdote. One day President Calvin Coolidge and his wife were visiting a government farm and were taken on separate tours. When Mrs. Coolidge was shown the chicken pens, she asked whether the rooster copulated more than once a day. "Dozens

of times," replied the guide. "Please tell that to the president," Mrs. Coolidge requested. When the president was shown the pens and told about the rooster, he asked, "Same hen every time?" "Oh, no, Mr. President, a different one each time." The president said, "Tell that to Mrs. Coolidge." Many male mammals are indefatigable when a new willing female is available after each copulation. They cannot be fooled by the experimenter cloaking a previous partner or masking her scent. This shows, incidentally, that male sexual desire is not exactly "undiscriminating." Males may not care *what kind of* female they mate with, but they are hypersensitive to *which* female they mate with. It is another example of the logical distinction between individuals and categories that I argued was so important when criticizing associationism in Chapter 2.

Men do not have the sexual stamina of roosters, but they show a kind of Coolidge effect in their desire over longer periods. In many cultures, including our own, men report that their sexual ardor for their wives wanes in the first years of marriage. It is the concept of the individual person, not her appearance or other qualities, that triggers the decline; the taste for new partners is not just an example of variety being the spice of life, as in getting bored with strawberry and wanting to try chocolate ripple. In Isaac Bashevis Singer's story "Schlemiel the First," a simpleton from the mythical village of Chelm sets out on a trip but loses his way and inadvertently returns home, thinking he has come across another village, which by an amazing coincidence looks just like his. He meets a woman who looks exactly like the wife he has grown tired of, and finds her ravishing.

Another part of the male sexual mind is an ability to be easily aroused by a possible sex partner—indeed, by the faintest hint of a possible sex partner. Zoologists have found that the males of many species will court an enormous range of objects having a vague resemblance to the female: other males, females of the wrong species, females of the right species that have been stuffed and nailed to a board, parts of stuffed females such as a head suspended in midair, even parts of stuffed females with important features missing like the eyes and the mouth. The male of the human species is aroused by the sight of a nude woman, not only in the flesh but in movies, photographs, drawings, postcards, dolls, and

bit-mapped cathode-ray-tube displays. He takes pleasure in this mistaken identity, supporting a worldwide pornography industry which in the United States alone grosses ten billion dollars a year, almost as much as spectator sports and the movies combined. In foraging cultures, young men make charcoal drawings of breasts and vulvas on rock overhangs, carve them on tree trunks, and scratch them in the sand. Pornography is similar the world over and was much the same a century ago as it is today. It depicts in graphic physical detail a succession of anonymous nude females eager for casual, impersonal sex.

It would make no sense for a woman to be easily aroused by the sight of a nude male. A fertile woman never has a shortage of willing sexual partners, and in that buyer's market she can seek the best husband available, the best genes, or other returns on her sexual favors. If she could be aroused by the sight of a naked man, men could induce her to have sex by exposing themselves and her bargaining position would be compromised. The reactions of the sexes to nudity are quite different: men see nude women as a kind of invitation, women see nude men as a kind of threat. In 1992 a Berkeley student known around campus as the Naked Guy chose to jog, attend class, and eat in the dining halls in the nude as a protest against the repressive sexual traditions of Western society. He was expelled when some female students protested that his behavior should be classified as sexual harassment.

Women do not seek the sight of naked male strangers or enactments of anonymous sex, and there is virtually no female market for pornography. (*Playgirl*, the supposed counterexample, is clearly for gay men. It has no ads for any product a woman would buy, and when a woman gets a subscription as a gag gift she finds herself on mailing lists for gay male pornography and sex toys.) In the laboratory, some early experiments claimed that men and women showed identical physiological arousal to a pornographic passage. The men, however, showed a bigger response to the *neutral* passage in the control condition than the women showed to the *pornography*. The so-called neutral passage, which had been chosen by the female investigators, described a man and a woman chatting about the relative merits of an anthropology major over pre-med. The men found it highly erotic! Women can sometimes be aroused when they have agreed to watch portrayals of intercourse, but they do not seek them out. (Symons points out that women are more choosy than men in *consenting* to sex, but once they have consented, there is no reason to believe they are any less responsive to sexual stimulation.) The closest

mass-market equivalents to pornography for women are the romance novel and the bodice-ripper, in which the sex is described in the context of emotions and relationships rather than as a succession of bumping bodies.

The desire for sexual variety is an unusual adaptation, for it is insatiable. Most commodities of fitness show diminishing returns or an optimal level. People do not seek mass quantities of air, food, and water, and they want to be not too hot and not too cold but just right. But the more women a man has sex with, the more offspring he leaves; too much is never enough. That gives men a limitless appetite for casual sex partners (and perhaps for the commodities that in ancestral environments would have led to multiple partners, such as power and wealth). Everyday life offers most men few opportunities to plumb the bottom of the desire, but occasionally a man is rich, famous, handsome, and amoral enough to try. Georges Simenon and Hugh Hefner claimed to have had thousands of partners; Wilt Chamberlain estimated that he had twenty thousand. Say we liberally adjust for braggadocio and assume that Chamberlain inflated his estimate by a factor of, say, ten. That would still mean that one thousand nine hundred and ninety-nine sex partners were not enough.

Symons notes that homosexual relations offer a clear window on the desires of each sex. Every heterosexual relationship is a compromise between the wants of a man and the wants of a woman, so differences between the sexes tend to be minimized. But homosexuals do not have to compromise, and their sex lives showcase human sexuality in purer form (at least insofar as the rest of their sexual brains are not patterned like those of the opposite sex). In a study of homosexuals in San Francisco before the AIDS epidemic, twenty-eight percent of gay men reported having had more than a thousand sex partners, and seventy-five percent reported having had more than a hundred. No gay woman reported a thousand partners, and only two percent reported as many as a hundred. Other desires of gay men, like pornography, prostitutes, and attractive young partners, also mirror or exaggerate the desires of heterosexual men. (Incidentally, the fact that men's sexual wants are the same whether they are directed at women or directed at other men refutes the

theory that they are instruments for oppressing women.) It's not that gay men are oversexed; they are simply men whose male desires bounce off other male desires rather than off female desires. Symons writes, "I am suggesting that heterosexual men would be as likely as homosexual men to have sex most often with strangers, to participate in anonymous orgies in public baths, and to stop off in public restrooms for five minutes of fellatio on the way home from work if women were interested in these activities. But women are not interested."

Among heterosexuals, if men want variety more than women do, Econ 101 tells us what should follow. Copulation should be conceived of as a female service, a favor that women can bestow on or withhold from men. Scores of metaphors treat sex with a woman as a precious commodity, whether they take the woman's perspective (*saving yourself, giving it away, feeling used*) or the man's (*getting any, sexual favors, getting lucky*). And sexual transactions often obey market principles, as cynics of all persuasions have long recognized. The feminist theorist Andrea Dworkin has written, "A man wants what a woman has—sex. He can steal it (rape), persuade her to give it away (seduction), rent it (prostitution), lease it over the long term (marriage in the United States) Or own it outright (marriage in most societies)." In all societies, it is mostly or entirely the men who woo, proposition, seduce, use love magic, give gifts in trade for sex, pay bride-prices (rather than collect dowries), hire prostitutes, and rape.

Sexual economics, of course, also depends on the desirability of the individuals, not just the average desires of the sexes. People "pay" for sex—in cash, commitment, or favors—when the partner is more desirable than they are. Since women are more discriminating than men, the average man has to pay for sex with the average woman. An average man can attract a higher-quality wife than casual sex partner (assuming that a marriage commitment is a kind of payment), whereas a woman can attract a higher-quality casual sex partner (who would pay nothing) than husband. The highest-quality men, in theory, should have a large number of women willing to have sex with them. A cartoon by Dan Wasserman shows a couple leaving the theater after having seen *Indecent Proposal*. The husband says, "Would you sleep with Robert Redford for a million dollars?" She replies, "Yes, but they'd have to give me some time to come up with the money."

The cartoonist's wit, though, exploits our sense of surprise. We don't expect real life to work that way. The men most attractive to women do

not hire themselves out as prostitutes; they may even hire prostitutes themselves. In 1995, the actor Hugh Grant, arguably the world's handsomest man, was arrested for having oral sex with a prostitute in the front seat of his car. The simple economic analysis fails here because money and sex are not completely fungible. As we shall see, part of men's attractiveness comes from their wealth, so the most attractive men don't need the money. And the "payment" that most women hope for is not cash but long-term commitment, which is a scarce resource even for the handsomest and wealthiest man. The economics of the Hugh Grant affair are well summed up by an exchange from another movie, based on the story of Heidi Fleiss, the Hollywood madame. A call girl asks her friend why her handsome tricks have to pay for sex. "They're not paying you for the sex," the friend explains. "They're paying you to go away afterwards."

Could it be that men *learn* to want sexual variety? Perhaps it is a means to an end, the end being status in our society. The Don Juan is revered as a dashing stud; the pretty woman on his arm is a trophy. Certainly anything that is desirable and rare can become a status symbol. But that does not mean that all desirable things are pursued *because* they are status symbols. I suspect that if men were given the hypothetical choice between clandestine sex with many attractive women and a *reputation* for sex with many attractive women, but without the sex, they would go for the sex. Not only because sex is incentive enough, but because a reputation for having sex is a disincentive. Don Juans do *not* inspire admiration, especially in women, though they may inspire envy in men, a different and not always welcome reaction. Symons remarks,

Human males appear to be so constituted that they resist learning *not* to desire variety despite impediments such as Christianity and the doctrine of sin; Judaism and the doctrine of mensch; social science and the doctrines of repressed homosexuality and psychosexual immaturity; evolutionary theories of monogamous pair-bonding; cultural and legal traditions that support and glorify monogamy; the fact that the desire for variety is virtually impossible to satisfy; the time and energy, and the innumerable kinds of risk—physical and emotional—that variety-seeking entails; and the obvious potential rewards of learning to be sexually satisfied with one woman.

A wandering eye, learned or not, is not the only component of a man's mind. Though desire often leads to behavior, it often does not, because,

other desires are stronger or because tactics of self-control (see Chapter 6) have been put into effect. Men's sexual tastes can be calibrated and overruled depending on the man's attractiveness, the availability of partners, and his assessment of the costs of a dalliance.

HUSBANDS AND WIVES

In evolutionary terms, a man who has a short-term liaison is betting that his illegitimate child will survive without his help or is counting on a cuckolded husband to bring it up as his own. For the man who can afford it, a surer way to maximize progeny is to seek several wives and invest in all their children. Men should want many wives, not just many sex partners. And in fact, men in power have allowed polygyny in more than eighty percent of human cultures. Jews practiced it until Christian times and outlawed it only in the tenth century. Mormons encouraged it until it was outlawed by the U.S. government in the late nineteenth century, and even today there are thought to be tens of thousands of clandestine polygynous marriages in Utah and other western states. Whenever polygyny is allowed, men seek additional wives and the means to attract them. Wealthy and prestigious men have more than one wife; ne'er-do-wells have none. Typically a man who has been married for some time seeks a younger wife. The senior wife remains his confidante and partner and runs the household; the junior one becomes his sexual interest.

In foraging societies wealth cannot accumulate, but a few fierce men, skilled leaders, and good hunters may have two to ten wives. With the invention of agriculture and massive inequality, polygyny can reach ridiculous proportions. Laura Betzig has documented that in civilization after civilization, despotic men have implemented the ultimate male fantasy: a harem of hundreds of nubile women, closely guarded (often by eunuchs) so no other man can touch them. Similar arrangements have popped up in India, China, the Islamic world, sub-Saharan Africa, and the Americas. King Solomon had a thousand concubines. Roman emperors called them slaves, and medieval European kings called them serving maids.

Polyandry, by comparison, is vanishingly rare. Men occasionally share a wife in environments so harsh that a man cannot survive without a woman, but the arrangement collapses when conditions improve. Eskimos have sporadically had polyandrous marriages, but the co-husbands

are always jealous and one often murders the other. As always, kinship mitigates enmity, and among Tibetan farmers two or more brothers sometimes marry a woman simultaneously in the hope of putting together a family that can survive in the bleak territory. The junior brother, though, aspires to have a wife of his own.

Marriage arrangements are usually described from the man's point of view, not because the desires of women are irrelevant but because powerful men have usually gotten their way. Men are bigger and stronger because they have been selected to fight one another, and they can form powerful clans because in traditional societies sons stay near their families and daughters move away. The most florid polygynists are always despots, men who could kill without fear of retribution. (According to the *Guinness Book of World Records*, the man with the most recorded children in history—888—was an emperor of Morocco with the evocative name Moulay Ismail The Bloodthirsty.) The hyperpolygynist not only must fend off the hundreds of men he has deprived of wives, but must oppress his harem. Marriages always have at least a bit of reciprocity, and in most polygynous societies a man may forgo additional wives because of their emotional and financial demands. A despot can keep them imprisoned and terrified.

But oddly enough, in a freer society polygyny is not necessarily bad for women. On financial and ultimately on evolutionary grounds, a woman may prefer to share a wealthy husband than to have the undivided attention of a pauper, and may even prefer it on emotional grounds. Laura Betzig summed up the reason: Would you rather be the third wife of John F. Kennedy or the first wife of Bozo the Clown? Co-wives often get along, sharing expertise and child-care duties, though jealousies among the subfamilies often erupt, much as in stepfamilies but with more factions and adult players. If marriage were genuinely a free market, then in a polygamous society men's greater demand for a limited supply of partners and their inflexible sexual jealousy would give the advantage to women. Laws enforcing monogamy would work to women's disadvantage. The economist Steven Landsburg explains the market principle, using labor instead of money in his example:

Today, when my wife and I argue about who should do the dishes, we start from positions of roughly equal strength. If polygamy were legal, my wife could hint that she's thought about leaving me to many Alan and Cindy down the block—and I might end up with dishpan hands.

. . . Antipolygamy laws are a textbook example of the theory of cartels. Producers, initially competitive, gather together in a conspiracy against the public or, more specifically, against their customers. They agree that each firm will restrict its output in an attempt to keep prices high. But a high price invites cheating, in the sense that each firm seeks to expand its own output beyond what is allowable under the agreement. Eventually, the cartel crumbles unless it is enforced by legal sanctions, and even then violations are legion.

That story, told in every economics textbook, is also the story of male producers in the romance industry. Initially fiercely competitive, they gather together in a conspiracy against their "customers"—the women to whom they offer their hands in marriage. The conspiracy consists of an agreement under which each man restricts his romantic endeavors in an attempt to increase the bargaining position of men in general. But the improved position of men invites cheating, in the sense that each man tries to court more women than allowed under the agreement. The cartel survives only because it is enforced by legal sanctions, and even so violations are legion.

Legal monogamy historically has been an agreement between more and less powerful men, not between men and women. Its aim is not so much to exploit the customers in the romance industry (women) as to minimize the costs of competition among the producers (men): Under polygyny, men vie for extraordinary Darwinian stakes—many wives versus none—and the competition is literally cutthroat. Many homicides and most tribal wars are directly or indirectly about competition for women. Leaders have outlawed polygyny when they needed less powerful men as allies and when they needed their subjects to fight an enemy instead of fighting one another. Early Christianity appealed to poor men partly because the promise of monogamy kept them in the marriage game, and in societies since, egalitarianism and monogamy go together as naturally as despotism and polygyny.

Even today, inequality has allowed a kind of polygyny to flourish. Wealthy men support a wife and a mistress, or divorce their wives at twenty-year intervals and pay them alimony and child support while marrying younger women. The journalist Robert Wright has speculated that easy divorce and remarriage, like overt polygyny, increases violence. Women of childbearing age are monopolized by well-to-do men, and the shortage of potential wives trickles down to the lower strata, forcing the poorest young men into desperate competition.

All of these intrigues come from a single difference between the sexes, men's greater desire for multiple partners. But men are not completely indiscriminate, and women are not voiceless in any but the most despotic societies. Each sex has criteria for picking partners for liaisons and for marriages. Like other staunch human tastes, they appear to be adaptations.

Both sexes want spouses, and men want liaisons more than women do, but that does not mean women never want liaisons. If they never did want them, the male urge to philander could not have evolved because it would never have been rewarded (unless the philanderer could always trick his conquest into thinking he was courting her as a wife—but even then, a married woman should never philander or be a target of philandering). Men's testicles would not have evolved to their larger-than-gorilla proportions, for their sperm would never be in danger of being outnumbered. And jealous feelings directed at wives would not exist; as we shall see, they do exist. The ethnographic record shows that in all societies, both sexes commit adultery, and the women do not always take arsenic or throw themselves under the 5:02 from St. Petersburg.

What could ancestral women have gained from liaisons that would have allowed the desire to evolve? One reward is resources. If men want sex for its own sake, women can make them pay for it. In foraging societies, women openly demand gifts from their lovers, usually meat. You may be offended at the thought that our foremothers gave themselves away for a steak dinner, but to foraging peoples in lean times when high-quality protein is scarce, meat is an obsession. (In *Pygmalion*, when Doolittle tries to sell his daughter Eliza to Higgins, Pickering shouts, "Have you no morals, man?" Doolittle replies, "Can't afford them, Governor. Neither could you if you was as poor as me.") From a distance it sounds like prostitution, but to the people involved it may feel more like ordinary etiquette, much as a woman in our own society might be offended if a wealthier lover never took her out to dinner or spent money on her, though both parties would deny there is a quid pro quo. In questionnaires, female college students report that an extravagant lifestyle and a willingness to give gifts are important qualities in picking a short-term lover, though not in picking a husband.

And like many birds, a woman could seek genes from the best-quality male and investment from her husband, because they are unlikely to be

the same man (especially under monogamy and when she has little say in her marriage). Women report that looks and strength matter more in a lover than in a husband; as we shall see, looks are an indicator of genetic quality. And when women go through with an affair, they generally pick men of higher status than their husbands; the qualities that lead to status are almost certainly heritable (though a taste for prestigious lovers may also help with the first motive, extracting resources). Liaisons with superior men also may allow a woman to test her ability to trade up in the marriage market, either as a prelude to doing so or to improve her bargaining position within the marriage. Symons' summary of the sex difference in adultery is that a woman has an affair because she feels that the man is in some way superior or complementary to her husband, and a man has an affair because the woman is not his wife.

Do men require *anything* in a casual sex partner other than two X chromosomes? Sometimes it would appear that the answer is no. The anthropologist Bronislaw Malinowski reported that some Trobriand Island women were considered so repulsive that they were absolutely debarred from sexual intercourse. These women nevertheless managed to have several offspring, which the Trobrianders interpreted as conclusive proof of virgin birth. But more systematic research has shown that men, at least American college students, do have some preferences in a short-term partner. They rate looks as important; as we shall see, beauty is a signal of fertility and genetic quality. Promiscuity and sexual experience are also rated as assets. As Mae West explained, "Men like women with a past because they hope history will repeat itself." But these assets turn into liabilities when the men are asked about long-term partners. They subscribe to the infamous madonna-whore dichotomy, which divides the female sex into loose women, who may be dismissed as easy conquests, and coy women, who are valued as potential wives. This mentality is often called a symptom of misogyny, but it is the optimal genetic strategy for males of any species that invest in their offspring: mate with any female that will let you, but make sure your consort does not mate with any other male.

What should women look for in a husband? A bumper sticker from the 1970s read, "A woman without a man is like a fish without a bicycle." But at least for women in foraging societies, that would have been an overstatement. When a foraging woman is pregnant, nursing, and bringing up children, she and the children are vulnerable to hunger, protein deficiency, predation, rape, kidnapping, and murder. Any man who

fathers her children should be put to good use in feeding and protecting them. From her point of view, he has nothing better to do, though from his point of view, there is an alternative: competing for and wooing other women. Men vary in their ability and willingness to invest in their children, so a woman should choose wisely. She should be impressed by wealth and status, or, in the case of men too young to have them, by portents that they will get them, such as ambition and industriousness. These are all useless unless the man hangs around once the woman becomes pregnant, and men have an interest in saying they will hang around whether or not they intend to. As Shakespeare wrote, "Men's vows are women's traitors." A woman therefore should look for signs of stability and sincerity. An aptitude for bodyguard duty would also come in handy.

What should men look for in a wife? Aside from faithfulness, which guarantees his paternity, she should be able to bear as many children as possible. (As always, that would be how our tastes were engineered; the reasoning does not imply that a man literally wants lots and lots of babies.) She should be fertile, which means she should be healthy and past the age of puberty but before the age of menopause. But a woman's current fertility is more relevant to a one-night stand than to a lifelong marriage. What counts is the number of offspring he can expect over the long term. Since a woman can bear and nurse one child every few years, and her childbearing years are finite, the younger the bride, the bigger the future family. That is true even though the youngest brides, teenagers, are somewhat less fertile than women in their early twenties. Ironically for the men-are-slime theory, an eye for nubile women may have evolved in the service of marriage and fatherhood, not one-night stands. Among chimpanzees, where a father's role ends with copulation, some of the wrinkled and saggy females are the sexiest.

Are the predictions just old-fashioned stereotypes? Buss designed a questionnaire asking about the importance of eighteen qualities of a mate and gave it to ten thousand people in thirty-seven countries on six continents and five islands—monogamous and polygynous, traditional and liberal, communist and capitalist. Men and women everywhere place the highest value of all on intelligence and on kindness and understanding. But in every country men and women differ on the other qualities. Women value earning capacity more than men do; the size of the difference varies from a third more to one and a half times more, but it's always there. In virtually every country, women place a greater value than

men on status, ambition, and industriousness. And in most, they value dependability and stability more than men do. In every country, men place a higher value on youth and on looks than women do. On average, men want a bride 2.66 years younger; women want a groom 3.42 years older. The results have been replicated many times.

People's actions tell the same story. According to the contents of personal advertisements, Men Seeking Women seek youth and looks, Women Seeking Men seek financial security, height, and sincerity. The owner of one dating service observed, "Women really read over our profile forms; guys just look at the pictures." Among married couples, the husband is 2.99 years older than the wife, as if they had split the difference between their preferences. In foraging cultures, everyone agrees that some people are sexier than others, and the sexpots are usually young women and prestigious men. Yanomamø men, for example, say that the most desirable women are *moho dudei*, an expression that when applied to fruit means perfectly ripe and when applied to women means between fifteen and seventeen years old. When shown slides, Western observers of both sexes agree with the Yanomamø men that the *moho dudei* women are the most attractive. In our society, the best predictor of a man's wealth is his wife's looks, and the best predictor of a woman's looks is her husband's wealth. Dumpy-looking cabinet secretaries like Henry Kissinger and John Tower are called sex symbols and womanizers. Octogenarian oil barons like J. Paul Getty and J. Howard Marshall marry women young enough to be their great-granddaughters, such as the model Anna Nicole Smith. Not-so-handsome rock stars like Billy Joel, Rod Stewart, Lyle Lovett, Rick Ocasek, Ringo Starr, and Bill Wyman marry gorgeous actresses and supermodels. But former Representative Patricia Schroeder says she has noticed that a middle-aged congresswoman does not radiate the same animal magnetism to the opposite sex that a middle-aged congressman does.

An obvious retort is that women value wealthy and powerful men because it is the men who have the wealth and power. In a sexist society, women have to marry up to get them. That alternative has been tested and refuted. Women with large salaries, postgraduate degrees, prestigious professions, and high self-esteem place a *greater* value on wealth and status in a husband than other women do. So do the leaders of feminist organizations. Poor men place no higher value on wealth or earning power in a wife than other men do. Among the Bakweri in Cameroon, the women are wealthier and more powerful than the men, and they still insist on men with money.

The humorist Fran Lebowitz once said in an interview, "People who get married because they're in love make a ridiculous mistake. It makes much more sense to marry your best friend. You *like* your best friend more than anyone you're ever going to be in love with. You don't choose your best friend because they have a cute nose, but that's all you're doing when you get married; you're saying, 'I will spend the rest of my life with you because of your lower lip.'"

It is a puzzle, and the obvious place to look for an answer is the fact that you don't make children with your best friend but you do with your spouse. Perhaps we care about a few millimeters of flesh here or there because it is a perceptual signal of a deeper trait that cannot be measured directly: how well equipped the person's body is to serve as the other parent of your children. Fitness as a dam or stud is like any other feature of the world. It is not written on a tag but has to be inferred from appearances, using assumptions about how the world works.

Could we really be equipped with an innate eye for beauty? What about the natives in *National Geographic* who file their teeth, stretch their necks with stacks of rings, burn scars into their cheeks, and put plates in their lips? What about the fat women in the Rubens paintings and Twiggy in the 60s? Don't they show that standards of beauty are arbitrary and vary capriciously? They do not. Who says that *everything* people do to their bodies is an attempt to look sexy? That is the tacit assumption behind the *National Geographic* argument, but it's obviously false. People decorate their bodies for many reasons: to look rich, to look well connected, to look tough, to look "in," to earn membership in an elite group by enduring a painful initiation. Sexual attractiveness is different. People outside a culture usually agree with the people inside about who is beautiful and who is not, and people everywhere want good-looking partners. Even three-month-old infants prefer to look at a pretty face.

What goes into sexiness? Both sexes want a spouse who has developed normally and is free of infection. Not only is a healthy spouse vigorous, non-contagious, and more fertile, but the spouse's hereditary resistance to the local parasites will be passed on to the children. We haven't evolved stethoscopes and tongue-depressors, but an eye for beauty does some of the same things. Symmetry, an absence of deformities, cleanliness, unblemished skin, clear eyes, and intact teeth are attractive in all cultures. Orthodontists

have found that a good-looking face has teeth and jaws in the optimal alignment for chewing. Luxuriant hair is always pleasing, possibly because it shows not only current health but a record of health in the years before. Malnutrition and disease weaken the hair as it grows from the scalp, leaving a fragile spot in the shaft. Long hair implies a long history of good health.

A subtler sign of good genes is being average. Not average in attractiveness, of course, but average in the size and shape of every part of the face. The average measurement of a trait in a local population is a good estimate of the optimal design favored by natural selection. If people form a composite of the opposite-sex faces around them, they would have an ideal of the fittest mate against which any candidate could be matched. The exact facial geometry of the local race or ethnic group would not need to be built in. In fact, composite faces, whether formed by superimposing negatives in an enlarger or by sophisticated computer-graphics algorithms, are prettier or handsomer than the individual faces that went into them.

Average faces are a good start, but some faces are even more attractive than the average face. When boys reach puberty, testosterone builds up the bone in their jaws, brows, and nasal region. Girls' faces grow more evenly. The difference in 3-D geometry allows us to tell a man's head from a woman's even when they are both bald and shaved. If the geometry of a woman's face is similar to a man's, she is homelier; if it is less similar, she is prettier. Beauty in a woman comes from a short, delicate, smoothly curved jawbone, a small chin, a small nose and upper jaw, and a smooth forehead without brow ridges. The "high cheekbones" of a beautiful woman are not bones at all but soft tissue, and contribute to beauty because the other parts of a beautiful face (the jaws, forehead, and nose) are small by comparison.

Why are masculine-looking women less "attractive? If a woman's face is masculinized, she probably has too much testosterone in her blood (a symptom of many diseases); if she has too much testosterone, she is likely to be infertile. Another explanation is that prettiness-detectors are really female-face detectors, designed to pick them out from every other object in the world and tuned to minimize the risk of a false alarm to a male face, which is the object most similar to a female face. The more unmanly the face, the louder the detector beeps. Similar engineering could explain why men with unfeminine faces are more handsome. A man with a large, angular jaw, a strong chin, and a prominent forehead and brow is undoubtedly an adult male with normal male hormones.

By the callous reckoning of natural selection, young women who have not yet had children are the best wives, because they have the longest reproductive career ahead of them and have no children from another man tagging along. Signs of youth and signs of never having been pregnant should make a woman prettier. Teenage women have larger eyes, fuller and redder lips, smoother, moister, and tighter skin, and firmer breasts, all long recognized as ingredients of pulchritude. Aging lengthens and coarsens a woman's facial bones, and so do pregnancies. Therefore a small-jawed, light-boned face is a clue to four reproductive virtues: being female, having the right hormones, being young, not having been pregnant. The equation of youth and beauty is often blamed on America's being obsessed with youth, but by that reasoning every culture is obsessed with youth. If anything, contemporary America is less youth-oriented. The age of *Playboy* models has *increased* over the decades, and in most times and places women in their twenties have been considered over the hill. Men's looks don't decline as quickly when they age, not because of a double standard in our society but because men's fertility doesn't decline as quickly when they age.

At puberty a girl's hips become wider because her pelvis grows and because fat is deposited on her hips, a reserve of calories available to supply the body during pregnancy. The ratio of waist size to hip size decreases in most fertile women to between .67 and .80, whereas the ratio for most men, children, and postmenopausal women is between .80 and .95. Among women, a low waist-to-hip ratio has been found to correlate with youth, health, fertility, not being pregnant, and never having been pregnant. The psychologist Devendra Singh has shown photographs and computer-generated pictures of female bodies of different sizes and shapes to hundreds of people of various ages, sexes, and cultures. Everyone finds a ratio of .70 or lower the most attractive. The ratio captures the old idea of the hourglass figure, the wasp waist, and the 36-24-36 ideal measurements. Singh also measured the ratio in *Playboy* centerfolds and winners of beauty contests over seven decades. Their weight has gone down, but their waist-to-hip ratio has stayed the same. Even most of the Upper Paleolithic Venus figurines, carved tens of thousands of years ago, have the right proportions.

The geometry of beauty once was an indicator of youth, health, and nonpregnancy, but it no longer has to be. Women today have fewer babies, have them later, are less exposed to the elements, and are better nourished and less disease-ridden than their ancestors. They can look

like an ancestral teenager well into middle age. Women also have a technology to simulate and exaggerate the clues to youth, femaleness, and health: eye makeup (to enlarge the eyes), lipstick, eyebrow plucking (to reduce the appearance of a masculine brow ridge), makeup (to exploit the shape-from-shading mechanism of Chapter 4), products that increase the luster, thickness, and color of hair, bras and clothing that simulate young breasts, and hundreds of potions alleged to keep the skin looking young. Dieting and exercise can keep the waist thinner and the waist-to-hip ratio lower, and an illusion can be engineered with bodices, corsets, hoops, crinolines, bustles, girdles, pleats, tapering, and wide belts. Women's fashion has never embraced bulky cummerbunds.

Outside the scientific literature, more has been written about women's weight than any other aspect of beauty. In the West, women in pictures have weighed less and less over the past decades. That has been taken as evidence for the arbitrariness of beauty and for the oppression of women, who are expected to conform to these standards no matter how unreasonable. Slender models are commonly blamed for anorexia nervosa in teenage girls, and a recent book was called *Fat Is a Feminist Issue*. But weight may be the least important part of beauty. Singh found that very fat women and very thin women are judged less attractive (and in fact they are less fertile), but there is a range of weights considered attractive, and shape (waist-to-hip ratio) is more important than size. The hoopla about thinness applies more to women who pose for other women than to women who pose for men. Twiggy and Kate Moss are fashion models, not pinups; Marilyn Monroe and Jayne Mansfield were pinups, not fashion models. Weight is a factor mostly in the competition among women for status in an age in which wealthy women are more likely to be slender than poor ones, a reversal of the usual relation.

Still, the women posing for both sexes today are slimmer today than their historical counterparts, and it may be for reasons other than just changes in the signs of status. My own conjecture is that today's slender centerfolds and supermodels would not have had trouble finding a date at any time in history, because they are *not* like the skinny women eschewed in centuries past. Body parts do not vary independently. Tall men tend to have big feet, people with thick waists tend to have double chins, and so on. Undernourished women may tend to have more masculine bodies, and well-nourished ones more feminine bodies, so historically attractive women may have tended to be heavier. Neither kind of woman has the most beautiful shape conceivable—say, Jessica

Rabbit's—because real bodies did not evolve as cartoon sex lures. They are compromises among the demands of attractiveness, running, lifting, childbearing, nursing, and surviving famines. Perhaps modern technology has fabricated a sex lure, not with a cartoonist's brush but with artificial selection. In a world of five billion people there are bound to be women with wide feet and small heads, men with big ears and scrawny necks, and any other combination of body parts you want to specify. There may be a few thousand women with freakish combinations of small waists, flat abdomens, large firm breasts, and curved but medium-sized hips—optical illusions that send the needles of people's fertility and childlessness gauges into the red. When word gets around that they can parlay their freaky bodies into fame and fortune, they come out of the woodwork, and enhance their gifts with makeup, exercise, and glamour photography. The bodies in the beer commercials may be unlike anything seen in history.

Beauty is not, as some feminists have claimed, a conspiracy by men to objectify and oppress women. The *really* sexist societies drape women in chadors from head to foot. Throughout history the critics of beauty have been powerful men, religious leaders, sometimes older women, and doctors, who can always be counted on to say that the latest beauty craze is hazardous to women's health. The enthusiasts are women themselves. The explanation is simple economics and politics (though not the orthodox feminist analysis—quite insulting to women, incidentally—in which women are dupes who have been brainwashed into striving for something they don't want). Women in open societies want to look good because it gives them an edge in competing for husbands, status, and the attention of powerful people. Men in closed societies hate beauty because it makes their wives and daughters indiscriminately attractive to other men, giving the women a measure of control over the profits from their own sexuality and taking it away from the men (and, in the case of daughters, away from their mothers). Similar economics make men want to look good, too, but the market forces are weaker or different because men's looks matter less to women than women's looks matter to men.

Though the beauty industry is not a conspiracy against women, it is not innocuous either. We calibrate our eye for beauty against the people we see, including our illusory neighbors in the mass media. A daily diet of freakishly beautiful virtual people may recalibrate the scales and make the real ones, including ourselves, look ugly.

For humans, like birds, life is complicated because of two of their reproductive habits. Males invest in their offspring, but fertilization happens out of sight inside the female's body, so a male never knows which offspring are his. A female, in contrast, can be certain that any egg or baby coming out of her body carries her genes. A cuckolded male is worse than a celibate one in the evolutionary struggle, and male birds have evolved defenses against it. So have humans. Sexual jealousy is found in all cultures.

Both sexes can feel intense jealousy at the thought of a dallying mate, but their emotions are different in two ways. Women's jealousy appears to be under the control of more sophisticated software, and they can appraise their circumstances and determine whether the man's behavior poses a threat to their ultimate interests. Men's jealousy is cruder and more easily triggered. (Once triggered, though, women's jealousy appears to be as intensely felt as men's.) In most societies, some women readily share a husband, but in no society do men readily share a wife. A woman having sex with another man is *always* a threat to the man's genetic interests, because it might fool him into working for a competitor's genes, but a man having sex with another woman is not necessarily a threat to the woman's genetic interests, because his illegitimate child is another woman's problem. It is only a threat if the man diverts investment from her and her children to the other woman and her children, either temporarily or, in the case of desertion, permanently.

So men and women should be jealous of different things. Men should squirm at the thought of their wives or girlfriends having sex with another man; women should squirm at the thought of their husbands or boyfriends giving time, resources, attention, and affection to another woman. Of course no one likes to think of their mate offering sex *or* affection to anyone else, but even then the reasons may differ: men may be upset about affection because it could lead to sex; women may be upset about sex because it could lead to affection. Buss found that men and women are made as jealous by the thought of alienated sex as by the thought of alienated affection, but when asked to pick their torture, most men said they were more upset by the thought of their partner being sexually unfaithful than emotionally unfaithful, and most women had the opposite reaction. (The same differences are found when men and women imagine their partners being both sexually *and* emotionally

unfaithful and are asked which aspect of the betrayal bothers them more. That shows that the sex difference is not just a matter of men and women having different expectations of their partners' behavior, the men worrying that a woman having sex must also be in love and the women worrying that a man in love must also be having sex.) Buss then pasted electrodes on people and asked them to imagine the two kinds of treachery. The men sweated, frowned, and palpitated more from images of sexual betrayal; the women sweated, frowned, and palpitated more from images of emotional betrayal. (I cited the experiment in Chapter 4 as an illustration of the power of mental images.) Similar results have been found in several countries in Europe and Asia.

It takes two to commit adultery, and men, always the more violent sex, have directed their anger at both parties. The largest cause of spousal abuse and spousal homicide is sexual jealousy, almost always the man's. Men beat and kill their wives and girlfriends to punish them for real or imagined infidelity and to deter them from becoming unfaithful or leaving them. Women beat and kill their husbands in self-defense or after years of abuse. Critics of feminism have made much of the occasional statistic that American men are victims of beating and homicide by their spouses almost as often as the women are. But that's not true in the vast majority of communities, and even in the few where it is, the husband's jealousy and intimidation are almost always the cause. Often a morbidly jealous man will imprison his wife in the house and interpret every incoming phone call as proof that she is unfaithful. Women are most at risk when they threaten to leave or do it. The forsaken man may stalk her, hunt her down, and execute her, always with the same rationale: "If I can't have her, no one can." The crime is pointless, but it is the undesired outcome of a paradoxical tactic, a doomsday machine. For every killing of an estranged wife or girlfriend there must be thousands of threats made credible by signs that the man is crazy enough to carry them out regardless of the cost.

Many pundits blame violence against women on this or that feature of American society, such as circumcision, war toys, James Bond, or football. But it happens worldwide, including in foraging societies. Among the Yanomamö, a man who suspects his wife of infidelity might slash her with a machete, shoot her with an arrow, hold an ember against her, cut off her ears, or kill her. Even among the idyllic !Kung San of the Kalahari Desert in southern Africa, men batter wives they suspect of being unfaithful. Incidentally, none of these points "condone" the violence or

imply that "it's not the man's fault," as it is sometimes claimed. Those non sequiturs could be attached to *any* explanation, such as the common feminist theory that men are brainwashed by media images that glorify violence against women.

All over the world, men also beat and kill cuckolds and suspected cuckolds. Recall that rivalry over women is the leading cause of violence, homicide, and warfare among foraging peoples. As it is written in Proverbs 6:34, "For jealousy is the rage of a man: therefore he will not spare in the day of vengeance."

Unlike birds, though, humans plug their sexual jealousy into a baroque cognitive machine. People think in metaphors, and the metaphor that men have always used for wives is property. In their essay "The Man Who Mistook His Wife for a Chattel," Wilson and Daly show that men do not merely aim to control their wives and fend off rivals; they assert an *entitlement* to wives, especially their reproductive capacity, identical to the right of an owner over inanimate property. An owner can sell, exchange, or dispose of his possessions, can modify them without interference, and can demand redress for theft or damage. These rights are recognized by the rest of society and can be enforced by collective reprisals. In culture after culture, men have deployed the full cognitive apparatus of ownership in conceiving of their relationship to their wives, and until recently they have formalized the metaphor in codes of law.

In most societies, marriage is a blatant transfer of ownership of a woman from her father to her husband. In our own marriage ceremony, the father of the bride still "gives her away," but more commonly he sells her. In seventy percent of societies, someone pays when two people get married. In ninety-six percent of these, the groom or his family pays the bride's family, sometimes in cash or a daughter, sometimes in bride-service, whereby the groom works for the bride's father for a fixed period. (In the Bible, Jacob worked for Laban for seven years for the right to marry his daughter Rachel, but Laban substituted his other daughter, Leah, at the wedding, so Jacob had to work *another* seven years to acquire Rachel as his second wife.) Dowries, which are more familiar to us, are not a mirror image of bride-wealth, because they go to the newlyweds, not to the bride's parents. The husband notifies other men of his ownership in customs retained by many modern couples. The woman, not the man, wears an engagement ring, bears her spouse's surname, and is given a new form of address, *Mrs.*, short for "mistress of."

People can control their property, and husbands (and before them,

fathers and brothers) have controlled women's sexuality. They have used chaperones, veils, wigs, chadors, segregation by sex, confinement, foot-binding, genital mutilation, and the many ingenious designs for chastity belts. Despots not only kept harems but kept them guarded. In traditional societies, "protecting a woman" was a euphemism for keeping her chaste. (Mae West observed, "Men always say they're protecting you, but they never say from what.") Only fertile women were controlled in these ways; children and postmenopausal women had more freedom.

The word *adultery* is related to the word *adulterate* and refers to making a woman impure by introducing an improper substance. The infamous double standard, in which a married woman's philandering is punished more severely than a married man's, is common in legal and moral codes in all kinds of societies. Its rationale was succinctly captured when James Boswell remarked, "There is a great difference between the offence of infidelity in a man and that of his wife," and Samuel Johnson replied, "The difference is boundless. The man imposes no bastards on his wife." Both the married woman and her lover are commonly punishable (often by death), but the symmetry is illusory, because it is the woman's marital status, not the man's, that makes it a crime, specifically, a crime against her husband. Until recently most of the world's legal systems treated adultery as a property violation or tort. The husband was entitled to damages, a refund of the bride-price, a divorce, or the right to violent revenge. Rape was an offense against the woman's husband, not against the woman. Elopement was considered an abduction of a daughter from her father. Until very recently, the rape of a woman by her husband was not a crime, or even a coherent concept: husbands were entitled to sex with their wives.

Throughout the English-speaking world, the common law recognizes three circumstances that reduce murder to manslaughter: self-defense, the defense of close relatives, and sexual contact with the man's wife. (Wilson and Daly observe that they are the three main threats to Darwinian fitness.) In several American states, including Texas as recently as 1974, a man who discovered his wife in flagrante delicto and killed her lover was not guilty of a crime. Even today, in many places those homicides are not prosecuted or the killer is treated leniently. Jealous rage at the sight of a wife's adultery is cited as one of the ways a "reasonable man" can be expected to behave.

I wish I could have discussed the evolutionary psychology of sexuality without the asides about feminist theory, but in today's intellectual climate that is impossible. The Darwinian approach to sex is often attacked as being antifeminist, but that is just wrong. Indeed, the accusation is baffling on the face of it, especially to the many feminist women who have developed and tested the theory. The core of feminism is surely the goal of ending sexual discrimination and exploitation, an ethical and political position that is in no danger of being refuted by any foreseeable scientific theory or discovery. Even the spirit of the research poses no threat to feminist ideals. The sex differences that have been documented are in the psychology of reproduction, not in economic or political worth, and they are invidious with regard to men, not women. The differences should heighten awareness of incest, exploitation, harassment, stalking, battering, rape (including date rape and marital rape), and legal codes that discriminate against women. If they show that men are especially tempted to commit certain crimes against women, the implication is that the deterrents should be surer and more severe, not that the crimes are somehow less odious. Even evolutionary explanations of the traditional division of labor by sex do not imply that it is unchangeable, "natural" in the sense of good, or something that should be forced on individual women or men who don't want it.

What evolutionary psychology challenges is not the goals of feminism, but parts of the modern orthodoxy about the mind that have been taken up by the intellectual establishment of feminism. One idea is that people are designed to carry out the interests of their class and sex, rather than to act out of their own beliefs and desires. A second is that the minds of children are formed by their parents, and the minds of adults are formed by language and by media images. A third is the romantic doctrine that our natural inclinations are good and that ignoble motives come from society.

The unstated premise that nature is nice lies behind many of the objections to the Darwinian theory of human sexuality. Carefree sex is natural and good, it is assumed, so if someone claims that men want it more than women do, it would imply that men are mentally healthy and women neurotic and repressed. That conclusion is unacceptable; so the claim that men want carefree sex more than women do cannot be correct. Similarly, sexual desire is good, so if men rape for sex (rather than to

express anger towards women), rape would not be as evil. Rape is evil; therefore the claim that men rape for sex cannot be correct. More generally, what people instinctively like is good, so if people like beauty, beauty would be a sign of worth. Beauty is not a sign of worth, so the claim that people like beauty cannot be correct.

These kinds of arguments combine bad biology (nature is nice), bad psychology (the mind is created by society), and bad ethics (what people like is good). Feminism would lose nothing by giving them up.

RIVALS

People everywhere strive for a ghostly substance called authority, cachet, dignity, dominance, eminence, esteem, face, position, preeminence, prestige, rank, regard, repute, respect, standing, stature, or status. People go hungry, risk their lives, and exhaust their wealth in pursuit of bits of ribbon and metal. The economist Thorstein Veblen noticed that people sacrificed so many necessities of life to impress one another that they appear to be responding to a "higher, spiritual need." Status and virtue are close in people's minds, as we see in words like *chivalrous*, *classy*, *courtly*, *gentlemanly*, *honorable*, *noble*, and *princely*, and their opposites *ill-bred*, *low-class*, *low-rent*, *mean*, *nasty*, *rude*, *shabby*, and *shoddy*. When it comes to the trifles of personal appearance, we express our admiration for the tasteful using ethical metaphors such as *right*, *good*, *correct*, and *faultless*, and censure the tacky with tones usually reserved for sin—an attitude that the art historian Quentin Bell dubbed "sartorial morality."

Is this any way to build an intelligent organism? Where do these powerful motives come from?

Many animals are moved by pointless decorations and rituals, and the selective causes are no longer mysterious. Here is the key idea. Creatures differ in their ability to hurt and help others. Some are stronger or fiercer or more poisonous; some have better genes or more largesse. These potent creatures want everyone to know they are potent, and the creatures they can impinge on *also* want to know which ones are potent. But it is impossible for every creature to probe every other one's DNA, muscle mass, biochemical composition, ferocity, and so on. So the consequential creatures advertise their worth with a signal. Unfortunately, the inconsequential creatures can counterfeit the signal and reap the

benefits, debasing its value to everyone else. The race is on for the consequential creatures to cook up a display that is hard to counterfeit, for the less consequential ones to become better counterfeiters, and for the third parties to sharpen their powers of discrimination. Like paper currency, the signals are inimitably gaudy and intrinsically worthless, but are treated as if they were valuable and *are* valuable because everyone treats them that way.

The precious stuff behind the displays can be divided into dominance—who can hurt you—and status—who can help you. They often go together, because people who can hurt you can also help you by their ability to hurt others. But it's convenient to look at them separately.

.Most people have heard of the dominance hierarchies, pecking orders, and alpha males that are widespread in the animal kingdom. Animals of the same species don't fight to the death every time they contest something of value. They have a ritualized fight or a show of arms or a staring match, and one backs down. Konrad Lorenz and other early ethologists thought that gestures of surrender helped preserve the species against internecine bloodshed, and that humans were in peril because we lost the gestures. But that idea comes from the fallacy that animals evolve to benefit the species. It cannot explain why a truculent mutant that never surrendered and that killed surrenderers would not walk over the competition and soon characterize the species. The biologists John Maynard Smith and Geoffrey Parker came up with a better explanation by modeling how the different aggressive strategies that animals might adopt would stack up against each other and against themselves.

Fighting every contest to the bitter end is a poor strategy for an animal, because chances are its adversary has evolved to do the same thing. A fight is costly to the loser, because it will be injured or dead and hence worse off than if it had relinquished the prize from the start. It also can be costly to the victor because he may sustain injuries in the course of victory. Both parties would have done better if they had assessed who was likely to win beforehand and if the underdog simply conceded. So animals size each other up to see who's bigger, or brandish their weapons to see whose are more dangerous, or wrestle until it's clear who's stronger. Though only one animal wins, both walk away. The loser con-

cedes because he can seek his fortunes elsewhere or bide his time until circumstances are more propitious. When animals size each other up, they evolve ways to exaggerate their size: ruffs, balloons, manes, bristling, rearing, and bellowing, whose low pitch shows off the size of the resonating cavity in the animal's body. If a fight is costly and a winner unpredictable, the faceoff may be decided by an arbitrary difference such as who arrived first, in the same way that human rivals may settle a dispute quickly by flipping a coin. If the animals are closely matched and the stakes are high enough (such as a harem), an all-out fight may ensue, sometimes to the death.

If both creatures walk away, they may remember the outcome and thereafter the loser will defer to the winner. When many animals in a group spar or size one another up in a round-robin, the outcome is a pecking order, which correlates with the probability that each animal would win an all-out duel. When the probabilities change—say, when a dominant animal gets old or injured, or an underling gains in strength or experience—the underling may mount a challenge and the rankings may change. In chimpanzees, dominance depends not only on fighting prowess but on political acumen: a pair in cahoots may depose a stronger animal going it alone. Many group-living primates settle into two dominance hierarchies, one for each sex. The females compete for food; the males compete for females. Dominant males mate more often, both because they can shove other males out of the way and because the females prefer to mate with them, if for no other reason than that a high-ranking sex partner will tend to sire high-ranking sons, who will give the female more grandchildren than low-ranking sons.

Humans don't have rigid pecking orders, but in all societies people recognize a kind of dominance hierarchy, particularly among men. High-ranking men are deferred to, have a greater voice in group decisions, usually have a greater share of the group's resources, and always have more wives, more lovers, and more affairs with other men's wives. Men strive for rank, and achieve it in some ways that are familiar from zoology books and other ways that are uniquely human. Better fighters have higher rank, and men who *look* like better fighters have higher rank. Sheer height is surprisingly potent in a species that calls itself the rational animal. The word for "leader" in most foraging societies is "big man," and in fact the leaders usually *are* big men. In the United States, taller men are hired more, are promoted more, earn more (\$600 per inch in annual salary), and are elected president more: the taller candidate won

twenty of the twenty-four elections between 1904 and 1996. A glance at the personal ads shows that women want taller men. As in other species whose males compete, the human male is bigger than the female, and has evolved ways of appearing bigger still, like a low voice and a beard (which makes the head look bigger and has evolved separately in lions and monkeys). Leonid Brezhnev claimed that he got to the top because of his eyebrows! Men everywhere exaggerate the size of their heads (with hats, helmets, headdresses, and crowns), their shoulders (with pads, boards, epaulettes, and feathers), and, in some societies, their penises (with impressive codpièces and sheaths, sometimes a yard long).

But humans also evolved language and a new way of propagating information about dominance: reputation. Sociologists have long been puzzled that the largest category of motives for homicide in American cities is not robbery, drug deals gone sour, or other tangible incentives. It is a category they call "altercation of relatively trivial origin; insult, curse, jostling, etc." Two young men argue over who gets to use the pool table in a bar. They shove each other and trade insults and obscenities. The loser, humiliated before onlookers, storms off and returns with a gun. The murders are the epitome of "senseless violence," and the men who commit them are often written off as madmen or animals.

Daly and Wilson point out that these men behave as if a great deal more is at stake than the use of a pool table. And a great deal more *is* at stake:

Men are known by their fellows as "the sort who can be pushed around" and "the sort who won't take any shit," as people whose word means action or people who are full of hot air, as guys whose girlfriends you can chat up with impunity or guys you don't want to mess with.

In most social milieus, a man's reputation depends in part upon the maintenance of a credible threat of violence. Conflicts of interest are endemic to society, and one's interests are likely to be violated by competitors unless those competitors are *deterred*. Effective deterrence is a matter of convincing our rivals that any attempt to advance their interests at our expense will lead to such severe penalties that the competitive gambit will end up a net loss which should never have been undertaken.

The credibility of the deterrent can be devalued by a public challenge that is not taken up, even if nothing tangible is at stake. Moreover, if a challenger knew that his target was a cool calculator of costs and bene-

fits, he could extort him into backing down with the threat of a fight that was dangerous to both. But a hothead who would stop at nothing to preserve his reputation (a doomsday machine) is unextortable.

The ghetto gang member who stabs the guy who dissed him has honorable counterparts in all the world's cultures. The very meaning of the word *honor* in many languages (including one of its senses in English) is a determination to avenge insults, with bloodshed if necessary. In many foraging societies a boy achieves manly status only after he has killed. A man's respect increases with his verified body count, giving rise to charming customs like scalping and headhunting. Dueling between "men of honor" was traditional in the American South, and many men rose to leadership with the help of their success in duels. The man on the ten-dollar bill, Secretary of the Treasury Alexander Hamilton, was killed in a duel by Vice President Aaron Burr, and the man on the twenty, President Andrew Jackson, won two duels and tried to provoke others.

Why don't we see periodontists or college professors dueling over a parking space? First, they live in a world in which the state has a monopoly on the legitimate use of violence. In places beyond the reach of the state, like urban underworlds or rural frontiers, or in times when the state did not exist, like the foraging bands in which we evolved, a credible threat of violence is one's only protection. Second, the assets of periodontists and professors, such as houses and bank accounts, are hard to steal. "Cultures of honor" spring up when a rapid response to a threat is essential because one's wealth can be carried away by others. They develop among herders, whose animals can be stolen, more often than among crop-growers, whose land stays put. And they develop among people whose wealth is in other liquid forms, like cash or drugs. But perhaps the biggest reason is that periodontists and professors are not male, poor, and young.

Maleness is by far the biggest risk factor for violence. Daly and Wilson report thirty-five samples of homicide statistics from fourteen countries, including foraging and preliterate societies and thirteenth-century England. In all of them, men kill men massively more often than women kill women—on average, twenty-six times more often.

Also, the poolhall avengers and their victims are nobodies: uneducated, unmarried, unprosperous, and often unemployed. Among polygynous mammals such as ourselves, reproductive success varies enormously among males, and the fiercest competition can be at the bottom, among

males whose prospects teeter between zero and nonzero. Men attract women by their wealth and status, so if a man doesn't have them and has no way of getting them he is on a one-way road to genetic nothingness. As with birds that venture into dangerous territories when they are near starvation, and hockey coaches that pull the goalie for an extra skater when they are a goal down with a minute to play, an unmarried man without a future should be willing to take any risk. As Bob Dylan pointed out, "When you got nothing, you got nothing to lose."

Youth makes matters even worse. The population geneticist Alan Rogers has calculated from actuarial data that young men should discount the future steeply, and so they do. Young men commit crimes, drive too fast, ignore illnesses, and pick dangerous hobbies like drugs, extreme sports, and surfing on the roofs of tram cars and elevators. The combination of maleness, youth, penury, hopelessness, and anarchy makes young men indefinitely reckless in defending their reputation.

And it's not so clear that professors (or people in any competitive profession) *don't* duel over pool tables, figuratively speaking. Academics are known by their fellows as "the sort who can be pushed around" and "the sort who won't take any shit," as people whose word means action or people who are full of hot air, as guys whose work you can criticize with impunity or guys you don't want to mess with. Brandishing a switchblade at a scholarly conference would somehow strike the wrong note, but there is always the stinging question, the devastating riposte, the moralistic outrage, the withering invective, the indignant rebuttal, and means of enforcement in manuscript reviews and grant panels. Scholarly institutions, of course, try to minimize this rutting, but it is hard to eradicate. The goal of argumentation is to make a case so forceful (note the metaphor) that skeptics are *coerced* into believing it—they are powerless to deny it while still claiming to be rational. In principle, it is the ideas themselves that are, as we say, compelling, but their champions are not always averse to helping the ideas along with tactics of verbal dominance, among them intimidation ("Clearly . . ."), threat ("It would be unscientific to . . ."), authority ("As Popper showed . . ."), insult ("This work lacks the necessary rigor for . . ."), and belittling ("Few people today seriously believe that . . ."). Perhaps this is why H. L. Mencken wrote that "college football would be more interesting if the faculty played instead of the students."

Status is the public knowledge that you possess assets that would allow you to help others if you wished to. The assets may include beauty, irreplaceable talent or expertise, the ear and trust of powerful people, and especially wealth. Status-worthy assets tend to be fungible. Wealth can bring connections and vice versa. Beauty can be parlayed into wealth (through gifts or marriage), can attract the attention of important people, or can draw more suitors than the beautiful one can handle. Asset-holders, then, are not just seen as holders of their assets. They exude an aura or charisma that makes people want to be in their graces. It's always handy to have people want to be in your graces, so status itself is worth craving. But there are only so many hours in the day, and sycophants must choose whom to fawn over, so status is a limited resource. If A has more, B must have less, and they must compete.

Even in the dog-eat-dog world of tribal leadership, physical dominance is not everything. Chagnon reports that some Yanomamö headmen are flamboyant bullies but others achieve their station by shrewdness and discretion. A man named Kaobawä, though no wimp, earned his authority by leaning on the support of his brothers and cousins and cultivating alliances with the men with whom he had traded wives. He conserved his authority by giving orders only when he was sure everyone would follow them, and magnified it by breaking up fights, disarming machete-wielding maniacs, and bravely scouting the village alone when raiders were in evidence. His quiet leadership was rewarded with six wives and as many affairs. In foraging societies, status also clings to good hunters and knowledgeable naturalists. Assuming that our ancestors, too, practiced occasional meritocracy, human evolution was not always the survival of the fiercest.

Romantic anthropologists used to claim that foraging peoples were unmoved by wealth. But that is because the foragers they studied didn't have any. Twentieth-century hunter-gatherers are unrepresentative of humanity in one respect. They live on land that no one else wants, land that cannot be farmed. They don't necessarily prefer their deserts, rainforests, and tundras, but farming peoples like us have taken the rest. Though foragers cannot achieve the massive inequality that comes from cultivating and storing food, they do have inequality, both of wealth and of prestige.

The Kwakiutl of the Canadian Pacific coast enjoyed annual runs of salmon and abundant sea mammals and berries. They settled in villages run by wealthy chiefs who tried to outdo one another in competitive feasts called potlatches. The guests at a potlatch were encouraged to gorge themselves on salmon and berries, and the chief boastfully showed them with boxes of oil, baskets of berries, and piles of blankets. The humiliated guests slunk back to their village and plotted revenge with an even bigger feast, in which they would not only give away valuables but ostentatiously destroy them. The chief would start a roaring fire in the center of his house and stoke it with fish oil, blankets, furs, canoe paddles, canoes, and sometimes the house itself, a spectacle of consumption the world would not see again until the American bar mitzvah.

Veblen proposed that the psychology of prestige was driven by three "pecuniary canons of taste": conspicuous leisure, conspicuous consumption, and conspicuous waste. Status symbols are flaunted and coveted not necessarily because they are useful or attractive (pebbles, daisies, and pigeons are quite beautiful, as we rediscover when they delight young children), but often because they are so rare, wasteful, or pointless that only the wealthy can afford them. They include clothing that is too delicate, bulky, constricting, or stain-prone to work in, objects too fragile for casual use or made from unobtainable materials, functionless objects made with prodigious labor, decorations that consume energy, and pale skin in lands where the plebeians work in the fields and suntans in lands where they work indoors. The logic is: You can't see all my wealth and earning power (my bank account, my lands, all my allies and flunkys), but you can see my gold bathroom fixtures. No one could afford them without wealth to spare, therefore you know I am wealthy.

Conspicuous consumption is counterintuitive because squandering wealth can only reduce it, bringing the squanderer down to the level of his or her rivals. But it works when other people's esteem is useful enough to pay for and when not *all* the wealth or earning power is sacrificed. If I have a hundred dollars and you have forty, I can give away fifty, but you can't; I will impress others and *still* be richer than you. The principle has been confirmed from an unlikely source, evolutionary biology. Biologists since Darwin had been puzzled by displays like the peacock's tail, which impresses the peahen but consumes nutrients, hinders movement, and attracts predators. The biologist Amotz Zahavi proposed that the displays evolved *because* they were handicaps. Only the healthiest animals could afford them, and females choose the healthiest birds to

mate with. Theoretical biologists were initially skeptical, but one of them, Alan Grafen, later proved that the theory was sound.

Conspicuous consumption works when only the richest can afford luxuries. When the class structure loosens, or sumptuous goods (or good imitations) become widely available, the upper middle class can emulate the upper class, the middle class can emulate the upper middle class, and so on down the ladder. The upper class cannot very well stand by as they begin to resemble the hoi polloi; they must adopt a new look. But then the look is emulated once again by the upper middle class and begins to trickle down again, prompting the upper class to leap to yet a different look, and so on. The result is fashion. The chaotic cycles of style, in which the chic look of one decade becomes dowdy or slutty, nerdy or foppish in the next, has been explained as a conspiracy of clothing makers, an expression of nationalism, a reflection of the economy, and much else. But Quentin Bell, in his classic analysis of fashion, *On Human Finery*, showed that only one explanation works: people follow the rule, "Try to look like the people above you; if you're at the top, try to look different from the people below you."

Once again animals discovered the trick first. The other dandies of the animal kingdom, butterflies, did not evolve their colors to impress the females. Some species evolved to be poisonous or distasteful, and warned their predators with gaudy colors. Other poisonous kinds copied the colors, taking advantage of the fear already sown. But then some nowpoisonous butterflies copied the colors, too, enjoying the protection while avoiding the expense of making themselves distasteful. When the mimics become too plentiful, the colors no longer conveyed information and no longer deterred the predators. The distasteful butterflies evolved new colors, which were then mimicked by the palatable ones, and so on.

Wealth is not the only asset that people flaunt and covet. In a complicated society, people compete in many leagues, not all of them dominated by plutocrats. Bell added a fourth canon to Veblen's list: conspicuous outrage. Most of us depend on the approval of others. We need the favor of bosses, teachers, parents, clients, customers, or prospective in-laws, and that requires a certain measure of respect and unobtrusiveness. Aggressive nonconformity is an advertisement that one is so confident in one's station or abilities that one can jeopardize the good will of others without ending up ostracized and destitute. It says, "I'm so talented, wealthy, popular, or well-connected that I can afford to offend you." The nineteenth century had the baroness George Sand smoking a cigar in trousers and Oscar

Wilde in knee breeches with long hair and a sunflower. In the last half of the twentieth century conspicuous outrage has become the convention, and we have been treated to a tedious parade of rebels, outlaws, wild ones, bohemians, freaks, punks, shock jocks, gender-benders, mau-maus, bad boys, gangstas, sex divas, bitch goddesses, vamps, tramps, and material girls. Hipness has replaced classiness as the motor of fashion, but the status psychology is the same. Trend-setters are members of upper classes who adopt the styles of lower classes to differentiate themselves from middle classes, who wouldn't be caught dead in lower-class styles because they're the ones in danger of being mistaken for them. The style trickles downward, sending the hip off in search of a new form of outrage. As the media and the merchandisers learn to market each new wave more efficiently, the avant-garde merry-go-round goes faster and more furiously. A regular feature of urban newspapers is the favorable notice of an "alternative" band followed by haughty letters advising that they were good when few had heard of them but that they have now sold out. Tom Wolfe's mordant social commentaries (*The Painted Word, From Bauhaus to Our House, Radical Chic*) document how a thirst for status in the form of hipness drives the worlds of art, architecture, and the politics of the cultural elite.

FRIENDS AND ACQUAINTANCES

People bestow favors on one another even when they are unrelated and have no sexual interest. It is easy to understand why even the most selfish organism might want to do so. If favors are traded, both parties profit as long as the value of what they get is greater to them than the value of what they give up. A clear example is a commodity whose benefit shows diminishing returns. If I have two pounds of meat and no fruit, and you have two pounds of fruit and no meat, the second pound of meat is worth less to me than the first (since there's only so much meat I can eat at a sitting), and you feel the same way about your second pound of fruit. We're both better off if we exchange a pound for a pound. Economists call the benefit a gain in trade.

When traders exchange goods simultaneously, cooperation is easy. If the other guy is renegeing, you hang on to your meat or grab it back. Most favors, however, cannot be retracted, such as sharing information, saving a drowning person, or helping in a fight. Also, most favors cannot change hands at

the same time. Needs may change; if I help you now in return for protection of my unborn child, I cannot collect until the child is born. And surpluses often are staggered; if you and I have just felled antelopes, there's no point in trading identical carcasses. Only if you felled one today and I fell one in a month does it make sense to trade. Money is one solution, but it is a recent invention and could not have figured in our evolution.

As we saw in Chapter 6, the problem with *delayed* exchanges, or reciprocation, is that it's possible to cheat, to accept a favor now and not return it later. Obviously everyone would be better off if no one cheated. But as long as the other guy *might* cheat (which is inevitable when individuals can vary), I may be discouraged from extending him a favor that in the long run would help us both. The problem has been compressed into a parable called the Prisoner's Dilemma. Partners in crime are held in separate cells, and the prosecutor offers each one a deal. If you rat on your partner and he stays mum, you go free and he gets ten years. If you both stay mum, you both get six months. If you both rat, you both get five years. The partners cannot communicate, and neither knows what the other will do. Each one thinks: If my partner rats and I stay mum, I'll do ten years; if he rats and I rat, too, I'll do five years. If he stays mum and I stay mum, I'll do six months; if he stays mum and I rat, I'll go free. Regardless of what he does, then, I'm better off betraying him. Each is compelled to turn in his partner, and they both serve five years—far worse than if each had trusted the other. But neither could take the chance because of the punishment he would incur if the other didn't. Social psychologists, mathematicians, economists, moral philosophers, and nuclear strategists have fretted over the paradox for decades. There is no solution.

Real life, however, is not a Prisoner's Dilemma in one respect. The mythical prisoners are placed in their dilemma once. Real people face each other in dilemmas of cooperation again and again, and can remember past treacheries or good turns and play accordingly. They can feel sympathetic and extend good will, feel aggrieved and seek revenge, feel grateful and return a favor, or feel remorseful and make amends. Recall that Trivers proposed that the emotions making up the moral sense could evolve when parties interacted repeatedly and could reward cooperation now with cooperation later and punish defection now with defection later. Robert Axelrod and William Hamilton confirmed the conjecture in a round-robin computer tournament that pitted different strategies for playing a repeated Prisoner's Dilemma game against each other. They

stripped the dilemma to its essentials and awarded points to a strategy for the equivalent of minimizing jail time. A simple strategy called tit-for-tat—cooperate on the first move, and then do what your partner did on the move before—beat sixty-two other strategies. Then they ran an artificial life simulation in which each strategy "reproduced" in proportion to its winnings and a new round-robin took place among the copies of the strategies. They repeated the process for many generations and found that the Tit for Tat strategy took over the population. Cooperativeness can evolve when the parties interact repeatedly, remember each other's behavior, and reciprocate it.

As we saw in Chapters 5 and 6, people are good at detecting cheaters and are fitted with moralistic emotions that prompt them to punish the cheaters and reward the cooperators. Does that mean that tit-for-tat underlies the widespread cooperation we find in the human species? It certainly underlies much of the cooperation we find in our society. Cash-register tapes, punch clocks, train tickets, receipts, accounting ledgers, and the other accoutrements of transactions that do not rely on the "honor system" are mechanical cheater-detectors. The cheaters, such as thieving employees, are sometimes charged with crimes, but more often they are simply cut off from further reciprocation, that is, fired. Similarly, the businesses that cheat their customers soon lose them. Footloose job applicants, fly-by-night businesses, and strangers calling with "investment opportunities" are often discriminated against because they look like they are playing a one-shot rather than an iterated game of cooperation, and so are immune to tit-for-tat. Even moderately good friends privately remember the most recent Christmas gifts and dinner-party invitations and calculate the proper way to reciprocate.

Does all this accounting come from our alienation and bourgeois values in a capitalist society? One of the fondest beliefs of many intellectuals is that there are cultures out there where everyone shares freely. Marx and Engels thought that preliterate peoples represented a first stage in the evolution of civilization called primitive communism, whose maxim was "From each according to his abilities, to each according to his needs." Indeed, people in foraging societies do share food and risk. But in many of them, people interact mainly with their kin, so in the biologist's sense they are sharing with extensions of themselves. Many cultures also have an *ideal* of sharing, but that means little. Of course I will proclaim how great it is for *you* to share; the question is, will *J* share when my turn comes?

Foraging peoples, to be sure, really do share with nonrelatives, but not out of indiscriminate largesse or a commitment to socialist principles. The data from anthropology show that the sharing is driven by cost-benefit analyses and a careful mental ledger for reciprocation. People share when it would be suicidal not to. In general, species are driven to share when the *variance* of success in gathering food is high. Say in some weeks I am lucky and have more food than I can eat, but in other weeks I am unlucky and in danger of starving. How can I store extra food in the fat weeks and draw on it in the lean weeks? Refrigeration is not an option. I could gorge on it now and store it as blubber, but that works only up to a point; I can't eat enough in a day to avoid hunger for a month. But I *can* store it in the bodies and minds of *other* people, in the form of a memory of my generosity they feel obliged to repay when fortunes reverse. When the prospects are risky, it pays to pool the risks.

The theory has been confirmed in nonhuman species, such as vampire bats, and it has also been confirmed in humans in two elegant studies that control for differences among cultures by contrasting the forms of sharing *within* a culture. The Ache of Paraguay hunt game and gather plant foods. Hunting is largely a matter of luck: on any given day an Ache hunter has a forty percent chance of coming home empty-handed. Gathering is largely a matter of effort: the longer you work, the more you bring home, and an empty-handed gatherer is probably lazy rather than unlucky. As predicted, the Ache share plant foods only within the nuclear family but share meat throughout the band.

The !Kung San of the Kalahari Desert are perhaps the closest thing the world has to primitive communists. Sharing is holy; boasting and hoarding are contemptible. They hunt and gather in a harsh, drought-prone ecosystem, and trade food and access to waterholes. The //Gana San, a neighboring branch of the same people, have taken to cultivating melons, which store water, and to herding goats. They do not yo-yo between good times and bad as much as their cousins, and unlike them, they hoard food and have developed inequalities in wealth and status. In both the Ache and the San, high-variance foods are shared, low-variance foods are hoarded.

These people do not pull out calculators and compute the variances. What goes through their minds when they decide to share? Cosmides and Tooby note that the psychology is hardly exotic; it matches our own sense of fairness and compassion. Consider what makes people more or less willing to help the homeless. Those who urge that we all share with

the homeless emphasize the random, variance-driven dimension to homelessness. Homeless people are worthy of aid because they are down on their luck. They are the unfortunate victims of circumstances like unemployment, discrimination, or mental illness. Advocates of the homeless urge us to think, "There but for fortune go I." Those who oppose sharing, on the other hand, emphasize the predictability of rewards in our society to anyone willing to put in the work. Homeless people are unworthy of aid because they are able-bodied but lazy, or brought it on themselves by choosing to drink or take drugs. Defenders of the homeless reply that drug use is itself an illness that could happen to anyone.

Even at their most munificent, foraging people do not act out of hearts filled with loving kindness. They enforce the sharing ethic with obsessively detailed memories of who has helped, a clear expectation of payback, and snide gossip about those who don't pitch in. And all this still does not expunge selfish feelings. The anthropologist Melvin Konner, who lived with the !Kung San for years and has written respectfully about their ways, tells his readers:

Selfishness, arrogance, avarice, cupidity, fury, covetousness, all these forms of gluttony are held in check in their traditional situation in the same way simple alimentary gluttony is: Namely, it doesn't happen because the situation does not allow it. Nor, as some suppose, because the people or their culture are somehow better. I will never forget the time a !Kung man—the father of a family, about forty years of age, well respected in the community, a good and substantial man in every way—asked me to hold on to a leg of antelope he had killed. He had given away most of it, as one had to. But he saw a chance to hide some of it, for later, for himself and his own family. Ordinarily, of course, there would be no place in the entire Kalahari to hide it; it would either be unsafe from scavengers or unsafe from predatory distant relatives. But the presence of foreigners presented an interface with another world, and he wanted to slip the meat, temporarily, through a chink in that interface, into the only conceivable hiding place.

When it comes to friendship, reciprocal altruism does not ring true. It would be in questionable taste for a dinner guest to pull out his wallet

and offer to pay the hosts for his dinner. Inviting the hosts back the very next night would not be much better. Tit-for-tat does not cement a friendship; it strains it. Nothing can be more awkward for good friends than a business transaction between them, like the sale of a car. The same is true for one's best friend in life, a spouse. The couples who keep close track of what each has done for the other are the couples who are the least happy.

Companionate love, the emotion behind close friendship and the enduring bond of marriage (the love that is neither romantic nor sexual), has a psychology of its own. Friends or spouses feel as if they are in each other's debt, but the debts are not measured and the obligation to repay is not onerous but deeply satisfying. People feel a spontaneous pleasure in helping a friend or a spouse, without anticipating repayment or regretting the favor if repayment never comes. Of course, the favors may be tabulated somewhere in the mind, and if the ledger has become too lopsided, a person might call in the debt or cut off future credit, that is, end the friendship. But the line of credit is long and the terms of repayment forgiving. Companionate love, then, does not literally contradict the theory of reciprocal altruism, but it does embody an elastic version in which the emotional guarantors—liking, sympathy, gratitude, and trust—are stretched to the limit.

The facts of companionate love are clear enough, but why did it evolve? Tooby and Cosmides have tried to reverse-engineer the psychology of friendship by calling attention to an aspect of the logic of exchange they call the Banker's Paradox. Many frustrated borrowers have learned that a bank will lend you exactly as much money as you can prove you don't need. As Robert Frost put it, "A bank is a place where they lend you an umbrella in fair weather and ask for it back when it begins to rain." The banks say they have only so much money to invest and every loan is a gamble. Their portfolio has to return a profit or they would go out of business, so they measure credit risks and weed out the worst.

The same cruel logic applies to altruism among our ancestors. A person mulling over whether to extend a large favor is like a bank. He must worry not only about cheaters (is the beneficiary willing to repay?) but about bad credit risks (is the beneficiary *able* to repay?). If the recipient dies, is disabled, becomes a pariah, or leaves the group, the favor would have been wasted. Unfortunately, it is the bad credit risks—the sick, starving, injured, and ostracized—who most *need* favors. Anyone can

suffer a reversal of fortune, especially in the harsh life of a forager. Once abandoned, a stricken forager is not long for that world. What kinds of thoughts and feelings might evolve as a kind of insurance in which other people would extend "credit" to you even if misfortune were to make you a risk?

One strategy is to make yourself irreplaceable. By cultivating expertise that no one in the group can duplicate, like toolmaking, wayfinding, or conflict resolution, you make yourself costly to abandon in times of need: everyone depends upon you too much to risk letting you die. People today do spend a lot of their social lives publicizing their unique and valuable talents or looking for a clique in which their talents would be unique and valuable. The quest for status is in part a motive for making oneself irreplaceable.

Another is to associate with people who benefit from the things that benefit you. Merely by going about your life and pursuing your own interests, you can advance someone else's interests as a side effect. Marriage is the clearest example: the husband and wife share an interest in their children's welfare. Another was pointed out by Mao Tse-tung in his little red book: "The enemy of my enemy is my friend." A third is to possess skills that benefit others at the same time that they benefit you, like being good at finding your way home. Other examples are living with a person who likes the room at the same temperature or who likes the same music. In all the examples, one delivers a benefit to someone without being altruistic in the biologist's sense of incurring a cost and thereby needing a repayment to make the act worthwhile. The challenge of altruism has attracted so much attention that a more direct form of helping in nature has often been downplayed: symbiosis, in which two organisms, such as the algae and fungi making up lichen, associate because the side effects of each one's lifestyle fortuitously benefit the other one. Symbionts give benefits and take them, but neither pays a cost. Roommates with the same taste in music are a kind of symbiotic pair, and each can value the other without an exchange of favors.

Once you have made yourself valuable to someone, the person becomes valuable to you. You value him or her because if you were ever in trouble, they would have a stake—albeit a selfish stake—in getting you out. But now that you value the person, they should value you even more. Not only are you valuable because of your talents or habits, but you are valuable because of your stake in rescuing him or her from hard times. The more you value the person, the more the person values you,

and so on. This runaway process is what we call friendship. If you ask people why they are friends, they are likely to say, "We like the same things, and we know we'll always be there for each other."

Friendship, like other kinds of altruism, is vulnerable to cheaters, and we have a special name for them: fair-weather friends. These sham friends reap the benefits of associating with a valuable person and mimic signs of warmth in an effort to become valued themselves. But when a little rain falls, they are nowhere in sight. People have an emotional response that seems designed to weed out fair-weather friends. When we are neediest, an extended hand is deeply affecting. We are moved, never forget the generosity, and feel compelled to tell the friend we will never forget it. Hard times show you who your real friends are. That is because the point of friendship, in evolutionary terms, is to save you in hard times when it's not worth anyone else's trouble.

Tooby and Cosmides go on to speculate that the design of our friendship emotions may explain the alienation and loneliness that so many people feel in modern society. Explicit exchanges and turn-taking reciprocity are the kinds of altruism we fall back on when friendship is absent and trust is low. But in modern market economies we trade favors with strangers at unprecedented rates. It may create the perception that we are not deeply engaged with our fellows and are vulnerable to desertion in difficult times. And ironically, the comfortable environment that makes us physically more secure may make us emotionally less secure, because it minimizes the crises that tell us who our real friends are.

ALLIES AND ENEMIES

No account of human relationships could be complete without a discussion of war. War is not universal, but people in all cultures feel that they are members of a group (a band, tribe, clan, or nation) and feel animosity toward other groups. And warfare itself is a major fact of life for foraging tribes. Many intellectuals believe that primitive warfare is rare, mild, and ritualized, or at least was so until the noble savages were contaminated by contact with Westerners. But this is romantic nonsense. War has always been hell.

Yanomamö villages raid one another endlessly. Seventy percent of all adults over forty have lost a family member to violence. Thirty percent of

the men are killed by other men. Forty-four percent of the men have killed someone. The Yanomamö call themselves the Fierce People, but other pristine tribes give similar numbers. The archeologist Lawrence Keeley has documented that New Guineans, Australian aborigines, Pacific Islanders, and Native Americans have been wracked by warfare, especially in the centuries before the Pax Britannica ended this nuisance to the colonial administrators in much of the world. In primitive warfare, mobilization was more complete, battles were more frequent, casualties higher, prisoners fewer, and weapons more damaging. War is, to put it mildly, a major selection pressure, and since it appears to have been a recurring event in our evolutionary history, it must have shaped parts of the human psyche.

Why would anyone be so stupid as to start a war? Tribal people can fight over anything of value, and the causes of tribal wars are as difficult to disentangle as the causes of World War I. But one motive that is surprising to Westerners appears over and over. In foraging societies, men go to war to get or keep women—not necessarily as a conscious goal of the warriors (though often it is exactly that), but as the ultimate payoff that allowed a willingness to fight to evolve. Access to women is the limiting factor on males' reproductive success. Having two wives can double a man's children, having three wives can triple it, and so on. For a man who is not at death's door, no other resource has as much impact on evolutionary fitness. The most common spoils of tribal warfare are women. Raiders kill the men, abduct the nubile women, gang-rape them, and allocate them as wives. Chagnon discovered that Yanomamö men who had killed an enemy had three times as many wives and three times as many children as those who had not. Most young men who had killed were married; most young men who had never killed were not. The difference is not an accident of other differences between the killers and the non-killers, such as size, strength, or number of kin. Killers are held in esteem in Yanomamö villages; they attract and are ceded more wives.

The Yanomamö sometimes plan raids just to abduct women. More frequently, they plan them to avenge a past killing or abduction, but they always try to abduct women, too. Blood feuds, in which relatives avenge a death with a death, either of the killer or of his relatives, are the major impetus to extended violence everywhere; the motive that drives them has an obvious deterrent function, as we saw in Chapter 6. Blood feuds can extend for decades or longer because each side counts the score differently, so at any time each remembers injustices that must be redressed. (Imagine your feelings toward a neighboring people that has murdered

your husband, your brothers, and your sons, or has raped and abducted your wife, your daughters, and your sisters.) But the feuders do not stop at an eye for an eye. If they see an opportunity to get rid of a headache once and for all by massacring their opponents, they may do so, with the women as an extra incentive. The desire for women not only helps to fuel blood feuds; it also helps to spark them in the first place. Usually the first killing was over a woman: a man seduces or abducts someone's wife, or reneges on a deal to trade a daughter.

Modern people have trouble believing that preliterate tribes go to war over women. One anthropologist wrote to Chagnon, "Women? Fighting over women? Gold and diamonds I can understand, but women? Never." The reaction, of course, is biologically topsy-turvy. Other anthropologists argued that the Yanomamö suffered from a protein shortage and were fighting over game. But their protein intake, when measured, turned out to be more than adequate. Across the world the best-fed foraging peoples are the *most* warlike. When Chagnon mentioned the meat-shortage hypothesis to his Yanomamö informants, they laughed incredulously and said, "Even though we like meat, we like women a whole lot more." Chagnon points out that they are not so different from us. "Some Saturday night just visit a hard-hat bar where fights are frequent. What are the fights usually about? Are they about the amount of meat in someone's hamburger? Or study the words of a dozen country-and-western songs. Do any of them say, 'Don't take your cow to town'?"

The similarities run deeper. Warfare among Western peoples is different from primitive warfare in many ways, but it is similar in at least one way: the invaders rape or abduct women. It was codified in the Bible:

And they warred against the Midianites, as the LORD commanded Moses; and they slew all the males. . . . And the children of Israel took all the women of Midian captives, and their little ones, and took the spoil of all their cattle, and all their flocks, and all their goods. . . . And Moses said unto them, Have ye saved all the women alive? . . . Now therefore kill every male among the little ones, and kill every woman that hath known man by lying with him. But all the women children, that have not known a man by lying with him, keep alive for yourselves. (Numbers 31)

When thou comest nigh unto a city to fight against it, then proclaim peace unto it. . . . And if it will make no peace with thee, but will make war against thee, then thou shalt besiege it: And when the LORD thy God hath delivered it into thine hands, thou shalt smite every male thereof

with the edge of the sword: But the women, and the little ones, and the cattle, and all that is in the city, even all the spoil thereof, shalt thou take unto thyself. (Deuteronomy 20)

When thou goest forth to war against thine enemies, and the LORD thy God hath delivered them into thine hands, and thou hast taken them captive, And seest among the captives a beautiful woman, and hast a desire unto her, that thou wouldest have her to thy wife; Then thou shalt bring her home to thine house; and she shall shave her head, and pare her nails; And she shall put the raiment of her captivity from off her, and shall remain in thine house, and bewail her father and her mother a full month; and after that thou shalt go in unto her, and be her husband, and she shall be thy wife. (Deuteronomy 21)

According to the *Iliad*, the Trojan War began with the abduction of Helen of Troy. During the First Crusade, Christian soldiers raped their way across Europe to Constantinople. Shakespeare has Henry V threatening a French village during the Hundred Years War that if they do not surrender, it will be their fault that their "pure maidens fall into the hand of hot and forcing violation":

If not, why, in a moment look to see
The blind and bloody soldier with foul hand
Defile the locks of your shrill-shrieking daughters;
Your fathers taken by the silver beards,
And their most reverend heads dash'd to the walls,
Your naked infants spitted upon pikes,
Whiles the mad mothers with their howls confused
Do break the clouds, as did the wives of Jewry
At Herod's bloody-hunting slaughtermen.

The feminist writer Susan Brownmiller has documented that rape was systematically practiced by the English in the Scottish Highlands, the Germans invading Belgium in World War I and eastern Europe in World War II, the Japanese in China, the Pakistanis in Bangladesh, the Cossacks during the pogroms, the Turks persecuting the Armenians, the Ku Klux Klan in the American South, and, to a lesser extent, Russian soldiers marching toward Berlin and American soldiers in Vietnam. Recently the Serbs in Bosnia and the Hutus in Rwanda have added themselves to this list. Prostitution, which in wartime is often hard to

distinguish from rape, is a ubiquitous perquisite of soldiers. Leaders may sometimes use rape as a terror tactic to attain other ends, as Henry V obviously did, but the tactic is effective precisely because the soldiers are so eager to implement it, as Henry took pains to remind the Frenchmen. In fact it often backfires by giving the defenders an incalculable incentive to fight on, and probably for that reason, more than out of compassion for enemy women, modern armies have outlawed rape. Even when rape is not a prominent part of our warfare, we invest our war leaders with enormous prestige, just as the Yanomamø do, and by now you know the effects of prestige on a man's sexual attractiveness and, until recently, his reproductive success.

War, or aggression by a coalition of individuals, is rare in the animal kingdom. You would think that the second-, third-, and fourth-strongest elephant seals would gang up, kill the strongest male, and divide his harem among them, but they never do. Aside from the social insects, whose unusual genetic system makes them a special case, only humans, chimpanzees, dolphins, and perhaps bonobos join up in groups of four or more to attack other males. These are some of the largest-brained species, hinting that war may require sophisticated mental machinery. Tooby and Cosmides have worked out the adaptive logic of coalitional aggression and the cognitive mechanisms necessary to support it. (That does not, of course, mean that they think war is unavoidable or "natural" in the sense of "good.")

People often are conscripted into armies, but sometimes they enlist with gusto. Jingoism is alarmingly easy to evoke, even without a scarce resource to fight over. In numerous experiments by Henri Tajfel and other social psychologists, people are divided into two groups, actually at random but ostensibly by some trivial criterion such as whether they underestimate or overestimate the number of dots on a screen or whether they prefer the paintings of Klee or Kandinsky. The people in each group instantly dislike and think worse of the people in the other group, and act to withhold rewards from them even if doing so is costly to their own group. This instant ethnocentrism can be evoked even if the experimenter drops the charade with the dots or paintings and divides people into groups by flipping a coin before their eyes! The behavioral

consequences are by no means minor. In a classic experiment, the social psychologist Muzafer Sherif carefully selected a group of well-adjusted, middle-class American boys for a summer camp, and randomly divided them into two groups which then competed in sports and skits. Within days the groups were brutalizing and raiding each other with sticks, bats, and rocks in socks, forcing the experimenters to intervene for the boys' safety.

The enigma of war is why people volunteer for an activity that has an excellent chance of getting them killed. How could a desire to play Russian roulette have evolved? Tooby and Cosmides explain it by the fact that natural selection favors traits that increase fitness *on average*. Every gene contributing to a trait is embodied in many individuals in many generations, so if one individual with the gene dies childless, the success of many others with the gene can make up for it. Imagine a game of Russian roulette where if you don't get killed you have one more offspring. A gene for joining in the game could be selected, because five-sixths of the time it would leave an extra copy in the gene pool and one-sixth of the time it would leave none. On average, that yields .83 more copies than staying out of the game. Joining a coalition of five other men that is certain to capture five women but suffer one fatality is in effect the same choice. The key idea is that the coalition acting together can gain a benefit that its members acting alone cannot, and that spoils are distributed according to the risks undertaken. (There are several complications, but they do not change the point.)

In fact, if the spoils are certain and divided up fairly, the level of danger doesn't matter. Say your coalition has eleven members and can ambush an enemy coalition of five, taking their women. If one member of your coalition is likely to be killed, you have a ten-in-eleven chance of surviving, which would entitle you to a one-in-two chance (five captive women, ten men) of gaining a wife, an expected gain of .45 wives (averaged over many situations with these payoffs). If two members will be killed, you have a smaller chance of surviving (nine in eleven), but if you do survive you have a larger chance of gaining a wife, since your dead allies won't be taking theirs. The average gain ($9/11 \times 5/9$) is the same, .45 wives. Even if *six* members are likely to be killed, so that your survival odds fall to less than even (five in eleven), the spoils are divided fewer ways (five women among five victors), so if you survive you are guaranteed a wife, for an expected gain, once again, of .45 wives.

Tooby and Cosmides' calculations assume that a man's children can do just fine when he is dead, so the loss of fitness with death is zero, not negative. Of course that is not true, but they point out that if the group is relatively prosperous the fatherless children's survival chances may not diminish too much and it still could pay men to raid. They predict that men should be more willing to fight when their group is secure in food than when it is hungry, contrary to the protein-shortage hypothesis. The data bear them out. Another implication is that females should never have an interest in starting a war (even if they had weapons or allies that made up for their smaller size). The reason that females never evolved an appetite to band together and raid neighboring villages for husbands is that a woman's reproductive success is rarely limited by the number of available males, so any risk to her life while pursuing additional mates is a sheer loss in expected fitness. (Foraging women do, however, encourage men to fight in defense of the group and to avenge slain family members.) The theory also explains why in modern warfare most people are unwilling to send women into combat and feel morally outraged when women are casualties, even though no ethical argument makes a woman's life more precious than a man's. It is hard to shake the intuition that war is a game that benefits men (which was true for most of our evolutionary history), so they should bear the risks.

The theory also predicts that men should be willing to fight collectively only if they are confident of victory and none of them knows in advance who will be injured or killed. If defeat is likely, it's pointless to fight on. And if you bear more than your share of the risk—say, if your platoonmates are exposing you to danger by looking out for their own hides—it's also pointless to fight on. These two principles shape the psychology of war.

Among foragers, warring bands are usually factions of the same people and have the same kinds of weaponry, so the predictor of victory in our evolutionary past would have been sheer numbers. The side with more warriors was invincible, and the odds of victory could be estimated from the manpower on each side. The Yanomamø are obsessed with the size of their villages for just that reason, and they often form alliances or rethink secessions because they know that smaller villages are helpless in wars. Even in modern societies, a mob of people on your side is emboldening and a mob on the other side terrifying. Mustering a crowd is a common tactic for whipping up patriotism, and a mass demonstration can incite panic even in a militarily secure ruler. A major principle

of battlefield strategy is to surround an enemy unit, making defeat look certain and causing panic and rout.

Just as important is an equitable distribution of risk. A war party faces the problem of altruism par excellence. Every member has an incentive to cheat by keeping himself out of harm's way and exposing the others to greater risk. Just as benevolent cooperation cannot evolve unless the favor-granter detects and punishes cheaters, aggressive cooperation cannot evolve unless the fighters detect and punish cowards or shirkers. Bravery and discipline are the obsessions of fighting men. They affect everything from a soldier's sense of whom he wants in his foxhole to the command structure that coerces soldiers into assuming risk equitably and that rewards bravery and punishes desertion. War is rare in the animal kingdom because animals, like humans, ought to be cowards unless they can enforce a multiparty contract to share the risks. Unlike ancestral humans, they did not have the cognitive machinery from which an enforcement calculator could easily evolve.

Here is another peculiarity of the logic and psychology of war. A man should agree to stay in a coalition for as long as he does not *know* that he is about to die. He may know the odds, but he cannot know whether the spinner of death is slowing down at him. But at *some* point he may see it coming. He may glimpse an archer who has him in his sights, or detect an impending ambush, or notice that he has been sent on a suicide mission. At that point everything changes, and the only rational move is to desert. Of course, if the uncertainty collapses only seconds before death, it's too late. The farther in advance a fighter can predict that he is about to become an unknown soldier, the more easily he can desert, and the more likely the coalition is to unravel. In a coalition of animals attacking another coalition or an individual, an attacker has some warning if he is being picked out for a counterattack, and can flee before they give chase. For that reason a coalition of animals would be especially prone to unraveling. But humans have invented weapons, from spears and arrows to bullets and bombs, that make fate unknowable until the last second. Behind this veil of ignorance, men can be motivated to fight to the last.

Decades before Tooby and Cosmides spelled out this logic, the psychologist Anatol Rapoport illustrated it with a paradox from World War II. (He believed the scenario was true but was unable to verify it.) At a bomber base in the Pacific, a flier had only a twenty-five percent chance of surviving his quota of missions. Someone calculated that if the fliers carried twice as many bombs, a mission could be carried out with half as

many flights. But the only way to increase the payload was to reduce the fuel, which meant that the planes would have to fly on one-way missions. If the fliers would be willing to draw lots and take a one-in-two chance of flying off to a certain death instead of hanging on to their three-in-four chance of flying off to an unpredictable death, they would *double* their chance of survival: only half of them would die instead of three-quarters. Needless to say, it was never implemented. Few of us would accept such an offer, though it is completely fair and would save many lives, including, possibly, our own. The paradox is an intriguing demonstration that our mind is equipped to volunteer for a risk of death in a coalition but only if we do not know when death will come.

HUMANITY

So should we all just take poison now and be done with it? Some people think that evolutionary psychology claims to have discovered that human nature is selfish and wicked. But they are flattering the researchers and anyone who would claim to have discovered the opposite. No one needs a scientist to measure whether humans are prone to knavery. The question has been answered in the history books, the newspapers, the ethnographic record, and the letters to Ann Landers. But people treat it like an open question, as if someday science might discover that it's all a bad dream and we will wake up to find that it is human nature to love one another. The task of evolutionary psychology is not to weigh in on human nature, a task better left to others. It is to add the satisfying kind of insight that only science can provide: to connect what we know about human nature with the rest of our knowledge of how the world works, and to explain the largest number of facts with the smallest number of assumptions. Already a large part of our social psychology, well documented in the lab and the field, can be shown to fall out of a few assumptions about kin selection, parental investment, reciprocal altruism, and the computational theory of mind.

So does human nature doom us to a nightmare of exploitation by ruthless fitness-maximizers? Again, it is silly to look to science for the answer. Everyone knows that people are capable of monumental kindness and sacrifice. The mind has many components, and accommodates not only ugly motives but love, friendship, cooperation, a sense of fair-

ness, and an ability to predict the consequences of our actions. The different parts of the mind struggle to engage or disengage the clutch pedal of behavior, so bad thoughts do not always cause bad deeds. Jimmy Carter, in his famous *Playboy* interview, said, "I have looked on a lot of women with lust. I've committed adultery in my heart many times." But the prying American press has found no evidence that he has committed it in real life even once.

And on the larger stage, history has seen terrible blights disappear permanently, sometimes only after years of bloodshed, sometimes as if in a puff of smoke. Slavery, harem-holding despots, colonial conquest, blood feuds, women as property, institutionalized racism and anti-Semitism, child labor, apartheid, fascism, Stalinism, Leninism, and war have vanished from expanses of the world that had suffered them for decades, centuries, or millennia. The homicide rates in the most vicious American urban jungles are twenty times lower than in many foraging societies. Modern Britons are twenty times less likely to be murdered than their medieval ancestors.

If the brain has not changed over the centuries, how can the human condition have improved? Part of the answer, I think, is that literacy, knowledge, and the exchange of ideas have undermined some kinds of exploitation. It's not that people have a well of goodness that moral exhortations can tap. It's that information can be framed in a way that makes exploiters look like hypocrites or fools. One of our baser instincts—claiming authority on a pretext of beneficence and competence—can be cunningly turned on the others. When everyone sees graphic representations of suffering, it is no longer possible to claim that no harm is being done. When a victim gives a first-person account in words the victimizer might use, it's harder to maintain that the victims are a lesser kind of being. When a speaker is shown to be echoing the words of his enemy or of a past speaker whose policies led to disaster, his authority can crumble. When peaceable neighbors are described, it's harder to insist that war is inevitable. When Martin Luther King said, "I have a dream that one day this nation will rise up and live out the true meaning of its creed: 'We hold these truths to be self-evident, that all men are created equal,'" he made it impossible for segregationists to maintain they were patriots without looking like charlatans.

And as I mentioned at the outset, though conflict is a human universal, so are efforts to reduce it. The human mind occasionally catches a

glimmering of the brute economic fact that often adversaries can both come out ahead by dividing up the surplus created by their laying down their arms. Even some of the Yanomamö see the futility of their ways and long for a means to break the cycle of vengeance. People throughout history have invented ingenious technologies that turn one part of the mind against another and eke increments of civility from a human nature that was not selected for niceness: rhetoric, exposés, mediation, face-saving measures, contracts, deterrence, equal opportunity, mediation, courts, enforceable laws, monogamy, limits on economic inequality, abjuring vengeance, and many others. Utopian theoreticians ought to be humble in the face of this practical wisdom. It is likely to remain more effective than "cultural" proposals to make over childrearing, language, and the media, and "biological" proposals to scan the brains and genes of gang members for aggression markers and to hand out antiviolence pills in the ghettos.

Tenzin Gyatso, the Dalai Lama of Tibet, was identified at the age of two as the fourteenth reincarnation of the Buddha of Compassion, Holy Lord, Gentle Glory, Eloquent, Compassionate, Learned Defender of the Faith, Ocean of Wisdom. He was taken to Lhasa and brought up by dotting monks, who tutored him in philosophy, medicine, and metaphysics. In 1950 he became the spiritual and secular leader in exile of the Tibetan people. Despite not having a power base, he is recognized as a world statesman on the sheer force of his moral authority, and in 1989 was awarded the Nobel Peace Prize. No human being could be more predisposed by his upbringing and by the role he has been thrust into to have pure and noble thoughts.

In 1993 an interviewer for the *New York Times* asked him about himself. He said that as a boy he loved war toys, especially his air rifle. As an adult, he relaxes by looking at battlefield photographs and had just ordered a thirty-volume Time-Life illustrated history of World War II. Like guys everywhere, he enjoys studying pictures of military hardware, like tanks, airplanes, warships, U-boats, submarines, and especially aircraft carriers. He has erotic dreams and finds himself attracted to beautiful women, often having to remind himself, "I'm a monk!" None of this has stood in the way of his being one of history's great pacifists. And despite the oppression of his people, he remains an optimist and predicts that the twenty-first century will be more peaceful than the twentieth. Why? asked the interviewer. "Because I believe," he said, "that in the 20th century, humanity has learned something from many, many experi-

ences. Some positive, and many negative. What misery, what destruction! The greatest number of human beings were killed in the two world wars of this century. But human nature is such that when we face a tremendous critical situation, the human mind can wake up and find some other alternative. That is a human capacity."

AGENCY PROBLEMS AND RESIDUAL CLAIMS*

EUGENE F. FAMA and MICHAEL C. JENSEN
University of Chicago University of Rochester

I. INTRODUCTION

A. Organizational Survival

SOCIAL and economic activities, such as religion, entertainment, education, research, and the production of other goods and services, are carried on by different types of organizations, for example, corporations, proprietorships, partnerships, mutuals, and nonprofits. Most goods and services can be produced by any form of organization, and there is competition among organizational forms for survival in any activity. Absent fiat, the form of organization that survives in an activity is the one that delivers the product demanded by customers at the lowest price while covering costs. This is the telling dimension on which the economic environment chooses among organizational forms.

An important factor in the survival of organizational forms is control of agency problems. Agency problems arise because contracts are not costlessly written and enforced. Agency costs include the costs of structuring, monitoring, and bonding a set of contracts among agents with conflicting interests, plus the residual loss incurred because the cost of full enforcement of contracts exceeds the benefits.¹ In this paper we explain the

* This paper is a revision of parts of our earlier paper, *The Survival of Organizations* (September 1980). In the course of this work, we have profited from the comments of R. Antle, R. Benne, F. Black, F. Easterbrook, A. Farber, W. Gavett, P. Hirsch, R. Hogarth, C. Holderness, R. Holthausen, C. Horne, J. Jeuck, R. Leftwich, S. McCormick, D. Mayers, P. Pashigian, M. Scholes, C. Smith, G. Stigler, R. Watts, T. Whisler, R. Yeaple, J. Zimmerman, and especially A. Alchian, W. Meckling, and C. Plosser. Financial support for Fama's participation is from the National Science Foundation. Jensen is supported by the Managerial Economics Research Center of the University of Rochester.

¹ This definition of agency costs first appears in Michael C. Jensen & William H. Meckling, *Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure*, 3 *J. Financial Econ.* 305 (1976).

[*Journal of Law & Economics*, vol. XXVI (June 1983)]

© 1983 by The University of Chicago. All rights reserved. 0022-2186/83/2602-0012\$01.50

special features of the residual claims of different organizational forms as efficient approaches to controlling special agency problems. We analyze only private organizations. In related papers we examine other features of the contract structures of different organizational forms that contribute to their survival; in particular, (1) the control of agency problems in the class of organizations characterized by separation of "ownership" and "control," and (2) the effects of special characteristics of residual claims on decision rules for resource allocation.²

B. Residual Claims: General Discussion

The contract structures of organizations limit the risks undertaken by most agents by specifying either fixed payoffs or incentive payoffs tied to specific measures of performance. The residual risk—the risk of the difference between stochastic inflows of resources and promised payments to agents—is borne by those who contract for the rights to net cash flows. We call these agents the residual claimants or residual risk bearers.

The characteristics of residual claims distinguish organizations from one another and help explain the survival of organizational forms in specific activities. We first analyze and contrast the relatively unrestricted residual claims of open corporations with the restricted residual claims of proprietorships, partnerships, and closed corporations. We then turn to the more specialized residual claims of professional partnerships, financial mutuals, and nonprofits.

II. OPEN CORPORATIONS

Most large nonfinancial organizations are open corporations. The common stock residual claims of such organizations are unrestricted in the sense that (1) stockholders are not required to have any other role in the organization, (2) their residual claims are freely alienable, and (3) the residual claims are rights in net cash flows for the life of the organization. Because of the unrestricted nature of the residual claims of open corporations, there is generally almost complete separation and specialization of decision functions and residual risk bearing.

A. Common Stock versus State Contingent Claims

One can imagine claims that are even less restricted than the common stocks of open corporations. There could be "state contingent claims"—

² Eugene F. Fama & Michael C. Jensen, *Separation of Ownership and Control*, in this issue. See also Eugene F. Fama & Michael C. Jensen, *Organizational Forms and Investment Decisions* (Working Paper No. MERC 83-03, Univ. Rochester, Managerial Economics Research Center 1983).

that is, claims of the sort discussed by Arrow and Debreu³ specifying payoffs for each possible future state of the world. Such state contingent claims allow any (hence generally "less restricted") allocation of risk. They are, nonetheless, fixed payoff promises. To specify the total payoffs to be obtained in all future states, one would need to identify all current and future decisions of an organization through state contingent claim contracts. Given the costs and information requirements this implies, it is not surprising that state contingent claims are not the dominant system for allocating risk.

We can also imagine state contingent claims that are true residual claims. The claim would cover a fraction of the organization's net cash flows in a given state rather than a specified payoff in that state. However, this type of claim generates conflicts among the claim holders of different states because alternative decisions shift payoffs across states and benefit some claim holders at the expense of others. Common stock that represents proportionate claims on the payoffs of all future states eliminates these agency problems, but at the sacrifice of some efficiency in the allocation of risk. Common stock and other common forms of residual claims also avoid most of the costs of defining and verifying states of the world.

B. *The Advantages of Common Stock Residual Claims*

1. *Unrestricted Risk Sharing among Residual Claimants.* The common stock of open corporations allows more efficient risk sharing than residual claims that are not separable from decision roles, as, for example, in proprietorships and partnerships where the proprietors and partners are the decision makers and the primary residual claimants. Common stock allows residual risk to be spread across many residual claimants who individually choose the extent to which they bear risk and who can diversify across organizations offering such claims. Other things equal, portfolio theory implies that such unrestricted risk sharing lowers the cost of risk-bearing services.⁴

2. *Specialized Risk Bearing by Residual Claimants.* The activities of large open nonfinancial corporations are typically complicated, involving contracts with many factors of production, for example, different types of labor, raw materials, and managers. When there is significant variation through time in the probability of default on these contracts, contracting costs increase. In addition, because the human capital of agents is gener-

³ Kenneth J. Arrow, *The Role of Securities in the Optimal Allocation of Risk Bearing*, 31 *Rev. Econ. Stud.* 91 (1964); Gerard Debreu, *Theory of Value* (1959).

⁴ See, for example, Arrow, *supra* note 3; or Eugene F. Fama, *Foundations of Finance* chs. 7 & 8 (1976).

ally employed in a single organization, risk aversion tends to cause them to charge more for any risk they bear than security holders who can diversify risk across many organizations.⁵

Efficient accommodation of large-scale specialized risk bearing by residual claimants is an advantage of corporate common stock. To bond contractual payments to other agents, the common stockholders put up wealth, which is used to purchase assets. If the wealth required to bond promised payments goes beyond the value of inputs optimally purchased rather than rented, common stock proceeds can be used to purchase liquid assets, for example, the securities of other organizations, that have no function except to bond specialization of risk bearing by residual claimants.

3. *Purchase of Organization-specific Assets.* Klein, Crawford, and Alchian and Jensen and Meckling argue that because of conflicts of interest with outside owners of organization-specific assets—assets that have lower value to other organizations—rental contracts for such assets generate higher agency costs than outright purchase.⁶ Common stock, with its capacity for raising wealth from residual claimants, is an efficient vehicle for financing such purchases in activities where using large amounts of organization-specific risky assets is efficient.

4. *Specialization of Management.* In the complicated production and distribution activities of large open corporations, coordinating the activities of agents, recontracting among them, and initiating and implementing resource allocation decisions are specialized tasks which are important to the survival of the organization and largely fall on its managers. However, managerial skills are not necessarily tied to wealth or willingness to bear risk, and incompetent managers who are important residual claimants can be difficult to remove. Thus, ignoring agency problems in the decision process, the survival of a complex organization is enhanced by common stock residual claims that allow specialization of management—in effect, the absence of a classical entrepreneur who is both decision maker and residual risk bearer.

5. *The Market Value Rule for Investment Decisions.* When common stocks are traded without transactions costs in a perfectly competitive capital market, the stockholders agree that resource allocation decisions

⁵ See Patricia B. Reagan & Rene M. Stulz, *Risk Bearing, Labor Contracts, and Capital Markets*, (Working Paper Series No. MERC 82-19 Univ. Rochester Managerial Economics Research Center 1982) for an analysis of risk sharing between internal agents and residual claimants and for references to the related literature.

⁶ Benjamin Klein, Robert Crawford, & Armen A. Alchian, *Vertical Integration, Appropriate Rents, and the Competitive Contracting Process*, 21 *J. Law & Econ.* 297 (1978); Michael C. Jensen & William H. Meckling, *Rights and Production Functions: An Application to Labor-managed Firms and Codetermination*, 52 *J. Bus.* 469 (1979).

should be evaluated according to their contribution to the current market value of their residual claims.⁷ The market value rule weighs current against future resources according to the opportunity costs at which resources can be traded across time in the capital market. For example, the market value rule favors expenditures to reduce the current and future costs of delivering products whenever the current market value of the future cost savings is greater than the current expenditure. Product prices can then be lowered while still covering costs.

In contrast, when the horizon of the residual claims is less than the life of the organization, residual claimants assign zero value to cash flows that occur beyond the horizon.⁸ Similarly, when residual claims are not freely alienable or separable from other roles in the organization, it is rational for risk bearers to attribute lower current value to uncertain cash flows than is implied by capital market prices for the future resources.⁹ As a consequence, ignoring agency problems in the decision process, organizations with common stock residual claims, investing according to the market value rule which is optimal for their residual claimants, will be able to deliver products at lower prices than organizations with restricted residual claims.

C. *The Agency Problems of Common Stock Residual Claims*

The unrestricted nature of the common stock residual claims of open corporations leads to an important agency problem. The decision process is in the hands of professional managers whose interests are not identical to those of residual claimants. This problem of separation of "ownership" and "control"—more precisely, the separation of residual risk bearing from decision functions—has troubled students of open corporations from Adam Smith to Berle and Means and Jensen and Meckling.¹⁰ In "Separation of Ownership and Control"¹¹ we argue that this agency problem is controlled by decision systems that separate the management (initiation and implementation) and control (ratification and monitoring) of important decisions at all levels of the organization.

⁷ See, for example, Eugene F. Fama, *The Effects of a Firm's Investment and Financing Decisions on the Welfare of its Security Holders*, 68 *Am. Econ. Rev.* 272 (1978).

⁸ See E. G. Furubotn & S. Pejovich, *Property Rights, Economic Decentralization and the Evolution of the Yugoslav Firm, 1965-1972*, 16 *J. Law & Econ.* 275 (1973); and Jensen & Meckling, *supra* note 8.

⁹ The details of the argument are in Fama & Jensen, *Organizational Forms*, *supra* note 2.

¹⁰ Adam Smith, *The Wealth of Nations* (Cannan ed. 1904) (1st ed. London 1776); Adolf A. Berle & Gardiner C. Means, *The Modern Corporation and Private Property* (1932); Jensen & Meckling, *supra* note 1.

¹¹ Fama & Jensen, in this issue.

Devices for separating decision management and decision control include (1) decision hierarchies in which the decision initiatives of lower level agents are passed on to higher level agents, first for ratification and then for monitoring, (2) boards of directors that ratify and monitor the organization's most important decisions and hire, fire, and compensate top-level decision managers, and (3) incentive structures that encourage mutual monitoring among decision agents. The costs of such mechanisms for separating decision management from decision control are part of the price that open corporations pay for the benefits of unrestricted common stock residual claims.

III. RESTRICTED VERSUS UNRESTRICTED RESIDUAL CLAIMS

The proprietorships, partnerships, and closed corporations observed in small-scale production activities differ in many ways both from one another and from open corporations. For example, proprietorships have a single residual claimant, whereas partnerships and closed corporations have multiple residual claimants. As a consequence, the residual claim contracts in partnerships and closed corporations must specify rights in net cash flows and procedures for transferring residual claims to new agents more explicitly than the residual claims in proprietorships.

However, for control of the agency problems in the decision process, the common characteristic of the residual claims of proprietorships, partnerships, and closed corporations that distinguishes them from open corporations is that the residual claims are largely restricted to important decision agents. This restriction avoids the agency problems between residual claimants and decision agents that arise because of separation of risk-bearing and decision functions in open corporations. Thus, costly mechanisms for separating the management and control of decisions are avoided.¹²

Restricting residual claims to decision makers controls agency problems between residual claimants and decision agents, but at the expense of the benefits of unrestricted common stock. The decision process suffers efficiency losses because decision agents must be chosen on the basis of wealth and willingness to bear risk as well as for decision skills. Residual claimants forgo optimal diversification so that residual claims and decision making can be combined in a small number of agents. Forgone diversification and limited alienability lower the value of the residual claims, raise the cost of risk-bearing services, and lead to less investment

¹² However, in partnerships and closed corporations, some mechanisms for resolving conflicts among residual claimant decision makers (for example, buy-out rules) are required.

in projects with uncertain payoffs than when residual claims are unrestricted. Finally, because decision agents have limited wealth, restricting residual claims to them also limits resources available for bonding contractual payoffs and for acquiring risky organization-specific assets.

An organizational form survives in an activity when the costs and benefits of its residual claims and the approaches it provides to controlling agency problems combine with available production technology to allow the organization to deliver products at lower prices than other organizational forms. The restricted residual claims of proprietorships, partnerships, and closed corporations are more likely to dominate when technology does not involve important economies of scale that lead to large demands for specialized decision skills, specialized risk bearing, and wealth from residual claimants. In these circumstances, the agency costs saved by restricting residual claims to decision agents outweigh the benefits that would be obtained from separation and specialization of decision and risk-bearing functions. On the other hand, unrestricted common stock residual claims are more likely to dominate when there are important economies of scale in production that (i) can be realized only with a complex decision hierarchy that makes use of specialized decision skills throughout the organization, (ii) generate large aggregate risks to be borne by residual claimants, and (iii) demand large amounts of wealth from residual claimants to purchase risky assets and to bond the payoffs promised to a wide range of agents in the organization. In such complex organizations the benefits of unrestricted common stock residual claims are likely to outweigh the costs of controlling the agency problems inherent in the separation and specialization of decision and risk-bearing functions. In these circumstances, the open corporation is more likely to win the competition for survival.¹³

IV. SPECIAL FORMS OF RESIDUAL CLAIMS

The restriction of residual claims to important decision agents distinguishes the residual claims of proprietorships, partnerships, and closed corporations from the unrestricted residual claims of open corporations. There are, however, other organizational forms, including professional partnerships, financial mutuals, and nonprofits, that offer more unusual residual claims. We explain the special characteristics of the residual claims of these organizations as effective devices for controlling special agency problems.

¹³ In Fama & Jensen, Separation of Ownership, in this issue, we discuss how the diffusion of information among decision agents influences the survival of organizational forms. For simplicity, we have ignored these issues here.

A. Professional Partnerships

Like the proprietorships, partnerships, and closed corporations discussed above, the residual claims of the professional partnerships observed in law, public accounting, medicine, and business consulting are restricted to important decision agents. However, in professional partnerships, a partner's share in net cash flows is renegotiated periodically, and his rights in net cash flows are often limited to his period of service in the organization. In effect, a professional partner's residual claim is a flexible and inalienable share of net cash flows for a limited horizon. Flexible sharing rules, inalienability, and limited horizons distinguish the residual claims of professional partnerships from those of the proprietorships, partnerships, and closed corporations observed in other activities. Moreover, these special features of professional partnership residual claims are generally retained when these organizations become professional service corporations for tax purposes.

1. *Decentralized Decision Making and Restricted Residual Claims.* In professional partnerships, large and small, individuals or small teams work on cases, audits, and so on. Because of the importance of specific knowledge about particular clients—knowledge that is costly to transfer among agents—it is efficient for the teams in large partnerships to make most decisions locally. Thus, with respect to the services rendered to customers, decision control takes place within teams, where interaction and mutual monitoring are heaviest. At this level, however, decision management (initiation and implementation) and decision control (ratification and monitoring) are not separate. To control the resulting agency problems, the residual claims in professional partnerships are restricted to the professional agents who are the important team members and who have major decision making roles. This is consistent with the hypothesis developed in "Separation of Ownership and Control"¹⁴ that combination of decision management and control functions in one or a few agents leads to restriction of residual claims to the important decision agents.

2. *The Demand for Monitoring, Bonding, and Consulting.* Lawyers, public accountants, physicians, and some business consultants provide services where one incompetent act can do large damage to a client. As a consequence, certification and pedigree are important to clients. Moreover, even in the largest professional service organizations, services are rendered in individual cases by one or a few professionals. Responsibility for variation in the quality of services is easily assigned to individual agents, and the performance of agents is often well known to clients. In these circumstances, the value of human capital is sensitive to perfor-

¹⁴ *Id.*

mance. In effect, unlimited liability is imposed on the human capital of professional agents by the market for their services. This gives the professional incentives to purchase monitoring and consulting to help limit losses in the value of human capital.

Since professional services are technical, a lawyer, physician, public accountant, or business consultant is efficiently monitored by others of the same training who can also provide valuable consulting services. Such mutual monitoring and consulting are encouraged when professional agents agree to pool net cash flows and to share liability for the actions of colleagues. Pooling of net cash flows and liability is attractive because it encourages mutual monitoring and consulting. Mutual monitoring and consulting improve the quality of services delivered, control liability losses, and enhance the human capital of the partners. Pooling of net cash flows and liability also has risk-sharing advantages.

The analysis is robust to the fact that partnerships sometimes purchase malpractice insurance. Insurance eliminates variability of liability payoffs by substituting a certain insurance premium. However, if premiums are renegotiated to reflect the malpractice experience of the insured, insurance does not destroy the professional's incentives to be monitored or to consult with other professionals.¹⁵ In addition, insurance covers liability to customers but not reductions in the value of human capital caused by incompetent or malfeasant acts.

3. *Large Professional Partnerships and Flexible Sharing Rules.* Some professional partnerships have hundreds and sometimes thousands of partners. Such large partnerships provide portfolios of specialized services that are marketed and delivered over a wide geographical area. They can also provide large bonds to protect clients against losses from malfeasance or incompetence.¹⁶ Large partnerships are also educational organizations, offering young professionals a wide range of opportunities and interaction with other professionals. We are more concerned, though, with the effects of size on the contract structures of these organizations than with explaining why they are large.

Having attained partner status, a professional may be tempted to free-ride on the efforts of colleagues. The residual claims of large partnerships take a direct approach to this agency problem. The residual claim is not generally a fixed share of net cash flows. Rather, a partner's share is renegotiated annually on the basis of past performance and estimates of likely contributions to future net cash flows. In these large partnerships

¹⁵ David Mayers & Clifford W. Smith, Jr., *On the Corporate Demand for Insurance*, 55 *J. Bus.* 281 (1982), argue that insurance itself is a way to purchase monitoring.

¹⁶ See Linda DeAngelo, *Auditor Size and Audit Quality*, 3 *J. Accounting & Econ.* 183 (1981).

service to a client is delivered by a small group of professionals who interact and monitor one another intensively. The composition of the teams changes from case to case to match specialized talents to specialized problems. As a result, the professionals develop knowledge of the talents and contributions of a range of colleagues. Flexible sharing rules add to partners' incentives to gather and communicate such knowledge to the renegotiation process.

Given flexible sharing rules and the way payoffs are tied to performance, large professional partnerships can be viewed as associations of proprietors who get together to obtain the benefits from marketing a portfolio of specialized skills both to clients and to young professionals who purchase specialized education. Or, since the partners often work in small teams that shift from case to case, a large partnership can be regarded as a fluid association of small partnerships.

4. *Limited Horizon Residual Claims.* Limitations on the horizon covered by residual claims cause organizations to bias decisions against alternatives that generate net cash flows beyond the horizon. In "Organizational Forms and Investment Decisions"¹⁷ we argue that the limited horizon feature of the residual claims of professional partnerships reflects the relative unimportance of assets that are not effectively capitalized in the human capital of existing partners. There are generally no important patents, specialized assets, or technologies to be passed from one generation of partners to the next. Each partner brings a depleting asset—human capital—to the partnership. The annual readjustments of shares in net cash flows that are typical, especially in large professional partnerships, calibrate a partner's payoffs to reflect the current and expected future contributions of his human capital. When a partner's human capital is used up or withdrawn from the organization, contributions to net cash flows cease, and this is reflected in the termination, without substantial compensation, of his residual claim.

This explanation of the limited horizon feature of the residual claims of professional partnerships gets support from several sources:

1. Professional human capital serves as a bond against malfeasance when its value is sensitive to performance. However, professional human capital cannot be sold to cover liability losses to customers. To satisfy the demand for reimbursement for such losses and to bond their services further, partners generally extend their liability to tangible assets held outside the organization (that is, they contract for unlimited liability), or they purchase insurance against liability losses to clients. Such use of

¹⁷ Fama & Jensen, *supra* note 2.

unlimited liability and insurance is consistent with the proposition that the dominant asset in a professional partnership is the inalienable human capital of the partners.

2. Unlike professional partnerships, the proprietorships, partnerships, and closed corporations observed in small-scale production activities commonly have mechanisms for transferring residual claims to the cash flows generated by assets other than human capital. Buy-out provisions with internal pricing rules for residual claims and first refusal rights are examples of such mechanisms. Moreover, the residual claims of these organizations are similar in other respects to those of professional partnerships, for example, restriction of the residual claims to important decision agents and periodic renegotiation of salaries to reflect variation through time in the contribution of human capital to net cash flows.

3. Most important, professional partners drop the limited horizon feature of their residual claims when there are substantial assets in the organization in addition to the human capital of existing partners. For example, a departing partner is generally compensated for his share in assets, such as cash and accounts receivable. More interesting, professional partnerships sometimes have devices for compensating a retiring partner for information about his clients that he passes along to remaining partners. Such payments for information reduce the incentives of partners to take actions that substitute near-term cash flows for long-term cash flows in a manner that inhibits organizational survival. It is also interesting that organizations in business and financial consulting that were once professional partnerships with limited horizon residual claims are tending to reorganize as open corporations. We hypothesize that this is largely caused by the pressure to transfer the rights to valuable nonhuman capital assets owned within the organization from one generation of residual claimants to the next.

B. Financial Mutuals

A common form of organization in financial activities is the mutual. In some financial activities, including life insurance, casualty insurance, and personal savings, mutuals exist side by side with open corporations, and there is no obvious tendency for one form of organization to dominate. Mutuals are dominant among investment mutual funds, but commercial banks are always corporations. Our task is to explain why mutuals survive in some financial activities but not in others.

1. *The Control Function of Redeemable Claims.* An unusual characteristic of mutuals is that the residual claimants are customers, for example, the policyholders of mutual insurance companies, the depositors

of mutual savings banks, and the shareholders of mutual funds. However, the unique characteristic of the residual claims of mutuals, which is important in understanding their survival value, is that the residual claims are redeemable on demand. The policyholder, depositor, or shareholder can, at his initiative, turn in his claim at a price determined by a prespecified rule. For example, the shareholder of an open-end mutual fund can redeem his claim for the market value of his share of the fund's assets, while the whole life or endowment insurance policyholder, like the shareholder of a mutual savings bank, can redeem his claim for its specified value plus accumulated dividends.

There is a special form of diffuse control inherent in the redeemable claims of financial organizations. The withdrawal decisions of redeemable claim holders affect the resources under the control of the organization's managers, and they do so in a more direct fashion than customer decisions in nonfinancial organizations. The decision of the claim holder to withdraw resources is a form of partial takeover or liquidation which deprives management of control over assets. This control right can be exercised independently by each claim holder. It does not require a proxy fight, a tender offer, or any other concerted takeover bid. In contrast, decisions of customers in open nonfinancial corporations, and the repricing of the corporation's securities in the capital market, provide signals about the performance of its decision agents, but without further action, either internal or from the corporate takeover market, the judgments of customers and of the capital market leave the assets owned within the organization under the control of the managers.

2. *The Limitations of Redeemable Claims.* Redeemable claims are not an efficient general financing instrument for nonfinancial organizations. Giving every claim holder the right to force contractions of assets would impose substantial costs on nonfinancial activities. For example, nonfinancial corporations typically have large demands for organization-specific assets that have lower value to other organizations. Substantial costs would be incurred in forced sales of such illiquid assets to accommodate redemptions of claims. In contrast, a financial organization purchases and sells financial assets to meet purchases and redemptions of claims. This is accomplished at low cost because financial assets are not organization specific and can be traded with low transactions costs.

There is a more subtle problem with redeemable residual claims in nonfinancial activities. The pricing rule used to redeem claims preempts development of an outside secondary market for the claims. No one will buy at a price higher than the redemption price or sell at a lower price. The absence of secondary markets for the redeemable claims of financial organizations is no problem since redemption price rules (for example, the net asset value rule for mutual fund shares) can be based on prices of

financial assets quoted in the capital market. In contrast, the residual claims of nonfinancial organizations are claims on uncertain future cash flows. Without a secondary market for the claims, accurate and inexpensive external indexes of their value would not exist, and any internal redemption pricing rule would be costly or arbitrary.

3. *Corporate Financial Organizations.* Our analysis should also explain why some financial organizations are mutuals and others are open corporations. The theory predicts that more of the business of financial mutuals is management of portfolios of financial assets whereas corporate financial organizations are more involved in business activities requiring organization-specific assets that are expensive to trade and that generate uncertain future net cash flows that are not easily priced.

Observation of different financial organizations is roughly consistent with these hypotheses. Most investment mutual funds manage portfolios of traded securities. The funds are open-end mutuals with redeemable residual claims, except for a handful of closed-end funds organized as open corporations with nonredeemable common stock residual claims. Consistent with our hypothesis, the closed-end funds often hold assets such as real estate or shares in new ventures that are expensive to value and to trade, though this is not universal.¹⁸

Commercial banks are required by law to be corporations. Our analysis suggests that they would be corporations in the absence of the requirement. A major part of bank business is providing transaction services. Depositors pay for these services directly or by forgoing returns on deposits. The primary assets of commercial banks are short-term loans. Granting and renewing these loans involves monitoring the borrowers and certifying credit worthiness—a service for which the borrowers pay. The capital value of the stochastic net cash flows from services to depositors and borrowers would not easily be captured in the internal pricing rule of a redeemable residual claim.

What survives in commercial banking is a contract structure involving deposits that, like all redeemable claims, allow the depositors to affect the resources under management control. Consistent with our model, variation in deposits is met by purchases and sales of government and private bonds traded at low cost in secondary markets. Since depositors do not have residual claims on net cash flows from service and other activities, redemption of deposits does not require internal valuation of these net cash flows. The rights to the residual net cash flows are assigned to

¹⁸ See Rex Thompson, *Capital Market Efficiency, Two-Parameter Asset Pricing and the Market for Corporate Control: The Implications of Closed-End Investment Company Discounts and Premiums* (1978) (Ph.D. dissertation, Univ. Rochester, Graduate School of Management).

TABLE I
BUSINESS RECEIPTS AND LONG-TERM NONFINANCIAL ASSETS OF CORPORATE AND MUTUAL
FINANCIAL ORGANIZATIONS, SELECTED YEARS

	1967	1969	1971	1973	1975
Business receipts as a percentage of total receipts:					
Corporate commercial banks	13.6	12.1	14.0	12.0	8.3
Savings and loans	4.7	4.7	6.3	5.4	5.6
Mutual savings banks	2.9	3.0	3.1	2.8	3.1
Corporate life insurance	82.7	82.7	83.0	82.0	81.0
Mutual life insurance	72.9	72.6	72.9	72.1	72.1
Corporate casualty insurance	91.5	89.2	89.7	87.7	87.1
Mutual casualty insurance	94.0	93.0	92.7	92.0	90.1
Long-term nonfinancial assets as a percentage of total assets:					
Corporate commercial banks	2.4	2.7	3.0	3.2	3.0
Savings and loans	2.4	2.4	2.4	2.4	2.4
Mutual savings banks	1.2	1.1	1.2	1.6	1.7
Corporate life insurance	4.9	6.1	5.4	5.4	6.5
Mutual life insurance	2.8	3.1	3.2	3.3	3.5
Corporate casualty insurance	5.3	7.6	9.0	9.5	9.5
Mutual casualty insurance	3.6	3.9	3.7	3.9	3.6

SOURCE.—U.S. Internal Revenue Service, computer tape of corporate statistics of income. Business receipts are revenues other than interest, dividends, and capital gains. Policy premiums are included in business receipts for insurance companies.

common stock. Since the common stock is not redeemable, there are incentives for development of a secondary market. The residual claims against uncertain future net cash flows are then priced more effectively than would be the case with redeemable residual claims for which there would be no secondary market. Such mixed capital structures, with fixed value redeemable claims (policies or deposits) and nonredeemable common stock residual claims, are also characteristic of the savings banks and insurance companies organized as open corporations.

Our analysis should also explain the differences between the corporate and mutual organizations observed in the same financial activity, for example, life insurance or personal saving. Relative to the mutuals, corporate financial organizations should be more involved in business activities other than management of financial assets, and these business activities should involve relatively more nonfinancial assets that can only be varied with large costs. The data on the business receipts (revenues other than interest, dividends, and capital gains) and long-term nonfinancial assets of banks and life insurance companies in Table I are consistent with these

hypotheses. Corporate commercial banks have more business receipts relative to total receipts and more long-term nonfinancial assets relative to total assets than mutual savings banks or savings and loan associations. More interesting, savings and loans, which are sometimes corporations, have relatively more business receipts and long-term nonfinancial assets than mutual savings banks. Likewise, corporate life insurance companies have higher ratios of business receipts to total receipts and higher ratios of long-term nonfinancial assets to total assets than mutual life insurance companies.¹⁹

The data for casualty insurance organizations are less supportive. Consistent with our analysis, mutual casualty companies show lower ratios of long-term nonfinancial assets to total assets than corporate casualty companies. However, contrary to our analysis, the mutuals have higher ratios of business receipts to total receipts.²⁰

Finally, an interesting organizational experiment is taking place in the banking sector. Although commercial banks are required to be corporations, regulations restricting commercial banks and savings banks to different activities are being relaxed. The direction is toward allowing savings banks to provide services such as checking privileges and short-term business loans, previously restricted to commercial banks. If the dominance of the corporate format in commercial banking is not the consequence of regulation, then as savings banks become involved in the service activities of commercial banking, they will tend to organize as corporations. On the other hand, if commercial banking services can be provided at lower prices with the mutual format, corporate commercial banks will not survive when mutual savings banks are allowed to compete with them.

C. *Nonprofit Organizations*

The familiar economic analysis of the entrepreneurial firm is of little help in explaining the dominance of nonprofits in some activities, such as religion, education, research, and classical music, but not in others, including automobile manufacturing, legal services, and popular music. We explain the survival of nonprofits in donor-financed activities as an efficient solution to the special agency problem posed by private donations.

¹⁹ Because policy premiums are included as business receipts, business receipts are a larger fraction of total receipts for insurance companies than for banks. Nevertheless, comparison of the business receipts of corporate and mutual insurance companies is relevant.

²⁰ See David Mayers & Clifford W. Smith, Jr., *Contractual Provisions, Organizational Structure, and Conflict Control in Insurance Markets*, 54 *J. Bus.* 407–33 (1981), for additional hypotheses regarding contract structures in the insurance industry.

1. *Nonprofit Organizations and Donations.* Donations per se do not imply dominance for the nonprofit form. When donations are applied directly to well-defined units of output, a for-profit producer perceives them as a reduction in variable costs or as an increase in demand and increases output accordingly. In fact, we observe unit subsidies both in activities organized on a nonprofit basis, for example, educational scholarships, and in activities organized on a for-profit basis, for example, free tickets to sports events for various groups.

However, some donors wish to provide general donations to particular producers (churches, universities, etc.) rather than unit subsidies. Such unrestricted donations pose agency problems for any organization with residual claimants. Residual claimants contract for rights to net cash flows. When activities are financed in part through donations, part of net cash flow is from resources provided by donors. Contracts that define the share of residual claimants in net cash flows are unlikely to assure donors that their resources are protected from expropriation by residual claimants. One solution to this agency problem is to have no alienable residual claims and to contract with donors to apply all net cash flows to output. Thus, our hypothesis is that the absence of residual claims avoids the donor-residual claimant agency problem and explains the dominance of nonprofits in donor-financed activities.²¹

The absence of alienable residual claims in nonprofits does not mean that residual risk is not borne. When net cash flows are used to expand outputs or to lower the prices of outputs, part of the risk of net cash flows is borne by consumers and part by the factors used to produce the outputs. Thus, residual net cash flows are allocated, but there are no specific residual claimants with alienable property rights in net cash flows. Moreover, the absence of residual claims does not mean that nonprofits make no profits. It means that alienable claims to profits do not exist.

Donations can substitute for the resources provided by residual claimants to purchase assets that are optimally owned rather than rented. When held as endowment, donations also help to bond contracts with

²¹ Henry B. Hansmann, *The Role of Nonprofit Enterprise*, 89 *Yale L. J.* 835 (1980), analyzes the nonprofit organization in detail, but he tends to attribute the nonprofit form more to the nature of products than to the agency problems of donations. He treats donors as customers and looks for product characteristics that would make for "contract failure" in a for-profit framework. For example, charity is delivered to third parties, and the customer (donor) has difficulty verifying delivery. Hansmann also argues that the nonprofit form is attractive for high technology goods (because the customer has difficulty verifying quality) and public goods. However, his approach predicts wider dominance for nonprofits (for example, all high technology or public goods) than is observed. The hypothesis that the nonprofit form is related to donor financing is more promising.

other agents in the organization. From a survival viewpoint the advantage of donations over resources provided by residual claimants is that donors forgo claims on their donations and on the returns earned on the donations, and this tends to allow the organization to deliver its products at lower prices.

Our nonprofit hypothesis deals only with activities financed by donations. Such donor-financed activities are dominated by nonprofits, for example, private universities, churches, hospitals, charities, and cultural performing groups (symphony orchestras, ballet companies, and opera companies). However, the limited scope of the hypothesis means that it cannot explain the nonprofits observed in activities where donations play no role, for example, country clubs.

2. *Other Explanations for Nonprofits.* One criticism of our hypothesis about the causal relation from donations to the nonprofit form is that it ignores the difficulty of measuring and selling the outputs of, for example, churches. The inference is that this explains the nonprofit form in these activities. It is difficult to measure all the things one gets from religion, education, research, or cultural activities. However, the same is true of products such as rock music and legal or psychiatric services marketed by organizations that have residual claims. Moreover, if donations disappeared, for-profit organizations, or more precisely organizations that have alienable residual claims, would arise to supply religion, research, and education. Some for-profit organizations supply these services now. For-profit educational organizations and research groups sell definable parts of their outputs; tuition for education and royalties to patents are examples. For-profit churches might sell ordinations, indulgences, or admission to services. Consistent with our hypothesis, when education and research are provided by organizations that have alienable residual claims, these organizations are not also financed with donations.

Some argue that sale of some products and services (for example, religion) is not acceptable and that this explains the nonprofit form in these activities. This is consistent with our hypothesis. When giving outputs away generates more resources through donations than sale, survival dictates the nonprofit form. Thus, universities generally make research freely available because this generates more resources through research grants and other donations than direct sale of the research. Churches usually do not insist on payment of admission charges or member taxes because they attract more total resources through voluntary contributions.

Coldly economic statements like these lead to the criticism that our analysis leaves no room for altruism. The opposite is true. Altruistic internal agents increase the willingness of altruistic customers and donors

to provide resources. In our terms, the altruism of internal agents allows low cost control of agency problems and acts to bond donors and customers against expropriation. Strong tastes for an organization's outputs on the part of internal agents and customers—what we call altruism in the case of nonprofits—contribute to the survival of any organization. All organizations try to develop such brand loyalty, but the nonprofits are especially successful, perhaps because of the nature of their products.

Some readers claim that donors, customers, and internal agents have tastes for the nonprofit form itself in some activities. To explain the complete dominance of nonprofits in an activity, however, this approach requires uniformity of tastes. If subgroups of customers, internal agents, and donors have no preference for the nonprofit form, we would expect more competition among profit and nonprofit organizations in donor-financed activities.

Finally, tax concessions are important to some nonprofits. However, the major activities dominated by nonprofits, such as religion, private education, research, hospital care, and certain cultural activities, were dominated by nonprofits before taxes were a major issue.²² Our hypothesis about the relation between unrestricted donations and the nonprofit form provides a more consistent explanation of the historical dominance of nonprofits in these activities. On the other hand, tax exemptions probably explain the nonprofits in activities where private donations are not a factor, including nursing homes, homes for the elderly, and private nursery schools.

3. *The General Control Problem in Nonprofits.* The donors of nonprofits have agency problems with internal decision agents similar to those faced by residual claimants in other organizations, such as open corporations and financial mutuals, where important decision managers do not bear a major share of the wealth effects of their decisions. We argue in "Separation of Ownership and Control"²³ that, like all other organizations characterized by separation of decision management from residual risk bearing, a nonprofit is on stronger footing in the competition for survival when it has a decision system that separates the management (initiation and implementation) and control (ratification and monitoring) of important decisions. For nonprofits the survival value of such decision systems is due to the assurances they provide that donations are used effectively and are not easily expropriated.

For example, like open corporations and financial mutuals, donor

²² See *id.*

²³ Fama & Jensen, in this issue.

nonprofits have boards of directors (or trustees) with the power to ratify and monitor important decisions and to hire, fire, and set the compensation of important decision agents. The similarities of the decision control systems of nonprofits, financial mutuals, and open corporations, along with the differences due to special agency problems and special features of residual claims (including the absence thereof), are discussed in "Separation of Ownership and Control."

V. SUMMARY AND CONCLUSIONS

Most goods and services can be produced by any form of organization. Organizations compete for survival, and the form of organization that survives in an activity is the one that delivers the product demanded by customers at the lowest price while covering costs.

The characteristics of residual claims are important both in distinguishing organizations from one another and in explaining the survival of specific organizational forms in specific activities. We explain the survival of organizational forms largely in terms of the comparative advantages of characteristics of residual claims in controlling the agency problems of an activity. The analysis identifies the underlying characteristics of activities that determine the organizational forms that survive.

A. Open Corporations

The common stock residual claims of open corporations are unrestricted in the sense that (1) they are freely alienable, (2) they are rights in net cash flows for the life of the organization, and (3) stockholders are not required to have any other role in the organization. Other things equal, the open corporation is more likely to survive in an activity the greater

1. the benefits of unrestricted risk sharing,
2. the benefits of specialized management,
3. the amount of organization-specific assets to be purchased,
4. the wealth required to bond contractual payoffs, and
5. the lower the cost of separating decision management (initiation and implementation) from decision control (ratification and monitoring).

For example, these factors favor the open corporate form when the technology in an activity implies economies of scale that involve (a) large aggregate residual risks to be shared among residual claimants, (b) large demands for specialized decision agents throughout the organization, and (c) large demands for wealth from residual claimants to bond contracts and to purchase organization-specific assets. Economies of scale are also likely to imply organizations that are complex in the sense that valuable specific knowledge—knowledge that is expensive to transfer across

agents—is widely diffused among agents.²⁴ Such complexity tends to favor unrestricted common stock residual claims which allow specialization of management and delegation of decision functions to agents with valuable relevant knowledge.

The benefits of unrestricted common stock residual claims in activities where optimal organizations are large and complex offset the agency costs resulting from the separation of decision functions and residual risk bearing. In "Separation of Ownership and Control" we contend that these agency costs are controlled by decision structures that separate the management and control of important decisions.

B. Proprietorships, Partnerships, and Closed Corporations

In a fictional world where contracts with decision agents were costlessly written and enforced, separation and specialization of decision and risk-bearing functions would involve no agency costs, and most if not all organizations would have unrestricted residual claims. However, actual organizations can realize the benefits of unrestricted residual claims only by incurring costs to control agency problems between specialized decision agents and specialized residual risk bearers. As a consequence, it is advantageous in some activities to trade the benefits of unrestricted common stock residual claims for the low-cost control of agency problems in the decision process obtained when residual claims are restricted to important decision agents. This restriction is a common characteristic of the residual claims of proprietorships, partnerships, and closed corporations. Other things equal, these organizations with their restricted residual claims are more likely to survive in activities where the costs of separating decision management from decision control are high. They are also more likely to survive when there are no important economies of scale and thus (a) no large demands for unrestricted risk sharing and specialized decision skills, and (b) no large demands for wealth from residual claimants to bond contracts and purchase organization-specific assets.

C. Special Forms of Residual Claims

Organizations such as professional partnerships, financial mutuals, and nonprofits have residual claims with unique characteristics that we explain as devices for controlling special agency problems.

1. *Professional Partnerships.* These are characterized by (1) restriction of residual claims to major decision agents, (2) periodic renegotiation

²⁴ The role of specific knowledge is discussed in Fama & Jensen, Separation of Ownership and Control, in this issue.

of partner shares in net cash flows (flexible sharing rules), and (3) inalienable residual claims in net cash flows with horizons that are often limited to a partner's period of service in the organization. Professional partnerships are more likely to survive in an activity when

1. valuable specific knowledge relevant to both the management and control of decisions is combined and diffused among agents,
2. there are no strong demands for organization-specific tangible assets, and
3. the benefits from consulting and mutual monitoring among decision agents are high.

These characteristics are observed in professional service activities (law, public accounting, and business consulting) where (1) restricting residual claims to important decision agents helps control the agency problems caused by delegating combined decision management and control rights with respect to cases, audits, and so forth, to agents with relevant specific knowledge; (2) the primary asset of the activity is professional human capital; and (3) mutual monitoring and consulting among agents are important to maintain the value of human capital, which is sensitive to performance.

2. *Financial Mutuals.* The distinguishing characteristic of the residual claims of financial mutuals is that the policyholder, depositor, or shareholder can sell his claim to the organization on demand at a price determined by a rule. The decision to withdraw resources by the holder of a redeemable claim is a form of partial takeover or liquidation that deprives management of control over assets. This mechanism for decision control can be exercised independently by each claim holder. It does not require a proxy fight, a tender offer, or any other concerted takeover bid. Mutuals are more likely to survive in an activity the lower the cost

1. of expanding and contracting assets and
2. of obtaining accurate indices of asset values.

These conditions occur in financial organizations where assets are primarily the securities of other organizations. Redeemable residual claims are a low-cost mechanism for controlling agency problems between the residual claimants and the decision agents of financial mutuals because accurate and inexpensive indexes for asset values are available and the assets are traded with low transactions costs. Redeemable claims are a high-cost mechanism for decision control in activities that involve large amounts of assets not traded in secondary markets. Redeemable residual claims are also inefficient in activities that involve large amounts of lumpy or organization-specific assets that can be varied only with large costs.

3. *Nonprofits.* The nonprofit organization is characterized by the

absence of alienable residual claims to net cash flows and contractual constraints on the distribution of net cash flows. Inalienable residual claims are vested in a board of trustees and net cash flows are committed to current and future output. Nonprofits are more likely to survive in an activity

1. the greater is the potential supply of donations and
2. the lower is the cost of separating decision management from decision control.

The nonprofit organization is a solution to the agency problem posed by donations. When the activities of an organization are financed in part through donations, part of stochastic net cash flow is due to the resources provided by donors. Contracts that define the share of residual claimants in net cash flows are unlikely to assure donors that their resources are protected against expropriation by residual claimants. One solution to this agency problem between donors and residual claimants is to have no residual claimants and to contract with donors to apply net cash flows to future output. The absence of alienable residual claims means that decision managers in nonprofits do not bear the wealth effects of their decisions. As in other organizations where residual risk bearing and decision management functions are separated, the resulting agency problems in the decision process are controlled by decision structures that separate the management and control of important decisions.

APPENDIX

BIBLIOGRAPHY

- Arrow, Kenneth J. "The Role of Securities in the Optimal Allocation of Risk Bearing." *Review of Economic Studies* 31, no. 86 (1964): 91-96.
- Berle, Adolf A., and Means, Gardiner C. *The Modern Corporation and Private Property*. New York: Macmillan Publishing Co., 1932.
- DeAngelo, Linda E. "Auditor Size and Audit Quality." *Journal of Accounting and Economics* 3, no. 3 (December 1981): 183-200.
- Debreau, Gerard. *Theory of Value*. New York: John Wiley & Sons, 1959.
- Fama, Eugene F. *Foundations of Finance*. New York: Basic Books, 1976.
- Fama, Eugene F. "The Effects of a Firm's Investment and Financing Decisions on the Welfare of Its Security Holders." *American Economic Review* 68, no. 3 (June 1978): 272-84.
- Fama, Eugene F., and Jensen, Michael C. "Separation of Ownership and Control." *Journal of Law & Economics* 26 (June 1983): 301-25.
- Fama, Eugene F., and Jensen, Michael C. "Organizational Forms and Investment Decisions." Managerial Economics Research Center Working Paper no. MERC 83-03. Rochester, N.Y.: University of Rochester, Graduate School of Management, 1983.
- Furubotn, E. G., and Pejovich, S. "Property Rights, Economic Decentralization

- and the Evolution of the Yugoslav Firm, 1965–1972.” *Journal of Law & Economics* 16 (October 1973): 275–302.
- Hansmann, Henry B. “The Role of Nonprofit Enterprise.” *Yale Law Journal* 89, no. 5 (April 1980): 835–901.
- Jensen, Michael C. “Organization Theory and Methodology.” *Accounting Review* 50 (April 1983).
- Jensen, Michael C., and Meckling, William H. “Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure.” *Journal of Financial Economics* 3, no. 4 (October 1976): 305–60.
- Jensen, Michael C., and Meckling, William H. “Rights and Production Functions: An Application to Labor-managed Firms and Codetermination.” *Journal of Business* 52, no. 4 (October 1979): 469–506.
- Klein, Benjamin; Crawford, Robert; and Alchian, Armen A. “Vertical Integration, Appropriable Rents, and the Competitive Contracting Process.” *Journal of Law & Economics* 21, no. 2 (October 1978): 297–326.
- Mayers, David, and Smith, Clifford D., Jr. “Contractual Provisions, Organizational Structure, and Conflict Control in Insurance Markets.” *Journal of Business* 54, no. 3 (July 1981): 407–33.
- Mayers, David, and Smith, Clifford W., Jr. “On the Corporate Demand for Insurance.” *Journal of Business* 55, no. 2 (April 1982): 281–96.
- Reagan, Patricia B., and Stulz, Rene M. “Risk Bearing, Labor Contracts and Capital Markets.” Managerial Economics Research Center Working Paper Series no. MERC 82-19. Rochester, N.Y.: University of Rochester, Graduate School of Management, November 1982.
- Smith, Adam. *The Wealth of Nations*, 1776. Edited by Edwin Cannan, 1904. Reprint. New York: Modern Library, 1937.
- Thompson, Rex. “Capital Market Efficiency, Two-Parameter Asset Pricing and the Market for Corporate Control: The Implications of Closed-End Investment Company Discounts and Premiums.” Ph.D. dissertation, University of Rochester, Graduate School of Management, 1978.

The Weasel Word 'Social'

The noun 'society', misleading as it is, is relatively innocuous compared with the adjective 'social', which has probably become the most confusing expression in our entire moral and political vocabulary. This has happened only during the past hundred years, during which time its modern usages, and its power and influence, have expanded rapidly from Bismarckian Germany to cover the whole world. The confusion that it spreads, within the very area wherein it is most used, is partly due to its describing not only phenomena produced by various modes of cooperation among men, such as in a 'society', but also the kinds of actions that promote and serve such orders. From this latter usage it has increasingly been turned into an exhortation, a sort of guide-word for rationalist morals intended to displace traditional morals, and now increasingly supplants the word 'good' as a designation of what is morally right. As a result of this 'distinctly dichotomous' character, as *Webster's New Dictionary of Synonyms* appropriately puts it, factual and normative meanings of the word 'social' constantly alternate, and what at first seems a description imperceptibly turns into a prescription.

- On this particular matter, German usage influenced the American language more than English; for by the eighteen-eighties a group of German scholars known as the historical or ethical school of economic research had increasingly substituted the term 'social policy' for the term 'political economy' to designate the study of human interaction. One of the few not to be swept away by this new fashion, Leopold von Wiese, later remarked that only those who were young in the 'social age' – in the decades immediately before the Great War – can appreciate how strong at that time was the inclination to regard the 'social' sphere as a surrogate for religion. One of the most dramatic manifestations of this was the appearance of the so-called social pastors. But 'to be "social"', Wiese insists, 'is not the same as being good or righteous or "righteous in the eyes of God"' (1917). To some of Wiese's students we owe instructive historical studies on the spreading of the term 'social' (see my references in 1976:180).

The extraordinary variety of uses to which the word 'social' has since been put in English is brought home vividly when in the *Fontana Dictionary of Modern Thought* (1977), cited earlier in another context, is found, appropriately preceded by 'Soap Opera', a series of no less than thirty-five combinations of 'social' with some noun or other, from 'Social Action' to 'Social Wholes'. In a similar effort, R. Williams's *Key Words* (1976), the author, although generally referring the reader, with the conventional 'q.v.', to corresponding entries, departed from this

practice with regard to 'social'. Apparently it would have been impractical for him to follow his policy here, and he simply had to abandon it. These examples led me for a while to note down all occurrences of 'social' that I encountered, thus producing the following instructive list of over one hundred and sixty nouns qualified by the adjective 'social':

accounting	action	adjustment
administration	affairs	agreement
age	animal	appeal
awareness	behaviour	being
body	causation	character
circle	climber	compact
composition	comprehension	concern
conception	conflict	conscience
consciousness	consideration	construction
contract	control	credit
cripples	critic (-que)	crusader
decision	demand	democracy
description	development	dimension
discrimination	disease	disposition
distance	duty	economy
end	entity	environment
epistemology	ethics	etiquette
event	evil	fact
factors	fascism	force
framework	function	gathering
geography	goal	good
graces	group	harmony
health	history	ideal
implication	inadequacy	independence
inferiority	institution	insurance
intercourse	justice	knowledge
laws	leader	life
market economy	medicine	migration
mind	morality	morals
needs	obligation	opportunity
order	organism	orientation
outcast	ownership	partner
passion	peace	pension
person	philosophy	pleasure
point of view	policy	position
power	priority	privilege

problem	process	product
progress	property	psychology
rank	realism	realm
Rechtsstaat	recognition	reform
relations	remedy	research
response	responsibility	revolution
right	role	rule of law
satisfaction	science	security
service	signals	significance
Soziolekt (group speech)	solidarity	spirit
structure	stability	standing
status	struggle	student
studies	survey	system
talent	teleology	tenets
tension	theory	thinkers
thought	traits	usefulness
utility	value	views
virtue	want	waste
wealth	will	work
worker	world	

Many of the combinations given here are even more widely used in a negative, critical form: thus 'social adjustment' becomes 'social maladjustment', and the same for 'social disorder', 'social injustice', 'social insecurity', 'social instability', and so on.

It is difficult to conclude from this list alone whether the word 'social' has acquired so many different meanings as to become useless as a tool of communication. However this may be, its practical effect is quite clear and at least threefold. First, it tends pervertedly to insinuate a notion that we have seen from previous chapters to be misconceived – namely, that what has been brought about by the impersonal and spontaneous processes of the extended order is actually the result of deliberate human creation. Second, following from this, it appeals to men to *redesign* what they never could have designed at all. And third, it also has acquired the power to empty the nouns it qualifies of their meaning.

In this last effect, it has in fact become the most harmful instance of what, after Shakespeare's 'I can suck melancholy out of a song, as a weasel suck eggs' (*As You Like It*, II,5), some Americans call a 'weasel word'. As a weasel is alleged to be able to empty an egg without leaving a visible sign, so can these words deprive of content any term to which they are prefixed while seemingly leaving them untouched. A weasel word is used to draw the teeth from a concept one is obliged to employ,

but from which one wishes to eliminate all implications that challenge one's ideological premises.

On current American usage of the expression see the late Mario Pei's *Weasel Words: The Art of Saying What You Don't Mean* (1978), which credits Theodore Roosevelt with having coined the term in 1918, thus suggesting that seventy years ago American statesmen were remarkably well educated. Yet the reader will not find in that book the prize weasel word 'social'.

Though abuse of the word 'social' is international, it has taken perhaps its most extreme forms in West Germany where the constitution of 1949 employed the expression *sozialer Rechtsstaat* (social rule of law) and whence the conception of 'social market economy' has spread – in a sense which its populariser Ludwig Erhard certainly never intended. (He once assured me in conversation that to him the market economy did not have to be *made* social but was so already as a result of its origin.) But while the rule of law and the market are, at the start, fairly clear concepts, the attribute 'social' empties them of any clear meaning. From these uses of the word 'social', German scholars have come to the conclusion that their government is constitutionally subject to the *Sozialstaatsprinzip*, which means little less than that the rule of law has been suspended. Likewise, such German scholars see a conflict between *Rechtsstaat* and *Sozialstaat* and entrench the *soziale Rechtsstaat* in their constitution – one, I may perhaps say, that was written by Fabian muddle-heads inspired by the nineteenth-century inventor of 'National Socialism', Friedrich Naumann (H. Maier, 1972:8).

Similarly, the term 'democracy' used to have a fairly clear meaning; yet 'social democracy' not only served as the name for the radical Austro-Marxism of the inter-war period but now has been chosen in Britain as a label for a political party committed to a sort of Fabian socialism. Yet the traditional term for what is now called the 'social state' was 'benevolent despotism', and the very real problem of achieving such despotism democratically, i.e., while preserving individual freedom, is simply wished away by the concoction 'social democracy'.

'Social Justice' and 'Social Rights'

Much the worst use of 'social', one that wholly destroys the meaning of any word it qualifies, is in the almost universally used phrase 'social justice'. Though I have dealt with this particular matter already at some length, particularly in the second volume on *The Mirage of Social Justice* in my *Law, Legislation and Liberty*, I must at least briefly state the point again here, since it plays such an important part in arguments for and against socialism. The phrase 'social justice' is, as a distinguished

man more courageous than I bluntly expressed it long ago, simply 'a semantic fraud from the same stable as People's Democracy' (Curran, 1958:8). The alarming extent to which the term seems already to have perverted the thinking of the younger generation is shown by a recent Oxford doctor's thesis on *Social Justice* (Miller, 1976), in which the traditional conception of justice is referred to by the extraordinary remark that 'there appears to be a category of private justice'.

I have seen it suggested that 'social' applies to everything that reduces or removes differences of income. But why call such action 'social'? Perhaps because it is a method of securing majorities, that is, votes in addition to those one expects to get for other reasons? This does seem to be so, but it also means of course that every exhortation to us to be 'social' is an appeal for a further step towards the 'social justice' of socialism. Thus use of the term 'social' becomes virtually equivalent to the call for 'distributive justice'. This is, however, irreconcilable with a competitive market order, and with growth or even maintenance of population and of wealth. Thus people have come, through such errors, to call 'social' what is the main obstacle to the very maintenance of 'society'. 'Social' should really be called 'anti-social'.

It is probably true that men would be happier about their economic conditions if they felt that the relative positions of individuals were just. Yet the whole idea behind distributive justice – that each individual ought to receive what he morally deserves – is meaningless in the extended order of human cooperation (or the catallaxy), because the available product (its size, and even its existence) depends on what is in one sense a morally indifferent way of allocating its parts. For reasons already explored, moral desert cannot be determined objectively, and in any case the adaptation of the larger whole to facts yet to be discovered requires that we accept that 'success is based on results, not on motivation' (Alchian, 1950:213). Any extended system of cooperation must adapt itself constantly to changes in its natural environment (which include the life, health and strength of its members); the demand that only changes with just effect should occur is ridiculous. It is nearly as ridiculous as the belief that deliberate organisation of response to such changes can be just. Mankind could neither have reached nor could now maintain its present numbers without an inequality that is neither determined by, nor reconcilable with, any deliberate moral judgements. Effort of course will improve individual chances, but it alone cannot secure results. The envy of those who have tried just as hard, although fully understandable, works against the common interest. Thus, if the common interest is *really* our interest, we must not give in to this very human instinctual trait, but instead allow the market process to determine the reward. Nobody can ascertain, save

through the market, the size of an individual's contribution to the overall product, nor can it otherwise be determined how much remuneration must be tendered to someone to enable him to choose the activity which will add most to the flow of goods and services offered at large. Of course if the latter should be considered morally good, then the market turns out to produce a supremely moral result.

Mankind is split into two hostile groups by promises that have no realisable content. The sources of this conflict cannot be dissipated by compromise, for every concession to factual error merely creates more unrealisable expectations. Yet, an anti-capitalist ethic continues to develop on the basis of errors by people who condemn the wealth-generating institutions to which they themselves owe their existence. Pretending to be lovers of freedom, they condemn several property, contract, competition, advertising, profit, and even money itself. Imagining that their reason can tell them how to arrange human efforts to serve their innate wishes better, they themselves pose a grave threat to civilisation.

The Social Responsibility of Business is to Increase its Profits

by Milton Friedman

The New York Times Magazine, September 13, 1970.

Copyright © 1970 by The New York Times Company.

When I hear businessmen speak eloquently about the "social responsibilities of business in a free-enterprise system," I am reminded of the wonderful line about the Frenchman who discovered at the age of 70 that he had been speaking prose all his life. The businessmen believe that they are defending free enterprise when they declaim that business is not concerned "merely" with profit but also with promoting desirable "social" ends; that business has a "social conscience" and takes seriously its responsibilities for providing employment, eliminating discrimination, avoiding pollution and whatever else may be the catchwords of the contemporary crop of reformers. In fact they are—or would be if they or anyone else took them seriously—preaching pure and unadulterated socialism. Businessmen who talk this way are unwitting puppets of the intellectual forces that have been undermining the basis of a free society these past decades.

The discussions of the "social responsibilities of business" are notable for their analytical looseness and lack of rigor. What does it mean to say that "business" has responsibilities? Only people can have responsibilities. A corporation is an artificial person and in this sense may have artificial responsibilities, but "business" as a whole cannot be said to have responsibilities, even in this vague sense. The first step toward clarity in examining the doctrine of the social responsibility of business is to ask precisely what it implies for whom.

Presumably, the individuals who are to be responsible are businessmen, which means individual proprietors or corporate executives. Most of the discussion of social responsibility is directed at corporations, so in what follows I shall mostly neglect the individual proprietors and speak of corporate executives.

In a free-enterprise, private-property system, a corporate executive is an employee of the owners of the business. He has direct responsibility to his employers. That responsibility is to conduct the business in accordance with their desires, which generally will be to make as much money as possible while conforming to the basic rules of the society, both those embodied in law and those embodied in ethical custom. Of course, in some cases his employers may have a different objective. A group of persons might establish a corporation for an eleemosynary purpose—for example, a hospital or a school. The manager of such a corporation will not have money profit as his objective but the rendering of certain services.

In either case, the key point is that, in his capacity as a corporate executive, the manager is the agent of the individuals who own the corporation or establish the eleemosynary institution, and his primary responsibility is to them.

Needless to say, this does not mean that it is easy to judge how well he is performing his task. But at least the criterion of performance is straightforward, and the persons among whom a voluntary contractual arrangement exists are clearly defined.

Of course, the corporate executive is also a person in his own right. As a person, he may have many other responsibilities that he recognizes or assumes voluntarily—to his family, his conscience, his feelings of charity, his church, his clubs, his city, his country. He may feel impelled by these responsibilities to devote part of his income to causes he regards as worthy, to refuse to work for particular corporations, even to leave his job, for example, to join his country's armed forces. If we wish, we may refer to some of these responsibilities as "social responsibilities." But in these respects he is acting as a principal, not an agent; he is spending his own money or time or energy, not the money of his employers or the time or energy he has contracted to devote to their purposes. If these are "social responsibilities," they are the social responsibilities of individuals, not of business.

What does it mean to say that the corporate executive has a "social responsibility" in his capacity as businessman? If this statement is not pure rhetoric, it must mean that he is to act in some way that is not in the interest of his employers. For example, that he is to refrain from increasing the price of the

product in order to contribute to the social objective of preventing inflation, even though a price increase would be in the best interests of the corporation. Or that he is to make expenditures on reducing pollution beyond the amount that is in the best interests of the corporation or that is required by law in order to contribute to the social objective of improving the environment. Or that, at the expense of corporate profits, he is to hire "hardcore" un-employed instead of better qualified available workmen to contribute to the social objective of reducing poverty.

In each of these cases, the corporate executive would be spending someone else's money for a general social interest. Insofar as his actions in accord with his "social responsibility" reduce returns to stockholders, he is spending their money. Insofar as his actions raise the price to customers, he is spending the customers' money. Insofar as his actions lower the wages of some employees, he is spending their money.

The stockholders or the customers or the employees could separately spend their own money on the particular action if they wished to do so. The executive is exercising a distinct "social responsibility," rather than serving as an agent of the stockholders or the customers or the employees, only if he spends the money in a different way than they would have spent it.

But if he does this, he is in effect imposing taxes, on the one hand, and deciding how the tax proceeds shall be spent, on the other.

This process raises political questions on two levels: principle and consequences. On the level of political principle, the imposition of taxes and the expenditure of tax proceeds are governmental functions. We have established elaborate constitutional, parliamentary and judicial provisions to control these functions, to assure that taxes are imposed so far as possible in accordance with the preferences and desires of the public—after all, "taxation without representation" was one of the battle cries of the American Revolution. We have a system of checks and balances to separate the legislative function of imposing taxes and enacting expenditures from the executive function of collecting taxes and administering expenditure programs and from the judicial function of mediating disputes and interpreting the law.

Here the businessman—self-selected or appointed directly or indirectly by stockholders—is to be simultaneously legislator, executive and, jurist. He is to decide whom to tax by how much and for what purpose, and he is to spend the proceeds—all this guided only by general exhortations from on high to restrain inflation, improve the environment, fight poverty and so on and on.

The whole justification for permitting the corporate executive to be selected by the stockholders is that the executive is an agent serving the interests of his principal. This justification disappears when the corporate executive imposes taxes and spends the proceeds for "social" purposes. He becomes in effect a public employee, a civil servant, even though he remains in name an employee of a private enterprise. On grounds of political principle, it is intolerable that such civil servants—insofar as their actions in the name of social responsibility are real and not just window-dressing—should be selected as they are now. If they are to be civil servants, then they must be elected through a political process. If they are to impose taxes and make expenditures to foster "social" objectives, then political machinery must be set up to make the assessment of taxes and to determine through a political process the objectives to be served.

This is the basic reason why the doctrine of "social responsibility" involves the acceptance of the socialist view that political mechanisms, not market mechanisms, are the appropriate way to determine the allocation of scarce resources to alternative uses.

On the grounds of consequences, can the corporate executive in fact discharge his alleged "social responsibilities?" On the other hand, suppose he could get away with spending the stockholders' or customers' or employees' money. How is he to know how to spend it? He is told that he must contribute to fighting inflation. How is he to know what action of his will contribute to that end? He is presumably an expert in running his company—in producing a product or selling it or financing it. But nothing about his selection makes him an expert on inflation. Will his holding down the price of his product reduce inflationary pressure? Or, by leaving more spending power in the hands of his customers, simply divert it elsewhere? Or, by forcing him to produce less because of the lower price, will it simply contribute to shortages? Even if he could answer these questions, how much cost is he

justified in imposing on his stockholders, customers and employees for this social purpose? What is his appropriate share and what is the appropriate share of others?

And, whether he wants to or not, can he get away with spending his stockholders', customers' or employees' money? Will not the stockholders fire him? (Either the present ones or those who take over when his actions in the name of social responsibility have reduced the corporation's profits and the price of its stock.) His customers and his employees can desert him for other producers and employers less scrupulous in exercising their social responsibilities.

This facet of "social responsibility" doctrine is brought into sharp relief when the doctrine is used to justify wage restraint by trade unions. The conflict of interest is naked and clear when union officials are asked to subordinate the interest of their members to some more general purpose. If the union officials try to enforce wage restraint, the consequence is likely to be wildcat strikes, rank-and-file revolts and the emergence of strong competitors for their jobs. We thus have the ironic phenomenon that union leaders—at least in the U.S.—have objected to Government interference with the market far more consistently and courageously than have business leaders.

The difficulty of exercising "social responsibility" illustrates, of course, the great virtue of private competitive enterprise—it forces people to be responsible for their own actions and makes it difficult for them to "exploit" other people for either selfish or unselfish purposes. They can do good—but only at their own expense.

Many a reader who has followed the argument this far may be tempted to remonstrate that it is all well and good to speak of Government's having the responsibility to impose taxes and determine expenditures for such "social" purposes as controlling pollution or training the hard-core unemployed, but that the problems are too urgent to wait on the slow course of political processes, that the exercise of social responsibility by businessmen is a quicker and surer way to solve pressing current problems.

Aside from the question of fact—I share Adam Smith's skepticism about the benefits that can be expected from "those who affected to trade for the public good"—this argument must be rejected on grounds of principle. What it amounts to is an assertion that those who favor the taxes and expenditures in question have failed to persuade a majority of their fellow citizens to be of like mind and that they are seeking to attain by undemocratic procedures what they cannot attain by democratic procedures. In a free society, it is hard for "evil" people to do "evil," especially since one man's good is another's evil.

I have, for simplicity, concentrated on the special case of the corporate executive, except only for the brief digression on trade unions. But precisely the same argument applies to the newer phenomenon of calling upon stockholders to require corporations to exercise social responsibility (the recent G.M. crusade for example). In most of these cases, what is in effect involved is some stockholders trying to get other stockholders (or customers or employees) to contribute against their will to "social" causes favored by the activists. Insofar as they succeed, they are again imposing taxes and spending the proceeds.

The situation of the individual proprietor is somewhat different. If he acts to reduce the returns of his enterprise in order to exercise his "social responsibility," he is spending his own money, not someone else's. If he wishes to spend his money on such purposes, that is his right, and I cannot see that there is any objection to his doing so. In the process, he, too, may impose costs on employees and customers. However, because he is far less likely than a large corporation or union to have monopolistic power, any such side effects will tend to be minor.

Of course, in practice the doctrine of social responsibility is frequently a cloak for actions that are justified on other grounds rather than a reason for those actions.

To illustrate, it may well be in the long run interest of a corporation that is a major employer in a small community to devote resources to providing amenities to that community or to improving its government. That may make it easier to attract desirable employees, it may reduce the wage bill or lessen losses from pilferage and sabotage or have other worthwhile effects. Or it may be that, given the laws about the deductibility of corporate charitable contributions, the stockholders can contribute more to charities they favor by having the corporation make the gift than by doing it themselves,

since they can in that way contribute an amount that would otherwise have been paid as corporate taxes.

In each of these—and many similar—cases, there is a strong temptation to rationalize these actions as an exercise of "social responsibility." In the present climate of opinion, with its wide spread aversion to "capitalism," "profits," the "soulless corporation" and so on, this is one way for a corporation to generate goodwill as a by-product of expenditures that are entirely justified in its own self-interest.

It would be inconsistent of me to call on corporate executives to refrain from this hyp-ocritical window-dressing because it harms the foundations of a free society. That would be to call on them to exercise a "social re-sponsibility"! If our institutions, and the atti-tudes of the public make it in their self-inter-est to cloak their actions in this way, I cannot summon much indignation to denounce them. At the same time, I can express admiration for those individual proprietors or owners of closely held corporations or stockholders of more broadly held corporations who disdain such tactics as approaching fraud.

Whether blameworthy or not, the use of the cloak of social responsibility, and the nonsense spoken in its name by influential and presti-gious businessmen, does clearly harm the foun-dations of a free society. I have been impressed time and again by the schizophrenic character of many businessmen. They are capable of being extremely farsighted and clearheaded in matters that are internal to their businesses. They are incredibly shortsighted and muddle-headed in matters that are outside their businesses but affect the possible survival of busi-ness in general. This shortsightedness is strikingly exemplified in the calls from many businessmen for wage and price guidelines or controls or income policies. There is nothing that could do more in a brief period to destroy a market system and replace it by a centrally con-trolled system than effective governmental con-trol of prices and wages.

The shortsightedness is also exemplified in speeches by businessmen on social respon-sibility. This may gain them kudos in the short run. But it helps to strengthen the already too prevalent view that the pursuit of profits is wicked and immoral and must be curbed and controlled by external forces. Once this view is adopted, the external forces that curb the market will not be the social consciences, however highly developed, of the pontificating executives; it will be the iron fist of Government bureaucrats. Here, as with price and wage controls, businessmen seem to me to reveal a suicidal impulse.

The political principle that underlies the market mechanism is unanimity. In an ideal free market resting on private property, no individual can coerce any other, all coopera-tion is voluntary, all parties to such coopera-tion benefit or they need not participate. There are no values, no "social" responsibilities in any sense other than the shared values and responsibilities of individuals. Society is a collection of individuals and of the various groups they voluntarily form.

The political principle that underlies the political mechanism is conformity. The indi-vidual must serve a more general social inter-est—whether that be determined by a church or a dictator or a majority. The individual may have a vote and say in what is to be done, but if he is overruled, he must conform. It is appropriate for some to require others to contribute to a general social purpose whether they wish to or not.

Unfortunately, unanimity is not always feasi-ble. There are some respects in which conformity appears unavoidable, so I do not see how one can avoid the use of the political mecha-nism altogether.

But the doctrine of "social responsibility" taken seriously would extend the scope of the political mechanism to every human activity. It does not differ in philosophy from the most explicitly collectivist doctrine. It differs only by professing to believe that collectivist ends can be attained without collectivist means. That is why, in my book *Capitalism and Freedom*, I have called it a "fundamentally subversive doctrine" in a free society, and have said that in such a society, "there is one and only one social responsibility of business—to use it resources and engage in activities designed to increase its profits so long as it stays within the rules of the game, which is to say, engages in open and free competition without deception or fraud."

La universidad, por detrás de la tecnología

Publicada el 20 de enero de 2022



Koldo Echebarria 

Director General en Esade Business & Law School

29 artículos

✓ Siguiendo

Hay dos puntos de intersección entre la educación superior y la tecnología que merecen una atención especial en este momento. Por un lado, la educación puede enriquecerse aplicando la tecnología en sus métodos de aprendizaje y, por otro lado, las nuevas tecnologías necesitan una intermediación educativa eficaz para crear capital humano y transferirse al sistema productivo. En ambos casos, el sistema educativo de nuestro país está perdiendo la carrera, con graves implicaciones en la erosión de la competitividad, pero también en el

El primer punto de intersección se refiere a la medida en que la tecnología es capaz de insertarse en el proceso de aprendizaje y hacerlo más efectivo, lo cual supone no solo formar a un mayor número de estudiantes, sino también incrementar la calidad del aprendizaje. Es sabido que una de las virtudes de la tecnología es crear nuevas abundancias, es decir, hacer accesible a muchos lo que antes era accesible solo a unos pocos. Hoy, por ejemplo, podemos disfrutar de clases gratuitas, impartidas por los mejores profesores del mundo a través de la red, o consultar un volumen ingente de materiales con valor educativo. La tecnología permite adaptar el ritmo del aprendizaje a las necesidades de cada estudiante, mediante materiales y ejercicios grabados o mediante el seguimiento a distancia, así como reproducir las clases, que se liberan de las limitaciones espaciotemporales del tradicional calendario presencial. Además, están surgiendo aplicaciones que permiten conocer la calidad del aprendizaje individual y colectivo, al proporcionar retroalimentación a los estudiantes sobre su adquisición efectiva de conocimientos y habilidades.

Mi observación, en este aspecto, coincide con la expresada por Kenneth Rogoff hace ya algunos años. La productividad del sistema universitario ha ido decayendo históricamente, al desentenderse de las posibilidades educativas que le ofrecen las nuevas tecnologías. El coste por alumno graduado en la universidad ha ido creciendo y se ha traspasado en su totalidad a los estudiantes y a sus familias, a través de unos precios más elevados, y al Estado, al asumir la financiación de la universidad pública. También ha aumentado el tiempo que los estudiantes dedican a la educación superior, al combinar grados y másteres que alargan sus estudios. Esta realidad convive en nuestro país con un porcentaje muy destacado de titulados subempleados conforme a su titulación, que la OCDE cifra en el 40% de los graduados.



diseñan los programas para satisfacer unos requisitos de titulación. Las empresas y las demás organizaciones que reciben a los titulados no forman parte de este proceso. Y los estudiantes no tienen más remedio que aceptar estas condiciones si quieren graduarse. Incluso en un mercado universitario como el norteamericano, con una amplia oferta, esta no ha servido para contener los precios, porque toda ella está sujeta a los mismos requisitos, que son los que ejercen más presión sobre la necesidad de recursos, como el número de profesores o de horas de enseñanza. Un ejemplo de este desequilibrio ha sido, en España, el rechazo de la autoridad del sistema universitario a informar positivamente de los grados de tres años (muy comunes en otros países) sin que se haya evidenciado en modo alguno su falta de eficacia, que el Gobierno ha terminado por elevarlo a decreto.

El propio Rogoff se preguntaba si la pandemia no podría generar una aceleración tecnológica de la universidad que lograra reducir el coste de la educación sin perjudicar la calidad de la enseñanza. Las posibilidades están ahí y hay operadores que van innovando para crear universidades paralelas con unos modelos de aprendizaje disruptivos que no arrastren el legado de las viejas universidades. Destaca alguna experiencia singular, como la de Minerva, una universidad norteamericana que ha reducido los costes a menos de la mitad de la media, utilizando nuevos modelos pedagógicos que aprovechan las nuevas tecnologías; esto no ha sido incompatible con una mayor personalización de la experiencia educativa, ni tampoco ha limitado las expectativas profesionales de los titulados, cuyas primeras colocaciones han sido excelentes. Está por ver, sin embargo, si estas experiencias se generalizan y permiten que el sistema universitario gane accesibilidad, reduciendo costes, y logre una producción más flexible y, al mismo tiempo, de mayor calidad.



permite hacer más productiva la creación y difusión del conocimiento, y combinar las tareas educativas con las propiamente investigadoras. Hay que estimular que los profesores presenten su investigación en el aula, en lugar de repetir conceptos básicos que los estudiantes pueden adquirir directamente, o en comunidades, con los materiales adecuados. Nuestra universidad abusa de las horas de clase, no necesariamente productivas, que contribuyen poco al aprendizaje de los estudiantes. Se ha demostrado que los estudiantes pueden aprender colectivamente de manera muy eficaz si están bien guiados y estructurados en su interacción, y si su trabajo recibe la oportuna retroalimentación.

El segundo punto de intersección es el que cruza la demanda de nuevos conocimientos y habilidades que se deriva de la adopción de las nuevas tecnologías con la oferta de enseñanzas universitarias que están ligadas a este fenómeno. Si la oferta no está alineada con la demanda, perderemos la oportunidad de acelerar la productividad de nuestras empresas, que retrasarán su adaptación tecnológica y, lo que es más importante, ampliaremos la brecha de ingresos entre quienes hayan recibido dicha educación y quienes no la hayan recibido, lo cual aumentará la desigualdad. Esta es la tesis que defienden Claudia Goldin y Laurence F. Katz en un libro muy documentado, *The Race between Education and Technology*, centrado en los Estados Unidos pero cuyas conclusiones, con muchos matices y con retraso en el tiempo, son relevantes para analizar el fenómeno en nuestro país.

Para examinar estos desajustes, no es suficiente considerar el número de titulados que salen de las universidades, sino también el de quienes disponen de conocimientos y habilidades relevantes para el proceso de incorporación de las nuevas tecnologías. Asistimos hoy a grandes desajustes entre los profesionales que las empresas necesitan y los graduados que salen de la universidad. Faltan conocimientos



ofertas de trabajo en titulaciones técnicas no se puede cubrir y, mientras tanto, el número de plazas que la universidad ofrece en estas materias está por debajo de la media europea y son precisamente estas carreras las que registran un mayor porcentaje de abandono, próximo al 20%. Adicionalmente, las empresas echan en falta en los titulados habilidades blandas, fundamentales para su inserción productiva en los puestos de trabajo.

Estos puntos de intersección suponen un enorme desafío para el sistema universitario, difícil de superar desde sus actuales coordenadas políticas y económicas. El sistema requiere reformas profundas que afectan la educación que ofrece, la forma en que la ofrece y su ritmo de adaptación al cambio tecnológico y social. Hay aspectos que deben perdurar en la universidad, como es el foco en el conocimiento científico, tanto en su creación como en su transmisión y el valor de descubrirlo, o la apuesta por una educación que establezca las bases de una trayectoria profesional larga y cambiante. Pero dudo que estos desafíos puedan afrontarse simplemente asignando más recursos, como argumentan a menudo las propias universidades. Como ocurre en otros sectores, la innovación es la respuesta y, con ella, una mayor productividad de los recursos empleados.

Es un hecho que, ante los desajustes entre oferta y demanda, las organizaciones innovan para satisfacer las necesidades no cubiertas. La brecha actual entre la demanda y la oferta de capacidades tecnológicas va a impulsar la aparición de nuevos actores, fuera de los cauces universitarios. Basta recordar, por ejemplo, que durante décadas las titulaciones de MBA no fueron oficiales en nuestro país, pero las aulas se llenaban de estudiantes que sabían que las empresas las valoraban y eran una vía para mejorar sus perspectivas profesionales. Mi apuesta, en este caso, es que, si la universidad no es capaz de correr al ritmo de



How will Language Modelers like ChatGPT Affect Occupations and Industries?

Ed Felten (Princeton)

Manav Raj (University of Pennsylvania)

Robert Seamans (New York University)

6 March 2023

Abstract: Recent dramatic increases in AI language modeling capabilities has led to many questions about the effect of these technologies on the economy. In this paper we present a methodology to systematically assess the extent to which occupations, industries and geographies are exposed to advances in AI language modeling capabilities. We find that the top occupations exposed to language modeling include telemarketers and a variety of post-secondary teachers such as English language and literature, foreign language and literature, and history teachers. We find the top industries exposed to advances in language modeling are legal services and securities, commodities, and investments.

Keywords: artificial intelligence, ChatGPT, language modeling, occupation, technology

1. Introduction

Artificial Intelligence (AI) will likely affect the economy in many ways, potentially boosting economic growth and changing the way people work and play. The effect of AI on work will likely be multi-faceted. In some cases, AI may substitute for work previously done by humans, and in other cases AI may complement work done by humans. The effect on work will likely also vary across industries. Recent research by Goldfarb et al (2020) document that adoption of AI is relatively high in some industries such as information technology and finance, but low in others such as health care and construction. Moreover, trying to understand how AI will affect work is like trying to hit a moving target because the capabilities of AI are still advancing.

A prominent example of how AI capabilities continue to advance are the recent improvements in AI language modeling. In particular, ChatGPT, a language modeler released by Open AI in late 2022, has garnered a huge amount of attention and controversy. Some worry about the negative effects of tools like ChatGPT on jobs, as in the *New York Post* article headlined “ChatGPT could make these jobs obsolete: ‘The wolf is at the door.’”¹ Others see practical and commercial promise from language modeling. For example, Microsoft announced a \$10 billion partnership with Open AI and has linked ChatGPT with its Bing search engine.² Google felt compelled to demonstrate its own language modeler, Bard, but mistakes during the demonstration led Google’s stock price to drop 7%.³ ChatGPT has been banned by J.P. Morgan.⁴ However, at present, most of this is speculation.

In order to better understand how language modelers such as ChatGPT will affect occupations, industries and geographies, we use a methodology developed by Felten et al (2018, 2021). Felten et al created the AI Occupational Exposure (AIOE) measure and used this measure to identify which occupations, industries and geographies are most exposed to AI. In this paper, we describe how the AIOE approach can be adapted to account for the recent advancement of language modeling.

¹ <https://nypost.com/2023/01/25/chat-gpt-could-make-these-jobs-obsolete/>

² <https://www.bloomberg.com/news/articles/2023-01-23/microsoft-makes-multibillion-dollar-investment-in-openai>

³ <https://www.cnbc.com/2023/02/08/alphabet-shares-slip-following-googles-ai-event-.html>

⁴ <https://www.cbsnews.com/news/chatgpt-jpmorgan-chase-bars-workers-from-using-ai-tool/>

We find that the top occupations affected include telemarketers and a variety of post-secondary teachers such as English language and literature, foreign language and literature, and history teachers. We also find the top industries exposed to advances in language modeling are legal services and securities, commodities, and investments.

This article contributes to several literatures. First, by providing a systematic examination of the effect of language modeling across occupations, industries and geographies, it contributes to a nascent literature on the effects of ChatGPT and other language modelers on the economy (e.g. Agarwal et al., 2022; Zarifhonarvar, 2023). More generally, the article builds on a broader set of literature studying the effect of AI on the economy (Furman and Seamans, 2019; Goldfarb et al., 2019). Second, the article builds on and extends a set of papers that provide systematic methodologies for studying how AI affects occupations (e.g., Brynjolfsson et al, 2018; Frey & Osborne, 2017; Tolan et al., 2021; Webb, 2020). The article specifically builds off and extends the methodology described in Felten et al. (2018, 2021). In so doing, the article demonstrates the flexibility of the original Felten et al methodology; it can be adjusted dynamically to assess the impact of changes in AI capabilities. Finally, the article adds to a large literature on the effect of automating technologies on labor (e.g., Acemoglu et al., 2022; Autor, 2015; Frank et al., 2019; Genz et al., 2021).

The article proceeds as follows. Section 2 describes the AI Occupational Exposure (AIOE) measure developed by Felten et al (2018, 2021). Section 3 extends the AIOE to account for recent advances in language modeling. Section 4 provides results, including listing the top 20 most affected occupations and industries. Section 5 concludes.

2. AI Occupational Exposure Methodology

According to Felten et al (2021), the AI Occupational Exposure (AIOE) is a measure of each occupation's "exposure" to AI. The term "exposure" is used so as to be agnostic as to the effects of AI on the occupation, which could involve substitution or augmentation depending on various factors associated with the occupation itself.

The AIOE measure was constructed by linking 10 AI applications (abstract strategy games, real-time video games, image recognition, visual question answering, image generation, reading comprehension, language modeling, translation, speech recognition, and instrumental track

recognition) to 52 human abilities (e.g., oral comprehension, oral expression, inductive reasoning, arm-hand steadiness, etc) using a crowd-sourced matrix that indicates the level of relatedness between each AI application and human ability. Data on the AI applications come from the Electronic Frontier Foundation (EFF) which collects and maintains statistics about the progress of AI across multiple applications. Data on human abilities comes from the Occupational Information Network (O*NET) database developed by the United States Department of Labor. O*NET uses these 52 human abilities to describe the occupational makeup of each of 800+ occupations that it tracks. Each of 800+ occupations can be thought of as a weighted combination of the 52 human abilities. O*NET uses two sets of weights: prevalence and importance.

Once the 10 AI categories and 52 human abilities are linked through the matrix, the AIOE can then be calculated for each occupation. To do this, first we calculate an ability-level exposure as follows:

$$A_{ij} = \sum_{i=1}^{10} x_{ij} \quad (1)$$

Where i indexes the AI application and j indexes the occupational ability. The ability-level exposure, A , is calculated as the sum of the 10 application-ability relatedness scores, x , as constructed using the matrix of crowd-sourced survey data.

We then calculate the AIOE for each occupation k as follows:

$$AIOE_k = \frac{\sum_{j=1}^{52} A_{ij} \times L_{jk} \times I_{jk}}{\sum_{j=1}^{52} L_{jk} \times I_{jk}} \quad (2)$$

In this equation, i indexes the AI application, j indexes the occupational ability, and k indexes the occupation. A_{ij} represents the ability-level exposure score. We weight the ability-level AI exposure by the ability's prevalence (L_{jk}) and importance (I_{jk}) within each occupation as measured by O*NET by multiplying the ability-level AI exposure by the prevalence and importance scores for that ability within each occupation, scaled so that they are equally weighted.

Felten et al (2021) explain the construction of the AIOE scores in more detail, describe how they can be weighted at the industry level to construct an AI Industry Exposure score, or weighted at the geographic level to construct an AI Geographic Exposure score. They also provide results

from a number of validation exercises and describe a number of ways in which the scores can be used by scholars and practitioners.⁵

3. Language Modeling AI Occupational Exposure

The original AIOE described in Felten et al (2021) explicitly weighted each of the AI applications the same. In order to update the AI Occupational Exposure score to account for advances in Language Modeling we modify equation (1) as follows.

$$A_{ij} = \sum_{i=1}^{10} \alpha_i x_{ij} \quad (3)$$

Where i indexes the AI application and j indexes the occupational ability. The ability-level exposure, A , is calculated as the weighted sum of the 10 application-ability relatedness scores, x , as constructed using the matrix of crowd-sourced survey data. α_i is the weight placed on each application i . The weights used in Felten et al (2021) set α_i equal to 1 for each application i .

Next, we set α_i equal to 0 for every AI application except for language modeling, which retains a weight of 1. This then constructs an ability-level exposure measure that only “counts” the value of abilities that are related to language modeling. We then proceed to calculate the $AIOE_k$ for each occupation k using this new “language modeling” weighted A_{ij} . The resulting $AIOE_k$ therefore captures the extent to which each occupation is exposed to advances in language modeling due to AI. A complete list of the occupations and their resulting AIOE language modeling score are listed in the appendix.

The resulting scores are highly correlated with the original AIOE scores (correlation coefficient: 0.979). This can be seen in Figure 1 which plots the original AIOE score and the new language modeling adjusted AIOE score for each occupation.

4. Results

⁵ The Felten et al (2021) paper is open access and available here: <https://onlinelibrary.wiley.com/doi/full/10.1002/smj.3286> The data and code used to create the AIOE scores described in Felten et al (2021) is available on GitHub: <https://github.com/AIOE-Data/AIOE>

In this section we present and briefly discuss tables of “top 20” occupations and industries exposed to language modeling.

4.1. Top 20 Occupations Exposed to Language Modeling

Table 1 provides the list of top 20 occupations exposed to AI based on the original Felten et al (2021) AI Occupational Exposure (AIOE) measure as well as the top 20 occupations exposed to AI enabled advances in language modeling capabilities.

Some occupations occur in both lists, including “clinical, counseling, and school psychologists”, and “history teachers, postsecondary”. Notably, the language modeling list includes more education-related occupations, indicating that occupations in the field of education are likely to be relatively more impacted by advances in language modeling than other occupations. This accords well with the recent spate of articles around how ChatGPT and other language modeling tools affect the way teachers assign work and detect cheating or could use language modeling tools to develop teaching materials.

Also of interest, the top occupation in the language modeling list is “telemarketer.” One might imagine that human telemarketers could benefit from language modeling being used to augment their work. For example, customer responses can be fed into a language modeling engine in real time and relevant, customer-specific prompts quickly fed to the telemarketer. Or, one might imagine that human telemarketers are substituted with language modeling enabled bots. The potential for language modeling to augment or substitute for human telemarketers work highlights one aspect of the AIOE measure: it measures “exposure” to AI, but whether that exposure leads to augmentation or substitution will depend on specifics of any given occupation.

4.2. Top 20 Industries Exposed to Language Modeling

Table 2 provides the list of 20 industries most exposed to AI based on the original Felten et al. (2021) AI Industry Exposure (AIIE) measure as well as the top 20 industries exposed to AI enabled advances in language modeling capabilities.

As before, we see some similarities in the industries categorized as most exposed to AI based on the original AIOE as well as the version that focuses on advances in language modeling

capabilities. For example, “Securities, Commodity Contracts, and Other Financial Investments and Related Activities” is categorized as the most exposed industry using the original AIOE and is the second most exposed industry using the language modeling-focused version of the AIOE. Legal services, insurance and employee benefit funds, and agencies, brokerages, and other insurance related activities are among the top five most exposed industries across both lists.

However, some differences emerge. One salient difference is that the language modeling-focused AIOE suggests a higher exposure to advances in AI within higher education and higher education-adjacent industries. Junior colleges, grantmaking and giving services, and business schools and computer and management training all appear within the top twenty exposed industries.

5. Conclusion

In this paper we present a methodology to systematically assess the extent to which occupations and industries are exposed to advances in AI language modeling capabilities. This methodology relies on the approach described in Felten et al (2021) but adapts it to account for recent advances in language modeling. We find that the top occupations exposed to language modeling include telemarketers and a variety of post-secondary teachers such as English language and literature, foreign language and literature, and history teachers. We also find the top industries exposed to advances in language modeling are legal services and securities, commodities, and investments.

At a broad level, this paper adds to a growing literature studying the effects of AI on labor and work. More specifically, the paper provides a systematic approach for understanding how ChatGPT and other language modelers will affect occupations, industries and geographies. We believe these results will be useful for other scholars as well as practitioners and policymakers.

References

- Acemoglu, D., Autor, D., Hazell, J., & Restrepo, P. (2022). Artificial intelligence and jobs: Evidence from online vacancies. *Journal of Labor Economics*, 40(S1), S293-S340.
- Agrawal, A., Gans, J., Goldfarb, A. 2022. "ChatGPT and How AI Disrupts Industries" *Harvard Business Review*. Available: <https://hbr.org/2022/12/chatgpt-and-how-ai-disrupts-industries>
- Autor, D. (2015). Why Are There Still So Many Jobs? The History and Future of Workplace Automation. *Journal of Economic Perspectives*, 29(3), 3–30.
- Brynjolfsson, E., Mitchell, T., & Rock, D. (2018). What can machines learn, and what does it mean for occupations and the economy? *AEA Papers and Proceedings*, 108, 43–47.
<https://doi.org/10.1257/pandp.20181019>
- Felten, E. W., Raj, M., & Seamans, R. (2018). A method to link advances in artificial intelligence to occupational abilities. In *AEA Papers and Proceedings* (Vol. 108, pp. 54-57).
- Felten, E., Raj, M., & Seamans, R. (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, 42(12), 2195-2217.
- Frank, M. R., Autor, D., Bessen, J. E., Brynjolfsson, E., Cebrian, M., Deming, D. J., ... & Rahwan, I. (2019). Toward understanding the impact of artificial intelligence on labor. *Proceedings of the National Academy of Sciences*, 116(14), 6531-6539.
- Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114, 254–280.
<https://doi.org/10.1016/j.techfore.2016.08.019>
- Furman, J., & Seamans, R. (2019). AI and the Economy. *Innovation policy and the economy*, 19(1), 161-191.

Genz, S., Gregory, T., Janser, M., Lehmer, F., & Matthes, B. (2021). How do workers adjust when firms adopt new technologies?. *ZEW-Centre for European Economic Research Discussion Paper*, (21-073).

Goldfarb, A., Gans, J., & Agrawal, A. (2019). *The economics of artificial intelligence: An agenda*. Chicago, IL: University of Chicago Press.

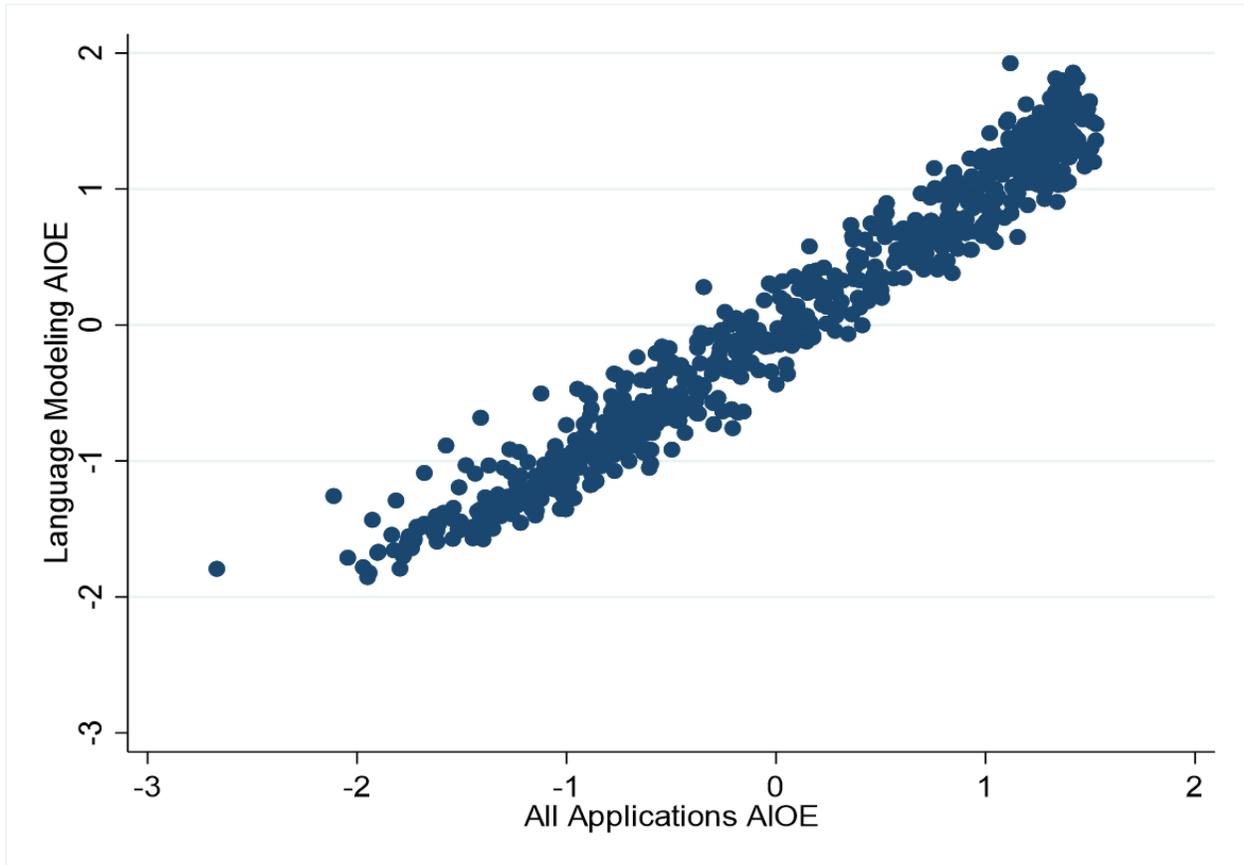
Goldfarb, A., Taska, B. Teodoridis, F. (2020). "Artificial Intelligence in Healthcare? Evidence from Online Job Postings", *AEA Papers and Proceedings*, 110 (5): 400-404

Tolan, S., Pesole, A., Martínez-Plumed, F., Fernández-Macías, E., Hernández-Orallo, J., & Gómez, E. (2021). Measuring the occupational impact of AI: tasks, cognitive abilities and AI benchmarks. *Journal of Artificial Intelligence Research*, 71, 191-236.

Webb, M. (2020). The impact of artificial intelligence on the labor market. *Stanford University working paper*.

Zarifhonarvar, A. 2023. "Economics of ChatGPT: A Labor Market View on the Occupational Impact of Artificial Intelligence" *Indiana University working paper*. Available: <http://dx.doi.org/10.2139/ssrn.4350925>

Figure 1: Comparison between Original AIOE and Language Modeling Adjusted AIOE



Notes: This figure plots the original AIOE score (x-axis) and the new language modeling adjusted AIOE score (y-axis) for each occupation.

Table 1: Top 20 Occupations Exposed to AI, Original and with Language Modeling Adjustment

Rank	Top 20 Occupations from Original AIOE	Top 20 Occupations after Language Modeling Adjustment
1	Genetic Counselors	Telemarketers
2	Financial Examiners	English Language and Literature Teachers, Postsecondary
3	Actuaries	Foreign Language and Literature Teachers, Postsecondary
4	Purchasing Agents, Except Wholesale, Retail, and Farm Products	History Teachers, Postsecondary
5	Budget Analysts	Law Teachers, Postsecondary
6	Judges, Magistrate Judges, and Magistrates	Philosophy and Religion Teachers, Postsecondary
7	Procurement Clerks	Sociology Teachers, Postsecondary
8	Accountants and Auditors	Political Science Teachers, Postsecondary
9	Mathematicians	Criminal Justice and Law Enforcement Teachers, Postsecondary
10	Judicial Law Clerks	Sociologists
11	Education Administrators, Postsecondary	Social Work Teachers, Postsecondary
12	Clinical, Counseling, and School Psychologists	Psychology Teachers, Postsecondary
13	Financial Managers	Communications Teachers, Postsecondary
14	Compensation, Benefits, and Job Analysis Specialists	Political Scientists
15	Credit Authorizers, Checkers, and Clerks	Area, Ethnic, and Cultural Studies Teachers, Postsecondary
16	History Teachers, Postsecondary	Arbitrators, Mediators, and Conciliators
17	Geographers	Judges, Magistrate Judges, and Magistrates
18	Epidemiologists	Geography Teachers, Postsecondary
19	Management Analysts	Library Science Teachers, Postsecondary
20	Arbitrators, Mediators, and Conciliators	Clinical, Counseling, and School Psychologists

Notes: This table lists the top 20 occupations most exposed to AI from the original AIOE (Felten et al., 2021) and the top 20 occupations most exposed to language modeling.

Table 2: Top 20 Industries Exposed to AI, Original and with Language Modeling Adjustment

Rank	Top 20 Industries from Original AIOE	Top 20 Industries after Language Modeling Adjustment
1	Securities, Commodity Contracts, and Other Financial Investments and Related Activities	Legal Services
2	Accounting, Tax Preparation, Bookkeeping, and Payroll Services	Securities, Commodity Contracts, and Other Financial Investments and Related Activities
3	Insurance and Employee Benefit Funds	Agencies, Brokerages, and Other Insurance Related Activities
4	Legal Services	Insurance and Employee Benefit Funds
5	Agencies, Brokerages, and Other Insurance Related Activities	Nondepository Credit Intermediation
6	Nondepository Credit Intermediation	Agents and Managers for Artists, Athletes, Entertainers, and Other Public Figures
7	Other Investment Pools and Funds	Insurance Carriers
8	Insurance Carriers	Other Investment Pools and Funds
9	Software Publishers	Accounting, Tax Preparation, Bookkeeping, and Payroll Services
10	Lessors of Nonfinancial Intangible Assets (except Copyrighted Works)	Business Support Services
11	Agents and Managers for Artists, Athletes, Entertainers, and Other Public Figures	Software Publishers
12	Credit Intermediation and Related Activities (5221 And 5223 only)	Lessors of Nonfinancial Intangible Assets (except Copyrighted Works)
13	Computer Systems Design and Related Services	Business Schools and Computer and Management Training
14	Management, Scientific, and Technical Consulting Services	Credit Intermediation and Related Activities (5221 And 5223 only)
15	Monetary Authorities-Central Bank	Grantmaking and Giving Services
16	Office Administrative Services	Travel Arrangement and Reservation Services
17	Other Information Services	Junior Colleges
18	Data Processing, Hosting, and Related Services	Computer Systems Design and Related Services
19	Business Schools and Computer and Management Training	Management, Scientific, and Technical Consulting Services
20	Grantmaking and Giving Services	Other Information Services

Notes: This table lists the top 20 industries most exposed to AI from the original AIOE (Felten et al., 2021) and the top 20 industries most exposed to language modeling.

APPENDIX

FULL LIST OF OCCUPATIONS SORTED BY LANGUAGE MODELING EXPOSURE SCORE

SOC Code	Occupation Title	Language Modeling AIOE
41-9041	Telemarketers	1.926
25-1123	English Language and Literature Teachers, Postsecondary	1.857
25-1124	Foreign Language and Literature Teachers, Postsecondary	1.814
25-1125	History Teachers, Postsecondary	1.813
25-1112	Law Teachers, Postsecondary	1.802
25-1126	Philosophy and Religion Teachers, Postsecondary	1.800
25-1067	Sociology Teachers, Postsecondary	1.770
25-1065	Political Science Teachers, Postsecondary	1.770
25-1111	Criminal Justice and Law Enforcement Teachers, Postsecondary	1.754
19-3041	Sociologists	1.747
25-1113	Social Work Teachers, Postsecondary	1.739
25-1066	Psychology Teachers, Postsecondary	1.716
25-1122	Communications Teachers, Postsecondary	1.702
19-3094	Political Scientists	1.687
25-1062	Area, Ethnic, and Cultural Studies Teachers, Postsecondary	1.669
23-1022	Arbitrators, Mediators, and Conciliators	1.647
23-1023	Judges, Magistrate Judges, and Magistrates	1.646
25-1064	Geography Teachers, Postsecondary	1.629
25-1082	Library Science Teachers, Postsecondary	1.626
19-3031	Clinical, Counseling, and School Psychologists	1.626
25-1081	Education Teachers, Postsecondary	1.624
25-1011	Business Teachers, Postsecondary	1.618
25-1053	Environmental Science Teachers, Postsecondary	1.603
43-3061	Procurement Clerks	1.590
25-1043	Forestry and Conservation Science Teachers, Postsecondary	1.563
13-1071	Human Resources Specialists	1.557
13-1111	Management Analysts	1.548
23-1021	Administrative Law Judges, Adjudicators, and Hearing Officers	1.547
43-4041	Credit Authorizers, Checkers, and Clerks	1.546
11-9033	Education Administrators, Postsecondary	1.545
21-1014	Mental Health Counselors	1.537
25-1061	Anthropology and Archeology Teachers, Postsecondary	1.534
25-1051	Atmospheric, Earth, Marine, and Space Sciences Teachers, Postsecondary	1.522
27-3031	Public Relations Specialists	1.518
23-1012	Judicial Law Clerks	1.513
41-9091	Door-To-Door Sales Workers, News and Street Vendors, and Related Workers	1.511
13-1023	Purchasing Agents, Except Wholesale, Retail, and Farm Products	1.496
25-1042	Biological Science Teachers, Postsecondary	1.493

19-3093	Historians	1.491
25-3099	Teachers and Instructors, All Other	1.486
43-4161	Human Resources Assistants, Except Payroll and Timekeeping	1.481
29-9092	Genetic Counselors	1.478
13-2072	Loan Officers	1.477
21-1013	Marriage and Family Therapists	1.473
27-3022	Reporters and Correspondents	1.471
21-2011	Clergy	1.470
25-1192	Home Economics Teachers, Postsecondary	1.468
11-3111	Compensation and Benefits Managers	1.467
25-1063	Economics Teachers, Postsecondary	1.462
23-1011	Lawyers	1.454
25-1052	Chemistry Teachers, Postsecondary	1.450
43-4111	Interviewers, Except Eligibility and Loan	1.449
25-3011	Adult Basic and Secondary Education and Literacy Teachers and Instructors	1.440
43-9081	Proofreaders and Copy Markers	1.436
13-1075	Labor Relations Specialists	1.431
25-1193	Recreation and Fitness Studies Teachers, Postsecondary	1.429
19-3032	Industrial-Organizational Psychologists	1.427
41-3021	Insurance Sales Agents	1.427
43-4061	Eligibility Interviewers, Government Programs	1.417
43-2021	Telephone Operators	1.412
25-1022	Mathematical Science Teachers, Postsecondary	1.407
13-1131	Fundraisers	1.405
13-2052	Personal Financial Advisors	1.396
11-3061	Purchasing Managers	1.393
25-1021	Computer Science Teachers, Postsecondary	1.390
23-2093	Title Examiners, Abstractors, and Searchers	1.384
11-2031	Public Relations and Fundraising Managers	1.383
25-1191	Graduate Teaching Assistants	1.376
29-1066	Psychiatrists	1.366
29-1031	Dietitians and Nutritionists	1.364
19-3092	Geographers	1.363
13-1141	Compensation, Benefits, and Job Analysis Specialists	1.362
11-3121	Human Resources Managers	1.359
13-2061	Financial Examiners	1.358
13-1161	Market Research Analysts and Marketing Specialists	1.356
13-2071	Credit Counselors	1.355
21-2021	Directors, Religious Activities and Education	1.355
27-3091	Interpreters and Translators	1.352
41-9021	Real Estate Brokers	1.348
29-1127	Speech-Language Pathologists	1.343

25-1032	Engineering Teachers, Postsecondary	1.337
11-3131	Training and Development Managers	1.337
43-4031	Court, Municipal, and License Clerks	1.336
25-2053	Special Education Teachers, Middle School	1.334
23-2011	Paralegals and Legal Assistants	1.333
19-3011	Economists	1.330
19-3039	Psychologists, All Other	1.330
25-9031	Instructional Coordinators	1.324
13-1151	Training and Development Specialists	1.323
27-3042	Technical Writers	1.317
11-9111	Medical and Health Services Managers	1.315
11-1011	Chief Executives	1.309
13-2031	Budget Analysts	1.300
43-6013	Medical Secretaries	1.298
11-3031	Financial Managers	1.295
11-2022	Sales Managers	1.294
25-2022	Middle School Teachers, Except Special and Career/Technical Education	1.284
19-1041	Epidemiologists	1.284
19-3022	Survey Researchers	1.279
41-3041	Travel Agents	1.278
43-4131	Loan Interviewers and Clerks	1.274
13-2051	Financial Analysts	1.273
43-3011	Bill and Account Collectors	1.270
13-2053	Insurance Underwriters	1.262
41-3031	Securities, Commodities, and Financial Services Sales Agents	1.260
19-2021	Atmospheric and Space Scientists	1.259
41-1012	First-Line Supervisors of Non-Retail Sales Workers	1.259
43-9041	Insurance Claims and Policy Processing Clerks	1.257
43-5011	Cargo and Freight Agents	1.255
43-6014	Secretaries and Administrative Assistants, Except Legal, Medical, and Executive	1.254
13-2081	Tax Examiners and Collectors, and Revenue Agents	1.253
21-1023	Mental Health and Substance Abuse Social Workers	1.253
13-2099	Financial Specialists, All Other	1.253
25-2031	Secondary School Teachers, Except Special and Career/Technical Education	1.249
13-1011	Agents and Business Managers of Artists, Performers, and Athletes	1.248
25-1031	Architecture Teachers, Postsecondary	1.246
43-2011	Switchboard Operators, Including Answering Service	1.244
15-2041	Statisticians	1.243
41-3099	Sales Representatives, Services, All Other	1.242
11-9121	Natural Sciences Managers	1.241

25-1054	Physics Teachers, Postsecondary	1.240
15-2031	Operations Research Analysts	1.234
27-3021	Broadcast News Analysts	1.230
25-1071	Health Specialties Teachers, Postsecondary	1.226
11-9151	Social and Community Service Managers	1.225
11-2011	Advertising and Promotions Managers	1.217
11-9141	Property, Real Estate, and Community Association Managers	1.215
13-2011	Accountants and Auditors	1.214
21-1011	Substance Abuse and Behavioral Disorder Counselors	1.204
11-2021	Marketing Managers	1.204
43-4141	New Accounts Clerks	1.201
15-2011	Actuaries	1.200
11-9032	Education Administrators, Elementary and Secondary School	1.199
21-1012	Educational, Guidance, School, and Vocational Counselors	1.196
31-9094	Medical Transcriptionists	1.196
27-3041	Editors	1.191
43-4021	Correspondence Clerks	1.184
27-3043	Writers and Authors	1.170
19-1042	Medical Scientists, Except Epidemiologists	1.167
15-2021	Mathematicians	1.166
15-1133	Software Developers, Systems Software	1.166
43-3051	Payroll and Timekeeping Clerks	1.155
25-1041	Agricultural Sciences Teachers, Postsecondary	1.154
43-6012	Legal Secretaries	1.149
29-1065	Pediatricians, General	1.146
13-2082	Tax Preparers	1.144
43-3021	Billing and Posting Clerks	1.136
43-1011	First-Line Supervisors of Office and Administrative Support Workers	1.134
13-1051	Cost Estimators	1.134
43-5032	Dispatchers, Except Police, Fire, and Ambulance	1.128
19-4061	Social Science Research Assistants	1.126
27-3011	Radio and Television Announcers	1.124
41-3011	Advertising Sales Agents	1.122
17-2011	Aerospace Engineers	1.117
29-1062	Family and General Practitioners	1.115
13-1031	Claims Adjusters, Examiners, and Investigators	1.106
43-6011	Executive Secretaries and Executive Administrative Assistants	1.106
19-2012	Physicists	1.097
21-1022	Healthcare Social Workers	1.094
13-2041	Credit Analysts	1.092
19-3099	Social Scientists and Related Workers, All Other	1.085
13-1081	Logisticians	1.081

43-4051	Customer Service Representatives	1.076
19-3051	Urban and Regional Planners	1.075
17-2161	Nuclear Engineers	1.073
39-6012	Concierges	1.071
43-9111	Statistical Assistants	1.053
15-1121	Computer Systems Analysts	1.046
15-1122	Information Security Analysts	1.045
17-2081	Environmental Engineers	1.044
15-1199	Computer Occupations, All Other	1.044
41-4012	Sales Representatives, Wholesale and Manufacturing, Except Technical and Scientific Products	1.039
19-2011	Astronomers	1.035
17-2171	Petroleum Engineers	1.033
43-4171	Receptionists and Information Clerks	1.033
27-2041	Music Directors and Composers	1.032
17-2151	Mining and Geological Engineers, Including Mining Safety Engineers	1.031
43-4011	Brokerage Clerks	1.026
41-9031	Sales Engineers	1.026
15-1131	Computer Programmers	1.025
25-2021	Elementary School Teachers, Except Special Education	1.017
39-1011	Gaming Supervisors	1.015
25-1121	Art, Drama, and Music Teachers, Postsecondary	1.007
29-1063	Internists, General	1.001
25-2054	Special Education Teachers, Secondary School	1.000
15-1141	Database Administrators	0.996
25-4021	Librarians	0.978
17-2061	Computer Hardware Engineers	0.977
19-1011	Animal Scientists	0.972
21-1091	Health Educators	0.968
15-1111	Computer and Information Research Scientists	0.968
11-9031	Education Administrators, Preschool and Childcare Center/Program	0.967
21-1021	Child, Family, and School Social Workers	0.957
41-4011	Sales Representatives, Wholesale and Manufacturing, Technical and Scientific Products	0.953
11-3021	Computer and Information Systems Managers	0.945
27-4013	Radio Operators	0.941
25-2052	Special Education Teachers, Kindergarten and Elementary School	0.937
17-2051	Civil Engineers	0.927
13-1199	Business Operations Specialists, All Other	0.925
19-2099	Physical Scientists, All Other	0.913
43-9061	Office Clerks, General	0.911

11-9199	Managers, All Other	0.910
17-2041	Chemical Engineers	0.906
17-2112	Industrial Engineers	0.905
11-3071	Transportation, Storage, and Distribution Managers	0.905
23-2091	Court Reporters	0.903
17-2071	Electrical Engineers	0.901
11-9161	Emergency Management Directors	0.901
21-1092	Probation Officers and Correctional Treatment Specialists	0.897
19-1029	Biological Scientists, All Other	0.883
15-1132	Software Developers, Applications	0.882
19-2041	Environmental Scientists and Specialists, Including Health	0.870
11-9131	Postmasters and Mail Superintendents	0.866
27-2012	Producers and Directors	0.860
17-1012	Landscape Architects	0.836
21-1015	Rehabilitation Counselors	0.835
25-3021	Self-Enrichment Education Teachers	0.825
27-1025	Interior Designers	0.822
43-3031	Bookkeeping, Accounting, and Auditing Clerks	0.818
19-1021	Biochemists and Biophysicists	0.794
13-1022	Wholesale and Retail Buyers, Except Farm Products	0.791
15-1134	Web Developers	0.789
19-2032	Materials Scientists	0.786
25-9021	Farm and Home Management Advisors	0.784
17-2131	Materials Engineers	0.780
43-5031	Police, Fire, and Ambulance Dispatchers	0.774
27-3012	Public Address System and Other Announcers	0.773
11-3011	Administrative Services Managers	0.768
17-2111	Health and Safety Engineers, Except Mining Safety Engineers and Inspectors	0.764
43-4151	Order Clerks	0.761
25-9041	Teacher Assistants	0.748
25-1072	Nursing Instructors and Teachers, Postsecondary	0.744
19-3091	Anthropologists and Archeologists	0.739
13-1021	Buyers and Purchasing Agents, Farm Products	0.737
39-7011	Tour Guides and Escorts	0.736
17-2031	Biomedical Engineers	0.734
27-1011	Art Directors	0.727
11-9071	Gaming Managers	0.719
43-4081	Hotel, Motel, and Resort Desk Clerks	0.711
39-9041	Residential Advisors	0.711
41-9022	Real Estate Sales Agents	0.701
29-1181	Audiologists	0.696
25-4011	Archivists	0.693

17-2021	Agricultural Engineers	0.690
29-1199	Health Diagnosing and Treating Practitioners, All Other	0.679
17-2141	Mechanical Engineers	0.678
11-1021	General and Operations Managers	0.678
17-3012	Electrical and Electronics Drafters	0.676
19-1020	Biologists	0.662
43-5061	Production, Planning, and Expediting Clerks	0.661
27-4032	Film and Video Editors	0.657
15-1152	Computer Network Support Specialists	0.657
25-4012	Curators	0.656
25-2051	Special Education Teachers, Preschool	0.655
25-2012	Kindergarten Teachers, Except Special Education	0.653
53-2021	Air Traffic Controllers	0.648
11-9041	Architectural and Engineering Managers	0.647
25-2023	Career/Technical Education Teachers, Middle School	0.647
11-9021	Construction Managers	0.636
13-2021	Appraisers and Assessors of Real Estate	0.628
39-1021	First-Line Supervisors of Personal Service Workers	0.628
11-9039	Education Administrators, All Other	0.628
29-1069	Physicians and Surgeons, All Other	0.619
33-9021	Private Detectives and Investigators	0.615
17-1011	Architects, Except Landscape and Naval	0.610
19-1013	Soil and Plant Scientists	0.609
29-9011	Occupational Health and Safety Specialists	0.598
13-1032	Insurance Appraisers, Auto Damage	0.582
17-2199	Engineers, All Other	0.581
27-2011	Actors	0.578
19-1012	Food Scientists and Technologists	0.573
13-1041	Compliance Officers	0.572
29-1051	Pharmacists	0.571
19-2042	Geoscientists, Except Hydrologists and Geographers	0.567
27-1021	Commercial and Industrial Designers	0.560
21-1093	Social and Human Service Assistants	0.557
19-2043	Hydrologists	0.555
17-3011	Architectural and Civil Drafters	0.554
29-1081	Podiatrists	0.554
17-3022	Civil Engineering Technicians	0.553
39-7012	Travel Guides	0.553
29-2071	Medical Records and Health Information Technicians	0.552
27-1024	Graphic Designers	0.527
13-1121	Meeting, Convention, and Event Planners	0.513
27-1022	Fashion Designers	0.508
43-9022	Word Processors and Typists	0.507

43-3041	Gaming Cage Workers	0.496
15-1143	Computer Network Architects	0.495
15-1142	Network and Computer Systems Administrators	0.475
19-1022	Microbiologists	0.474
17-1021	Cartographers and Photogrammetrists	0.472
11-9081	Lodging Managers	0.464
29-2092	Hearing Aid Specialists	0.458
17-2072	Electronics Engineers, Except Computer	0.457
15-2091	Mathematical Technicians	0.438
33-9031	Gaming Surveillance Officers and Gaming Investigators	0.433
27-1014	Multimedia Artists and Animators	0.432
43-3071	Tellers	0.431
43-4181	Reservation and Transportation Ticket Agents and Travel Clerks	0.423
25-2032	Career/Technical Education Teachers, Secondary School	0.422
51-8011	Nuclear Power Reactor Operators	0.409
43-9011	Computer Operators	0.408
29-1125	Recreational Therapists	0.401
25-1194	Vocational Education Teachers, Postsecondary	0.390
17-3013	Mechanical Drafters	0.382
39-3012	Gaming and Sports Book Writers and Runners	0.365
29-1064	Obstetricians and Gynecologists	0.364
41-2031	Retail Salespersons	0.357
11-3051	Industrial Production Managers	0.354
17-3026	Industrial Engineering Technicians	0.347
17-2121	Marine Engineers and Naval Architects	0.343
27-4014	Sound Engineering Technicians	0.338
11-9061	Funeral Service Managers	0.337
41-9011	Demonstrators and Product Promoters	0.334
29-9012	Occupational Health and Safety Technicians	0.329
19-1023	Zoologists and Wildlife Biologists	0.327
19-2031	Chemists	0.327
15-1151	Computer User Support Specialists	0.326
25-2011	Preschool Teachers, Except Special Education	0.322
27-2022	Coaches and Scouts	0.319
29-1128	Exercise Physiologists	0.307
29-1161	Nurse Midwives	0.305
29-1171	Nurse Practitioners	0.302
27-2042	Musicians and Singers	0.290
39-3093	Locker Room, Coatroom, and Dressing Room Attendants	0.279
29-9099	Healthcare Practitioners and Technical Workers, All Other	0.279
43-4121	Library Assistants, Clerical	0.276
29-1141	Registered Nurses	0.272
25-9011	Audio-Visual and Multimedia Collections Specialists	0.272

29-1122	Occupational Therapists	0.266
51-8012	Power Distributors and Dispatchers	0.257
31-9099	Healthcare Support Workers, All Other	0.252
53-2022	Airfield Operations Specialists	0.246
43-4071	File Clerks	0.240
31-9092	Medical Assistants	0.236
29-1061	Anesthesiologists	0.212
41-1011	First-Line Supervisors of Retail Sales Workers	0.201
17-3021	Aerospace Engineering and Operations Technicians	0.201
29-1041	Optometrists	0.200
35-9031	Hosts and Hostesses, Restaurant, Lounge, and Coffee Shop	0.182
25-4031	Library Technicians	0.180
19-4099	Life, Physical, and Social Science Technicians, All Other	0.176
43-9021	Data Entry Keyers	0.172
33-3021	Detectives and Criminal Investigators	0.151
41-2021	Counter and Rental Clerks	0.143
53-1021	First-Line Supervisors of Helpers, Laborers, and Material Movers, Hand	0.138
21-1094	Community Health Workers	0.135
29-1151	Nurse Anesthetists	0.134
19-4021	Biological Technicians	0.133
53-6041	Traffic Technicians	0.129
29-2091	Orthotists and Prosthetists	0.126
39-9032	Recreation Workers	0.097
29-2057	Ophthalmic Medical Technicians	0.083
17-3023	Electrical and Electronic Engineering Technicians	0.078
19-1031	Conservation Scientists	0.074
53-1031	First-Line Supervisors of Transportation and Material-Moving Machine and Vehicle Operators	0.068
47-4011	Construction and Building Inspectors	0.068
11-9051	Food Service Managers	0.061
29-1011	Chiropractors	0.052
29-1071	Physician Assistants	0.040
39-3011	Gaming Dealers	0.030
29-2081	Opticians, Dispensing	0.029
39-1012	Slot Supervisors	0.027
29-1023	Orthodontists	0.021
19-4092	Forensic Science Technicians	0.010
29-2053	Psychiatric Technicians	-0.001
51-5111	Prepress Technicians and Workers	-0.001
45-2011	Agricultural Inspectors	-0.008
41-2022	Parts Salespersons	-0.014
33-9099	Protective Service Workers, All Other	-0.017

29-2033	Nuclear Medicine Technologists	-0.023
29-2052	Pharmacy Technicians	-0.023
19-4041	Geological and Petroleum Technicians	-0.027
31-9095	Pharmacy Aides	-0.028
51-1011	First-Line Supervisors of Production and Operating Workers	-0.037
39-5094	Skincare Specialists	-0.037
29-2011	Medical and Clinical Laboratory Technologists	-0.042
19-4011	Agricultural and Food Science Technicians	-0.049
29-1123	Physical Therapists	-0.059
43-9031	Desktop Publishers	-0.065
31-2011	Occupational Therapy Assistants	-0.075
39-5092	Manicurists and Pedicurists	-0.079
17-3029	Engineering Technicians, Except Drafters, All Other	-0.079
29-1131	Veterinarians	-0.084
29-2054	Respiratory Therapy Technicians	-0.088
19-4051	Nuclear Technicians	-0.091
19-4031	Chemical Technicians	-0.094
29-9091	Athletic Trainers	-0.095
41-2012	Gaming Change Persons and Booth Cashiers	-0.104
41-2011	Cashiers	-0.107
27-4012	Broadcast Technicians	-0.111
25-4013	Museum Technicians and Conservators	-0.111
39-3031	Ushers, Lobby Attendants, and Ticket Takers	-0.111
49-1011	First-Line Supervisors of Mechanics, Installers, and Repairers	-0.114
39-4031	Morticians, Undertakers, and Funeral Directors	-0.114
53-6061	Transportation Attendants, Except Flight Attendants	-0.118
27-4011	Audio and Video Equipment Technicians	-0.120
17-1022	Surveyors	-0.121
17-3025	Environmental Engineering Technicians	-0.124
29-2051	Dietetic Technicians	-0.128
29-2031	Cardiovascular Technologists and Technicians	-0.131
29-2035	Magnetic Resonance Imaging Technologists	-0.137
19-4091	Environmental Science and Protection Technicians, Including Health	-0.143
29-1067	Surgeons	-0.145
27-2023	Umpires, Referees, and Other Sports Officials	-0.145
35-1012	First-Line Supervisors of Food Preparation and Serving Workers	-0.150
33-2021	Fire Inspectors and Investigators	-0.152
51-9061	Inspectors, Testers, Sorters, Samplers, and Weighers	-0.156
39-4021	Funeral Attendants	-0.158
11-9013	Farmers, Ranchers, and Other Agricultural Managers	-0.158
27-1027	Set and Exhibit Designers	-0.160
39-9011	Childcare Workers	-0.167

33-3011	Bailiffs	-0.171
35-1011	Chefs and Head Cooks	-0.172
19-1032	Foresters	-0.196
33-1012	First-Line Supervisors of Police and Detectives	-0.196
29-1126	Respiratory Therapists	-0.199
39-9021	Personal Care Aides	-0.206
29-1024	Prosthodontists	-0.214
27-1023	Floral Designers	-0.219
39-3091	Amusement and Recreation Attendants	-0.229
25-2059	Special Education Teachers, All Other	-0.236
31-2021	Physical Therapist Assistants	-0.247
31-9097	Phlebotomists	-0.253
49-2091	Avionics Technicians	-0.264
17-3027	Mechanical Engineering Technicians	-0.274
29-2032	Diagnostic Medical Sonographers	-0.274
31-1013	Psychiatric Aides	-0.275
29-2099	Health Technologists and Technicians, All Other	-0.281
29-2056	Veterinary Technologists and Technicians	-0.282
29-2012	Medical and Clinical Laboratory Technicians	-0.283
43-5051	Postal Service Clerks	-0.284
17-3031	Surveying and Mapping Technicians	-0.290
39-3092	Costume Attendants	-0.296
29-2061	Licensed Practical and Licensed Vocational Nurses	-0.298
35-3021	Combined Food Preparation and Serving Workers, Including Fast Food	-0.308
33-9092	Lifeguards, Ski Patrol, and Other Recreational Protective Service Workers	-0.316
29-1124	Radiation Therapists	-0.316
53-6051	Transportation Inspectors	-0.318
35-3022	Counter Attendants, Cafeteria, Food Concession, and Coffee Shop	-0.321
29-1022	Oral and Maxillofacial Surgeons	-0.330
51-6092	Fabric and Apparel Patternmakers	-0.333
51-4012	Computer Numerically Controlled Machine Tool Programmers, Metal and Plastic	-0.341
33-2022	Forest Fire Inspectors and Prevention Specialists	-0.342
31-9091	Dental Assistants	-0.346
29-1021	Dentists, General	-0.348
29-2021	Dental Hygienists	-0.352
43-5111	Weighers, Measurers, Checkers, and Samplers, Recordkeeping	-0.353
31-2012	Occupational Therapy Aides	-0.357
49-9061	Camera and Photographic Equipment Repairers	-0.360
47-1011	First-Line Supervisors of Construction Trades and Extraction Workers	-0.360

33-1011	First-Line Supervisors of Correctional Officers	-0.361
53-2031	Flight Attendants	-0.365
37-1011	First-Line Supervisors of Housekeeping and Janitorial Workers	-0.367
27-4021	Photographers	-0.382
33-9011	Animal Control Workers	-0.390
35-2011	Cooks, Fast Food	-0.391
33-9091	Crossing Guards	-0.397
13-1074	Farm Labor Contractors	-0.404
31-1011	Home Health Aides	-0.405
29-2034	Radiologic Technologists	-0.405
35-3011	Bartenders	-0.411
33-9032	Security Guards	-0.423
51-3092	Food Batchmakers	-0.427
51-9081	Dental Laboratory Technicians	-0.437
51-8013	Power Plant Operators	-0.441
35-3041	Food Servers, Nonrestaurant	-0.445
49-2011	Computer, Automated Teller, and Office Machine Repairers	-0.454
35-3031	Waiters and Waitresses	-0.469
51-3011	Bakers	-0.474
37-1012	First-Line Supervisors of Landscaping, Lawn Service, and Groundskeeping Workers	-0.486
39-5091	Makeup Artists, Theatrical and Performance	-0.497
51-9151	Photographic Process Workers and Processing Machine Operators	-0.498
39-5093	Shampooers	-0.501
41-9012	Models	-0.503
33-3031	Fish and Game Wardens	-0.506
31-1014	Nursing Assistants	-0.515
33-9093	Transportation Security Screeners	-0.523
43-5081	Stock Clerks and Order Fillers	-0.523
39-2021	Nonfarm Animal Caretakers	-0.528
49-9062	Medical Equipment Repairers	-0.537
29-2055	Surgical Technologists	-0.542
39-5011	Barbers	-0.543
51-9082	Medical Appliance Technicians	-0.548
45-1011	First-Line Supervisors of Farming, Fishing, and Forestry Workers	-0.558
35-2012	Cooks, Institution and Cafeteria	-0.559
51-8031	Water and Wastewater Treatment Plant and System Operators	-0.559
51-8091	Chemical Plant and System Operators	-0.560
53-5031	Ship Engineers	-0.567
35-2015	Cooks, Short Order	-0.567
17-3024	Electro-Mechanical Technicians	-0.568
51-8092	Gas Plant Operators	-0.574
43-9071	Office Machine Operators, Except Computer	-0.577

53-4031	Railroad Conductors and Yardmasters	-0.579
51-8021	Stationary Engineers and Boiler Operators	-0.580
49-9063	Musical Instrument Repairers and Tuners	-0.584
31-9096	Veterinary Assistants and Laboratory Animal Caretakers	-0.590
51-6052	Tailors, Dressmakers, and Custom Sewers	-0.607
31-2022	Physical Therapist Aides	-0.615
39-5012	Hairdressers, Hairstylists, and Cosmetologists	-0.616
51-3093	Food Cooking Machine Operators and Tenders	-0.619
49-9064	Watch Repairers	-0.619
33-3041	Parking Enforcement Workers	-0.627
27-4031	Camera Operators, Television, Video, and Motion Picture	-0.629
51-2093	Timing Device Assemblers and Adjusters	-0.635
33-1021	First-Line Supervisors of Fire Fighting and Prevention Workers	-0.636
51-9071	Jewelers and Precious Stone and Metal Workers	-0.637
35-2013	Cooks, Private Household	-0.639
33-3051	Police and Sheriff's Patrol Officers	-0.640
53-2012	Commercial Pilots	-0.644
39-3021	Motion Picture Projectionists	-0.649
51-8093	Petroleum Pump System Operators, Refinery Operators, and Gaugers	-0.651
51-2023	Electromechanical Equipment Assemblers	-0.652
39-4011	Embalmers	-0.657
49-3052	Motorcycle Mechanics	-0.666
33-3012	Correctional Officers and Jailers	-0.669
49-2095	Electrical and Electronics Repairers, Powerhouse, Substation, and Relay	-0.672
51-9011	Chemical Equipment Operators and Tenders	-0.678
31-9011	Massage Therapists	-0.682
49-2098	Security and Fire Alarm Systems Installers	-0.685
51-4111	Tool and Die Makers	-0.687
27-1013	Fine Artists, Including Painters, Sculptors, and Illustrators	-0.698
29-2041	Emergency Medical Technicians and Paramedics	-0.699
49-2094	Electrical and Electronics Repairers, Commercial and Industrial Equipment	-0.703
45-4023	Log Graders and Scalers	-0.703
43-9051	Mail Clerks and Mail Machine Operators, Except Postal Service	-0.711
51-9141	Semiconductor Processors	-0.713
53-3031	Driver/Sales Workers	-0.722
37-2021	Pest Control Workers	-0.724
53-3022	Bus Drivers, School or Special Client	-0.729
39-2011	Animal Trainers	-0.729
49-2097	Electronic Home Entertainment Equipment Installers and Repairers	-0.730

27-1026	Merchandise Displayers and Window Trimmers	-0.735
43-5071	Shipping, Receiving, and Traffic Clerks	-0.738
49-2021	Radio, Cellular, and Tower Equipment Installers and Repairers	-0.743
47-4041	Hazardous Materials Removal Workers	-0.744
33-3052	Transit and Railroad Police	-0.745
53-1011	Aircraft Cargo Handling Supervisors	-0.751
51-2022	Electrical and Electronic Equipment Assemblers	-0.752
53-2011	Airline Pilots, Copilots, and Flight Engineers	-0.758
35-2014	Cooks, Restaurant	-0.759
19-4093	Forest and Conservation Technicians	-0.761
51-9194	Etchers and Engravers	-0.775
31-9093	Medical Equipment Preparers	-0.776
49-9012	Control and Valve Installers and Repairers, Except Mechanical Door	-0.781
51-4062	Patternmakers, Metal and Plastic	-0.788
51-3091	Food and Tobacco Roasting, Baking, and Drying Machine Operators and Tenders	-0.790
53-5021	Captains, Mates, and Pilots of Water Vessels	-0.793
49-9099	Installation, Maintenance, and Repair Workers, All Other	-0.793
53-6011	Bridge and Lock Tenders	-0.799
51-7031	Model Makers, Wood	-0.802
51-8099	Plant and System Operators, All Other	-0.811
49-3011	Aircraft Mechanics and Service Technicians	-0.814
49-3091	Bicycle Repairers	-0.823
51-4193	Plating and Coating Machine Setters, Operators, and Tenders, Metal and Plastic	-0.828
51-2092	Team Assemblers	-0.832
51-5112	Printing Press Operators	-0.844
53-3011	Ambulance Drivers and Attendants, Except Emergency Medical Technicians	-0.844
51-4081	Multiple Machine Tool Setters, Operators, and Tenders, Metal and Plastic	-0.845
53-7011	Conveyor Operators and Tenders	-0.846
51-9051	Furnace, Kiln, Oven, Drier, and Kettle Operators and Tenders	-0.847
49-9091	Coin, Vending, and Amusement Machine Servicers and Repairers	-0.852
51-4035	Milling and Planing Machine Setters, Operators, and Tenders, Metal and Plastic	-0.867
51-4061	Model Makers, Metal and Plastic	-0.868
51-2021	Coil Winders, Tapers, and Finishers	-0.869
49-3092	Recreational Vehicle Service Technicians	-0.869
49-2096	Electronic Equipment Installers and Repairers, Motor Vehicles	-0.870
47-4099	Construction and Related Workers, All Other	-0.871
53-7071	Gas Compressor and Gas Pumping Station Operators	-0.873

47-5012	Rotary Drill Operators, Oil and Gas	-0.880
51-4022	Forging Machine Setters, Operators, and Tenders, Metal and Plastic	-0.881
27-2032	Choreographers	-0.886
51-4032	Drilling and Boring Machine Tool Setters, Operators, and Tenders, Metal and Plastic	-0.887
51-4011	Computer-Controlled Machine Tool Operators, Metal and Plastic	-0.888
51-6011	Laundry and Dry-Cleaning Workers	-0.889
49-9021	Heating, Air Conditioning, and Refrigeration Mechanics and Installers	-0.896
49-2093	Electrical and Electronics Installers and Repairers, Transportation Equipment	-0.910
53-4041	Subway and Streetcar Operators	-0.913
51-5113	Print Binding and Finishing Workers	-0.914
39-6011	Baggage Porters and Bellhops	-0.914
53-4011	Locomotive Engineers	-0.916
47-2111	Electricians	-0.916
27-1012	Craft Artists	-0.917
51-9196	Paper Goods Machine Setters, Operators, and Tenders	-0.917
51-9083	Ophthalmic Laboratory Technicians	-0.918
43-5021	Couriers and Messengers	-0.922
51-6061	Textile Bleaching and Dyeing Machine Operators and Tenders	-0.923
51-6091	Extruding and Forming Machine Setters, Operators, and Tenders, Synthetic and Glass Fibers	-0.924
51-2011	Aircraft Structure, Surfaces, Rigging, and Systems Assemblers	-0.927
47-4021	Elevator Installers and Repairers	-0.928
49-9031	Home Appliance Repairers	-0.928
37-3012	Pesticide Handlers, Sprayers, and Applicators, Vegetation	-0.934
35-2021	Food Preparation Workers	-0.934
51-4192	Layout Workers, Metal and Plastic	-0.936
49-3051	Motorboat Mechanics and Service Technicians	-0.938
53-3041	Taxi Drivers and Chauffeurs	-0.938
49-2022	Telecommunications Equipment Installers and Repairers, Except Line Installers	-0.939
51-9023	Mixing and Blending Machine Setters, Operators, and Tenders	-0.941
45-2021	Animal Breeders	-0.941
51-9031	Cutters and Trimmers, Hand	-0.946
47-4091	Segmental Pavers	-0.949
51-9123	Painting, Coating, and Decorating Workers	-0.951
47-5031	Explosives Workers, Ordnance Handling Experts, and Blasters	-0.955
49-3053	Outdoor Power Equipment and Other Small Engine Mechanics	-0.961
43-5052	Postal Service Mail Carriers	-0.967
53-7072	Pump Operators, Except Wellhead Pumpers	-0.972

51-3021	Butchers and Meat Cutters	-0.977
51-4041	Machinists	-0.978
49-2092	Electric Motor, Power Tool, and Related Repairers	-0.980
49-9081	Wind Turbine Service Technicians	-0.984
49-9094	Locksmiths and Safe Repairers	-0.991
47-5021	Earth Drillers, Except Oil and Gas	-0.999
45-4011	Forest and Conservation Workers	-1.000
51-6041	Shoe and Leather Workers and Repairers	-1.007
43-5053	Postal Service Mail Sorters, Processors, and Processing Machine Operators	-1.007
51-9111	Packaging and Filling Machine Operators and Tenders	-1.011
53-7063	Machine Feeders and Offbearers	-1.013
49-9043	Maintenance Workers, Machinery	-1.015
51-6051	Sewers, Hand	-1.019
53-4013	Rail Yard Engineers, Dinkey Operators, and Hostlers	-1.021
51-3022	Meat, Poultry, and Fish Cutters and Trimmers	-1.026
43-5041	Meter Readers, Utilities	-1.027
37-2012	Maids and Housekeeping Cleaners	-1.031
51-9193	Cooling and Freezing Equipment Operators and Tenders	-1.031
49-3023	Automotive Service Technicians and Mechanics	-1.031
45-2041	Graders and Sorters, Agricultural Products	-1.032
47-2231	Solar Photovoltaic Installers	-1.037
53-6021	Parking Lot Attendants	-1.038
51-2031	Engine and Other Machine Assemblers	-1.045
51-4122	Welding, Soldering, and Brazing Machine Setters, Operators, and Tenders	-1.048
51-9012	Separating, Filtering, Clarifying, Precipitating, and Still Machine Setters, Operators, and Tenders	-1.048
47-2042	Floor Layers, Except Carpet, Wood, and Hard Tiles	-1.049
53-3021	Bus Drivers, Transit and Intercity	-1.050
51-6042	Shoe Machine Operators and Tenders	-1.052
49-9052	Telecommunications Line Installers and Repairers	-1.054
49-3031	Bus and Truck Mechanics and Diesel Engine Specialists	-1.058
47-5013	Service Unit Operators, Oil, Gas, and Mining	-1.068
49-9071	Maintenance and Repair Workers, General	-1.070
53-4012	Locomotive Firers	-1.073
51-2041	Structural Metal Fabricators and Fitters	-1.078
47-2132	Insulation Workers, Mechanical	-1.080
49-9092	Commercial Divers	-1.082
31-1015	Orderlies	-1.088
45-2093	Farmworkers, Farm, Ranch, and Aquacultural Animals	-1.089
37-2011	Janitors and Cleaners, Except Maids and Housekeeping Cleaners	-1.093
51-4194	Tool Grinders, Filers, and Sharpeners	-1.097

53-7073	Wellhead Pumpers	-1.099
51-4072	Molding, Coremaking, and Casting Machine Setters, Operators, and Tenders, Metal and Plastic	-1.108
53-6031	Automotive and Watercraft Service Attendants	-1.108
51-6062	Textile Cutting Machine Setters, Operators, and Tenders	-1.116
49-9051	Electrical Power-Line Installers and Repairers	-1.119
47-2152	Plumbers, Pipefitters, and Steamfitters	-1.120
47-2011	Boilermakers	-1.122
51-9192	Cleaning, Washing, and Metal Pickling Equipment Operators and Tenders	-1.124
51-4034	Lathe and Turning Machine Tool Setters, Operators, and Tenders, Metal and Plastic	-1.126
53-7041	Hoist and Winch Operators	-1.127
51-9021	Crushing, Grinding, and Polishing Machine Setters, Operators, and Tenders	-1.136
51-9041	Extruding, Forming, Pressing, and Compacting Machine Setters, Operators, and Tenders	-1.138
51-4052	Pourers and Casters, Metal	-1.143
49-9097	Signal and Track Switch Repairers	-1.148
51-4031	Cutting, Punching, and Press Machine Setters, Operators, and Tenders, Metal and Plastic	-1.150
51-6063	Textile Knitting and Weaving Machine Setters, Operators, and Tenders	-1.156
51-4191	Heat Treating Equipment Setters, Operators, and Tenders, Metal and Plastic	-1.171
47-2073	Operating Engineers and Other Construction Equipment Operators	-1.174
53-5022	Motorboat Operators	-1.177
47-2151	Pipelayers	-1.181
51-4023	Rolling Machine Setters, Operators, and Tenders, Metal and Plastic	-1.186
53-7064	Packers and Packers, Hand	-1.193
51-4033	Grinding, Lapping, Polishing, and Buffing Machine Tool Setters, Operators, and Tenders, Metal and Plastic	-1.194
47-2044	Tile and Marble Setters	-1.196
51-9032	Cutting and Slicing Machine Setters, Operators, and Tenders	-1.199
49-3021	Automotive Body and Related Repairers	-1.199
53-3033	Light Truck or Delivery Services Drivers	-1.200
51-7042	Woodworking Machine Setters, Operators, and Tenders, Except Sawing	-1.204
51-7032	Patternmakers, Wood	-1.208
49-9045	Refractory Materials Repairers, Except Brickmasons	-1.209
49-9041	Industrial Machinery Mechanics	-1.217
51-9022	Grinding and Polishing Workers, Hand	-1.220
49-3022	Automotive Glass Installers and Repairers	-1.235

51-7021	Furniture Finishers	-1.238
49-3041	Farm Equipment Mechanics and Service Technicians	-1.239
45-2091	Agricultural Equipment Operators	-1.240
51-9195	Molders, Shapers, and Casters, Except Metal and Plastic	-1.241
47-2211	Sheet Metal Workers	-1.245
47-2031	Carpenters	-1.245
53-7031	Dredge Operators	-1.256
39-9031	Fitness Trainers and Aerobics Instructors	-1.257
49-3042	Mobile Heavy Equipment Mechanics, Except Engines	-1.259
47-4071	Septic Tank Servicers and Sewer Pipe Cleaners	-1.260
53-7081	Refuse and Recyclable Material Collectors	-1.263
51-7041	Sawing Machine Setters, Operators, and Tenders, Wood	-1.265
47-2121	Glaziers	-1.267
53-5011	Sailors and Marine Oilers	-1.269
51-6031	Sewing Machine Operators	-1.273
51-2091	Fiberglass Laminators and Fabricators	-1.277
49-9096	Riggers	-1.278
51-9121	Coating, Painting, and Spraying Machine Setters, Operators, and Tenders	-1.279
47-4051	Highway Maintenance Workers	-1.282
49-9011	Mechanical Door Repairers	-1.282
47-5042	Mine Cutting and Channeling Machine Operators	-1.283
49-9044	Millwrights	-1.284
51-9191	Adhesive Bonding Machine Operators and Tenders	-1.285
33-2011	Firefighters	-1.287
27-2021	Athletes and Sports Competitors	-1.290
51-4121	Welders, Cutters, Solderers, and Brazers	-1.298
47-2071	Paving, Surfacing, and Tamping Equipment Operators	-1.303
51-4051	Metal-Refining Furnace Operators and Tenders	-1.304
53-4021	Railroad Brake, Signal, and Switch Operators	-1.306
49-9098	Helpers--Installation, Maintenance, and Repair Workers	-1.321
51-4021	Extruding and Drawing Machine Setters, Operators, and Tenders, Metal and Plastic	-1.325
51-7011	Cabinetmakers and Bench Carpenters	-1.327
47-2082	Tapers	-1.345
53-7021	Crane and Tower Operators	-1.352
45-4022	Logging Equipment Operators	-1.355
47-5041	Continuous Mining Machine Operators	-1.355
51-9198	Helpers--Production Workers	-1.355
51-9199	Production Workers, All Other	-1.363
53-7032	Excavating and Loading Machine and Dragline Operators	-1.364
47-3012	Helpers--Carpenters	-1.370
45-3021	Hunters and Trappers	-1.374

47-2141	Painters, Construction and Maintenance	-1.381
53-7033	Loading Machine Operators, Underground Mining	-1.390
47-5011	Derrick Operators, Oil and Gas	-1.391
53-3032	Heavy and Tractor-Trailer Truck Drivers	-1.400
53-7121	Tank Car, Truck, and Ship Loaders	-1.403
47-2061	Construction Laborers	-1.405
51-6093	Upholsterers	-1.405
47-2041	Carpet Installers	-1.414
51-6064	Textile Winding, Twisting, and Drawing Out Machine Setters, Operators, and Tenders	-1.417
49-3093	Tire Repairers and Changers	-1.423
45-2092	Farmworkers and Laborers, Crop, Nursery, and Greenhouse	-1.432
35-9011	Dining Room and Cafeteria Attendants and Bartender Helpers	-1.433
49-9095	Manufactured Building and Mobile Home Installers	-1.440
47-3013	Helpers--Electricians	-1.445
47-2072	Pile-Driver Operators	-1.455
47-3015	Helpers--Pipelayers, Plumbers, Pipefitters, and Steamfitters	-1.462
51-9197	Tire Builders	-1.463
49-9093	Fabric Menders, Except Garment	-1.464
47-2022	Stonemasons	-1.483
49-3043	Rail Car Repairers	-1.485
47-2081	Drywall and Ceiling Tile Installers	-1.496
47-4061	Rail-Track Laying and Maintenance Equipment Operators	-1.497
53-7111	Mine Shuttle Car Operators	-1.497
53-7062	Laborers and Freight, Stock, and Material Movers, Hand	-1.498
47-5071	Roustabouts, Oil and Gas	-1.500
47-2131	Insulation Workers, Floor, Ceiling, and Wall	-1.510
45-3011	Fishers and Related Fishing Workers	-1.511
53-7061	Cleaners of Vehicles and Equipment	-1.535
51-3023	Slaughterers and Meat Packers	-1.543
47-2181	Roofers	-1.553
35-9021	Dishwashers	-1.561
47-2142	Paperhangers	-1.566
51-9122	Painters, Transportation Equipment	-1.568
47-5081	Helpers--Extraction Workers	-1.572
47-2043	Floor Sanders and Finishers	-1.574
53-7051	Industrial Truck and Tractor Operators	-1.575
47-2161	Plasterers and Stucco Masons	-1.578
37-3013	Tree Trimmers and Pruners	-1.581
47-5061	Roof Bolters, Mining	-1.593
47-2021	Brickmasons and Blockmasons	-1.607
47-5051	Rock Splitters, Quarry	-1.614
51-4071	Foundry Mold and Coremakers	-1.642

37-3011	Landscaping and Groundskeeping Workers	-1.656
47-3016	Helpers--Roofers	-1.668
47-4031	Fence Erectors	-1.675
47-2051	Cement Masons and Concrete Finishers	-1.678
47-2053	Terrazzo Workers and Finishers	-1.685
47-2221	Structural Iron and Steel Workers	-1.701
47-3014	Helpers--Painters, Paperhangers, Plasterers, and Stucco Masons	-1.711
47-2171	Reinforcing Iron and Rebar Workers	-1.781
45-4021	Fallers	-1.791
27-2031	Dancers	-1.793
47-3011	Helpers--Brickmasons, Blockmasons, Stonemasons, and Tile and Marble Setters	-1.822
51-6021	Pressers, Textile, Garment, and Related Materials	-1.854

5. Professional career — Due: TBA

References:

- Alsop, Ron (2008), “The ‘Trophy Kids’ Go to Work,” *The Wall Street Journal*, October 21. (Adapted from Alsop, Ron, *The Trophy Kids Grow Up: How the Millennial Generation Is Shaking Up the Workplace*, Jossey-Bass, New York, 2008). See for a discussion, <http://www.wsj.com/articles/SB122455219391652725>.
- Callan, Erin (2013), “Is There Life After Work?” *The New York Times*, March 9.
- Newport, Cal (2012), “Solving Gen Y’s Passion Problem,” *Harvard Business Review Blog Network*, September 18.
- Hamori, Monika (2010), “Managing Yourself: Job-Hopping to the Top and Other Career Fallacies,” *Harvard Business Review*, July 29. (Includes “Why Top Young Managers Are in a Nonstop Job Hunt,” by Monika Hamori, Jie Cao, and Burak Koyuncu).

THE WALL STREET JOURNAL.

The 'Trophy Kids' Go to Work

When Gretchen Neels, a Boston-based consultant, was coaching a group of college students for job interviews, she asked them how they believe employers view them. She gave them a clue, telling them that the word she was looking for begins with the letter "e." One young man shouted out, "excellent." Other students chimed in with "enthusiastic" and "energetic." Not even close. The correct answer, she said, is "entitled." "Huh?" the students responded, surprised and even hurt to think that managers are offended by their highfalutin opinions of themselves.

If there is one overriding perception of the millennial generation, it's that these young people have great -- and sometimes outlandish -- expectations. Employers realize the millennials are their future work force, but they are concerned about this generation's desire to shape their jobs to fit their lives rather than adapt their lives to the workplace. Although members of other generations were considered somewhat spoiled in their youth, millennials feel an unusually strong sense of entitlement. Older adults criticize the high-maintenance rookies for demanding too much too soon. "They want to be CEO tomorrow," is a common refrain from corporate recruiters.

More than 85% of hiring managers and human-resource executives said they feel that millennials have a stronger sense of entitlement than older workers, according to a survey by CareerBuilder.com. The generation's greatest expectations: higher pay (74% of respondents); flexible work schedules (61%); a promotion within a year (56%); and more vacation or personal time (50%). "They really do seem to want everything, and I can't decide if it's an inability or an unwillingness to make trade-offs," says Derrick Bolton, assistant dean and M.B.A. admissions director at Stanford University's Graduate School of Business. "They want to be CEO, for example, but they say they don't want to give up time with their families."

Millennials, of course, will have to temper their expectations as they seek employment during this deep economic slump. But their sense of entitlement is an ingrained trait that will likely resurface in a stronger job market. Some research studies indicate that the millennial generation's great expectations stem from feelings of superiority. Michigan State University's Collegiate Employment Research Institute and MonsterTrak, an online careers site, conducted a research study of 18- to 28-year-olds and found that nearly half had moderate to high superiority beliefs about themselves. The superiority factor was measured by responses to such statements as "I deserve favors from others" and "I know that I have more natural talents than most."

For their part, millennials believe they can afford to be picky, with talent shortages looming as baby boomers retire. "They are finding that they have to adjust work around our lives instead of us adjusting our lives around work," a teenage blogger named Olivia writes on the Web site Xanga.com. "What other option do they have? We are hard working and utilize tools to get the job done. But we don't want to work more than 40 hours a week, and we want to wear clothes that are comfortable. We want to be able to spice up the dull workday by listening to our iPods. If corporate America doesn't like that, too bad."

Where do such feelings come from? Blame it on doting parents, teachers and coaches. Millennials are truly "trophy kids," the pride and joy of their parents. The millennials were lavishly praised and often received trophies when they excelled, and sometimes when they didn't, to avoid damaging their self-esteem. They and their parents have placed a high premium on success, with not only academic accolades but also sports and other extracurricular activities.

Now what happens when these trophy kids arrive in the workplace with greater expectations than any generation before them? "Their attitude is always 'What are you going to give me,'" says Natalie Griffith, manager of human-resource programs at Eaton Corp. "It's not necessarily arrogance; it's simply their mindset." Millennials want loads of attention and guidance from employers. An annual or even semiannual evaluation isn't enough. They want to know how they're doing weekly, even daily. "The millennials were raised with so much affirmation and positive reinforcement that they come into the workplace needy for more," says Subha Barry, managing director and head of global diversity and inclusion at Merrill Lynch.

But managers must tread lightly when making a critique. This generation was treated so delicately that many schoolteachers stopped grading papers and tests in harsh-looking red ink. Some managers have seen millennials break down in tears after a negative performance review and even quit their jobs. "They like the constant positive reinforcement, but don't always take suggestions for improvement well," says Steve Canale, recruiting manager at General Electric Co. In performance evaluations, "it's still important to give the good, the bad and the ugly, but with a more positive emphasis." Millennials also want things spelled out clearly. Many flounder without precise guidelines but thrive in structured situations that provide clearly defined rules and the order that they crave. Managers will need to give step-by-step directions for handling everything from projects to voice-mail messages to client meetings. It may seem obvious that employees should show up on time, limit lunchtime to an hour and turn off cellphones during meetings. But those basics aren't necessarily apparent to many millennials.

Gail McDaniel, a corporate consultant and career coach for college students, spoke to managers at a health-care company who were frustrated by some of their millennial employees. It seems that one young man missed an important deadline, and when his manager asked him to explain, he said, "Oh, you forgot to remind me." Parents and teachers aren't doing millennials any favors by constantly adapting to their needs, Ms. McDaniel says. "Going into the workplace, they have an expectation that companies will adapt for them, too." Millennials also expect a flexible work routine that allows them time for their family and personal interests. "For this generation, work is not a place you go; work is a thing you do," says Kaye Foster-Cheek, vice president for human resources at Johnson & Johnson. Although millennials have high expectations about what their employers should provide them, companies shouldn't expect much loyalty in return. If a job doesn't prove fulfilling, millennials will forsake it in a flash. Indeed, many employers say it's retention that worries them most.

In the Michigan State/MonsterTrak study, about two-thirds of the millennials said they would likely "surf" from one job to the next. In addition, about 44% showed their lack of loyalty by stating that they would renege on a job-acceptance commitment if a better offer came along. These workplace nomads don't see any stigma in listing three jobs in a single year on their resumes. They are quite confident about landing yet another job, even if it will take longer in this dismal economy. In the meantime, they needn't worry about their next paycheck because they have their parents to cushion them. They're comfortable in the knowledge

that they can move back home while they seek another job. The weak job market may make millennials think twice about moving on, but once jobs are more plentiful, they will likely resume their job-hopping ways.

Justin Pfister, the founder of Open Yard, an online retailer of sports equipment, believes he and his fellow millennials will resist having their expectations deflated. If employers fail to provide the opportunities and rewards millennials seek, he says, they're likely to drop out of the corporate world as he did and become entrepreneurs. "We get stifled when we're offered single-dimensional jobs," he says. "We are multi-dimensional people living and working in a multi-dimensional world." These outspoken young people tend to be highly opinionated and fearlessly challenge recruiters and bosses. Status and hierarchy don't impress them much. They want to be treated like colleagues rather than subordinates and expect ready access to senior executives, even the CEO, to share their brilliant ideas. Recruiters at such companies as investment-banking firm Goldman Sachs Group Inc. and Amazon.com describe "student stalkers" who brashly fire off emails to everyone from the CEO on down, trying to get an inside track to a job. Companies have a vested interest in trying to slow the millennial mobility rate. They not only will need millennials to fill positions left vacant by retiring baby boomers but also will benefit from this generation's best and brightest, who possess significant strengths in teamwork, technology skills, social networking and multitasking. Millennials were bred for achievement, and most will work hard if the task is engaging and promises a tangible payoff.

Clearly, companies that want to compete for top talent must bend a bit and adapt to the millennial generation. Employers need to show new hires how their work makes a difference and why it's of value to the company. Smart managers will listen to their young employees' opinions, and give them some say in decisions. Employers also can detail the career opportunities available to millennials if they'll just stick around awhile. Indeed, it's the wealth of opportunities that will prove to be the most effective retention tool.

In the final analysis, the generational tension is a bit ironic. After all, the grumbling baby-boomer managers are the same indulgent parents who produced the millennial generation. Ms. Barry of Merrill Lynch sees the irony. She is teaching her teenage daughter to value her own opinions and to challenge things. Now she sees many of those challenging millennials at her company and wonders how she and other managers can expect the kids they raised to suddenly behave differently at work. "It doesn't mean we can be as indulgent as managers as we are as parents," she says. "But as parents of young people just like them, we can treat them with respect."

Adapted from "The Trophy Kids Grow Up: How the Millennial Generation Is Shaking Up the Workplace" by Ron Alsop. Published by Jossey-Bass, a Wiley imprint.

<http://online.wsj.com/news/articles/SB12245219391652725>

March 9, 2013

Is There Life After Work?

By ERIN CALLAN
SANIBEL, Fla.

AT an office party in 2005, one of my colleagues asked my then husband what I did on weekends. She knew me as someone with great intensity and energy. “Does she kayak, go rock climbing and then run a half marathon?” she joked. No, he answered simply, “she sleeps.” And that was true. When I wasn’t catching up on work, I spent my weekends recharging my batteries for the coming week. Work always came first, before my family, friends and marriage — which ended just a few years later.

In recent weeks I have been following with interest the escalating debate about work-life balance and the varying positions of Facebook’s Sheryl Sandberg, Marissa Mayer of Yahoo and the academic [Anne-Marie Slaughter](#), among others. Since I resigned my position as chief financial officer of Lehman Brothers in 2008, amid mounting chaos and a cloud of public humiliation only months before the company went bankrupt, I have had ample time to reflect on the decisions I made in balancing (or failing to balance) my job with the rest of my life. The fact that I call it “the rest of my life” gives you an indication where work stood in the pecking order.

I don’t have children, so it might seem that my story lacks relevance to the work-life balance debate. Like everyone, though, I did have relationships — a spouse, friends and family — and none of them got the best version of me. They got what was left over.

I didn’t start out with the goal of devoting all of myself to my job. It crept in over time. Each year that went by, slight modifications became the new normal. First I spent a half-hour on Sunday organizing my e-mail, to-do list and calendar to make Monday morning easier. Then I was working a few hours on Sunday, then all day. My boundaries slipped away until work was all that was left.

Inevitably, when I left my job, it devastated me. I couldn’t just rally and move on. I did not know how to value who I was versus what I did. What I did *was* who I was.

I have spent several years now living a different version of my life, where I try to apply my energy to my new husband, Anthony, and the people whom I love and care about. But I can’t make up for lost time. Most important, although I now have stepchildren, I missed having a child of my own. I am 47 years old, and Anthony and I have been trying in vitro fertilization for several years. We are still hoping.

Sometimes young women tell me they admire what I’ve done. As they see it, I worked hard for 20 years and can now spend the next 20 focused on other things. But that is not balance. I do not wish that for anyone. Even at the best times in my career, I was never deluded into thinking I had achieved any sort of rational allocation between my life at work and my life outside.

I have often wondered whether I would have been asked to be C.F.O. if I had not worked the way that I did. Until recently, I thought my singular focus on my career was the most powerful ingredient in my success. But I am beginning to realize that I sold myself short. I was talented, intelligent and energetic. It didn’t have to be so extreme. Besides, there were diminishing returns to that kind of labor.

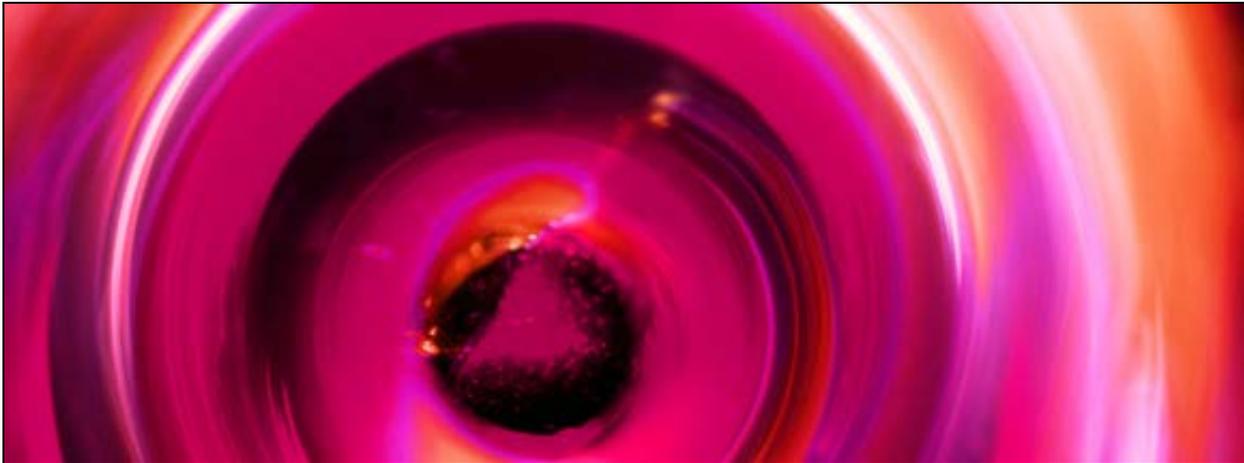
I didn’t have to be on my BlackBerry from my first moment in the morning to my last moment at night. I didn’t have to eat the majority of my meals at my desk. I didn’t have to fly overnight to a meeting in Europe on my birthday. I now believe that I could have made it to a similar place with at least some better version of a personal life. Not without sacrifice — I don’t think I could have “had it all” — but with somewhat more harmony.

I have also wondered where I would be today if Lehman Brothers hadn’t collapsed. In 2007, I did start to have my doubts about the way I was living my life. Or not really living it. But I felt locked in to my career. I had just been asked to be C.F.O. I had a responsibility. Without the crisis, I may never have been strong enough to step away. Perhaps I needed what felt at the time like some of the worst experiences in my life to come to a place where I could be grateful for the life I had. I had to learn to begin to appreciate what was left.

At the end of the day, that is the best guidance I can give. Whatever valuable advice I have about managing a career, I am only now learning how to manage a life.

Erin Callan is the former chief financial officer of Lehman Brothers.

HBR Blog Network



Solving Gen Y's Passion Problem

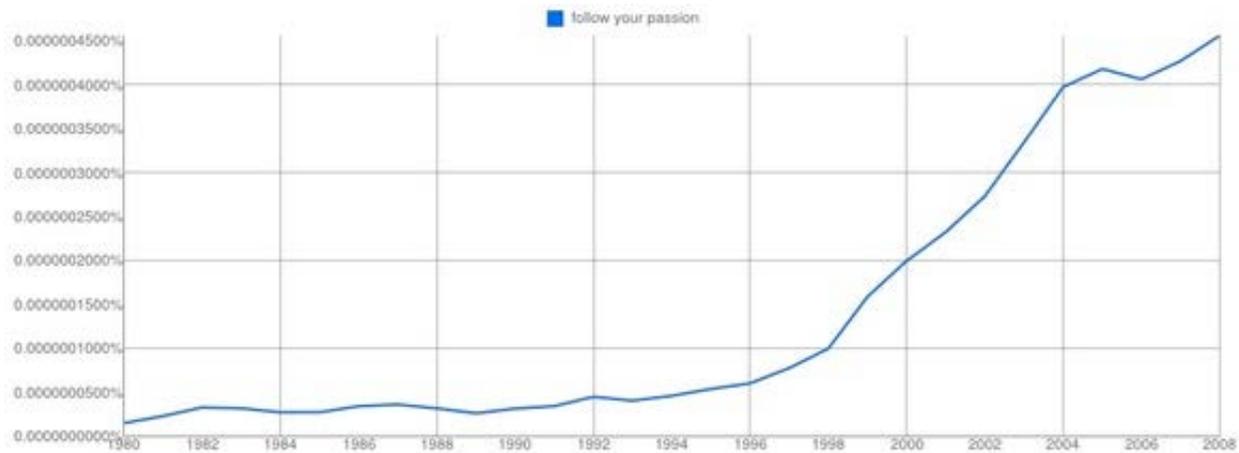
by Cal Newport | 9:00 AM September 18, 2012

Generation Y, of which I'm a member, is entering the job market in record numbers, and according to many commentators things are not going well.

One of the best-known books about my cohort, for instance, is titled *Generation Me* (<http://www.generationme.org/>). The *New York Post* called us "The Worst Generation" (http://www.nypost.com/p/news/business/jobs/the_worst_generation_ZHtISjvJY3GgIWGTIWa0gO), while *USA Today* noted that we are (http://www.usatoday.com/money/workplace/2005-11-06-gen-y_x.htm) "pampered" and "high maintenance." Earlier this year, a *New York Times* op-ed called us "Generation Why Bother" (<http://www.nytimes.com/2012/03/11/opinion/sunday/the-go-nowhere-generation.html>), noting that we're "perhaps...too happy at home checking Facebook," when we could be out aggressively seeking new jobs and helping the economy recover. The fact that up to a third of 25-34 year-olds now live with their parents (<http://www.pewsocialtrends.org/2012/03/15/the-boomerang-generation/>) only supports these gripes.

To many, the core problem of this generation is clear: we're entitled (http://www.unh.edu/news/cj_nr/2010/may/lw17gen-y.cfm). I don't deny these behaviors, but having recently finished researching and writing a book on career advice, I have a different explanation. The problem is not that we're intrinsically selfish or entitled. It's that we've been misinformed.

Generation Y was raised during the period when "follow your passion" became pervasive career advice. The chart below, generated using Google's N-Gram Viewer (http://books.google.com/ngrams/graph?content=follow+your+passion&year_start=1980&year_end=2008&corpus=0&smoothing=3), shows the occurrences of this phrase in printed English over time.



(http://blogs.hbr.org/cs/assets_c/2012/09/Passion_pic_newport-2312.html)

Notice that the phrase begins its rise in the 1990s and skyrockets in the 2000s: the period when Generation Y was in its formative schooling years.

Why is this a problem? This simple phrase, "follow your passion," turns out to be surprisingly pernicious. It's hard to argue, of course, against the general idea that you should aim for a fulfilling working life. But this phrase requires something more. The verb "follow" implies that you start by identifying a passion and then match this preexisting calling to a job. Because the passion precedes the job, it stands to reason that you should love your work from the very first day.

It's this final implication that causes damage. When I studied people who love what they do for a living, I found that in most cases their passion developed slowly, often over unexpected and complicated paths. It's rare, for example, to find someone who loves their career before they've become very good at it — expertise generates many different engaging traits, such as respect, impact, autonomy — and the process of becoming good can be frustrating and take years.

The early stages of a fantastic career might not feel fantastic at all, a reality that clashes with the fantasy world implied by the advice to "follow your passion" — an alternate universe where there's a perfect job waiting for you, one that you'll love right away once you discover it. It shouldn't be surprising that members of Generation Y demand a lot from their working life right away and are frequently disappointed about what they experience instead.

The good news is that this explanation yields a clear solution: we need a more nuanced conversation surrounding the quest for a compelling career. We currently lack, for example, a good phrase for describing those tough first years on a job where you grind away at building up skills while being shoved less-than-inspiring entry-level work. This tough skill-building phase can provide the foundation for a wonderful career, but in this common scenario the "follow your passion" dogma would tell you that this work is not immediately enjoyable and therefore is not your passion. We need a deeper way to discuss the value of this early period in a long working life.

We also lack a sophisticated way to discuss the role of serendipity in building a passionate pursuit. Steve Jobs, for example, in his oft-cited Stanford Commencement address (<http://www.npr.org/blogs/thetwo-way/2011/10/06/141120359/read-and-watch-steve-jobs-stanford-commencement-address>), told the crowd to not "settle" for anything less than work they loved. Jobs clearly loved building Apple, but as his biographers reveal, he stumbled into this career path at a time when he was more concerned with issues of philosophy and Eastern mysticism. This is a more complicated story than him simply following a clear preexisting passion, but it's a story we need to tell more.

These are just two examples among many of the type of nuance we could inject into our cultural conversation surrounding satisfying work — a conversation that my generation, and those that follow us, need to hear. We're ambitious and ready to work hard, but we need the right direction for investing this energy. "Follow your passion" is an inspiring slogan, but its reign as the cornerstone of modern American career advice needs to end.

We don't need slogans, we need information — concrete, evidence-based observations about how people really end up loving what they do.



Managing Yourself: Job-Hopping to the Top and Other Career Fallacies

by Monika Hamori

Climbing the hierarchy used to be a reward for loyalty. But in the 1980s, as firms stripped out layers of management, promotions became fewer and farther between. To get ahead, executives started moving from company to company. A 2009 survey by career network ExecuNet found that executives now stay with an organization for only 3.3 years, on average, before moving on. Outside job changes outnumber internal ones by about two to one.

The Real Story

As the data show, some career moves make more sense than others and the conventional wisdom doesn't necessarily hold true.

Fallacy 1: Job-Hoppers Prosper

30% of moves from one organization to another are demotions

10% of inside moves are demotions

Fallacy 2: A Move Should Be a Move Up

4% of job changes are large promotions

34% are modest promotions

Fallacy 3: Big Fish Swim in Big Ponds

8% of moves from a big name to a small name involve a step down in title

24% of moves from a small name to a big name involve a step down

Fallacy 4: Industry Switchers Are Penalized

10% started career with no industry experience

49% of CEOs at the largest firms in Europe and Asia had experience in more than one industry

17% had experience in three or more industries

But is it true that switching employers offers a fast track to the top jobs? According to my research, the answer is no. In fact, that's one of four career fallacies I identified in a study examining how managers get ahead. Understanding the reality behind job moves gives executives a leg up when planning for the future.

About the Research

These findings on career fallacies come from my eight-year research project using four sources of data:

- 1) 14,000 career histories of executives in four sectors of the financial services industry, drawn from the records of one of the largest multinational executive-search firms;
- 2) the career histories of the CEOs of the 2005 Financial Times Europe 500 and the U.S. Standard & Poor's 500 (a total of 1,001 CEOs because one firm has co-CEOs). The CEOs are located in the United States and in 21 European countries;
- 3) semistructured interviews with 45 executive search consultants at both large, multinational search firms and specialized boutiques (all U.S.-based);
- 4) interviews and online discussions with more than 20 alumni of IE Business School's executive MBA program. Interviewees were typically mid-career professionals (late thirties, early forties) living in Europe, Asia, or North or South America. Their work

experience ranged from 10 to 20 years.

Fallacy 1: Job-Hoppers Prosper

The notion that you get ahead faster by switching companies is reinforced by career counselors, who advise people to keep a constant eye on outside opportunities. But the data show that footloose executives are not more upwardly mobile than their single-company colleagues.

My analysis of the career histories of 1,001 chief executive officers who lead the largest corporations in Europe and the U.S. reveals that CEOs have worked, on average, for just three employers during their careers. And although lifetime employment is increasingly rare, a quarter of the CEOs I looked at spent an entire career with the same firm. Overall, the more years people stayed with a company, the faster they made it to the top.

CEOs are arguably a special population, so I also analyzed the job changes of 14,000 non-CEO executives to compare the outcomes of their inside and outside moves. Again, inside moves produced a considerably higher percentage and faster pace of promotions.

One likely reason that internal candidates do better is that companies know more about them; promoting an insider poses less risk than hiring somebody from the outside, no matter how extensive the CV or how detailed the reference. Executive search firms show a preference for stability as well—which is ironic, given that they're the ones in the business of shuttling professionals from job to job. One U.S. boutique firm specializing in IT evaluates candidates on two axes: stability and "performance and capability indicators." Candidates have to score well on both to be selected for interviews. A consultant at another firm told me that a short stint—less than three years or so—probably wouldn't be sufficient to produce any meaningful contribution to a firm and thus wouldn't do much to demonstrate a candidate's value. Search consultants also tend to interpret frequent moves as a sign of bad decision making, whereas long organizational tenure is rarely seen as reaching a plateau.

There are exceptions, of course. In smaller industries, for instance, where "everybody knows everybody," companies that recruit from competitors can be stigmatized as poachers. And frequent moves are unacceptable in certain countries. A midcareer Spanish manager who has worked in Japan for almost 10 years told me that leaving a job is culturally seen as treachery. Expat professionals are particularly limited in their movements because their working visas are sponsored by their employers.

Lessons for executives.

First, know that search firms are looking for résumés that demonstrate a balance between external and internal moves. One finance-search-firm recruiter I interviewed put it this way: "We like people with two or three companies. And then you look at the patterns: ideally, 10 years in one employer, two or three years in the next, but then we want to see another eight-year run." Many search firms are looking for evidence that an executive is integrating with and being rewarded by the people who work with him or her.

Second, remember that a significant proportion of executives succeed by sticking it out with one company, so consider cross-employer moves only if they'll considerably increase your employability.

Fallacy 2: A Move Should Be a Move Up

A job change, whether internal or external, doesn't necessarily mean a promotion, despite the perception that careers generally follow an upward trajectory. In reality, many changes are lateral moves, even among relatively successful executives.

In my research, the moves that constituted promotions met at least one of two criteria: They resulted in a better title with more responsibility or propelled the executive to a larger firm. Such job changes represented about 40% of the data set. Lateral moves—across division, geography, or industry—were equally common. And 20% of the job changes reflected downward moves—a lesser title or narrower scope of responsibility or a lateral move to a much smaller organization. (Smaller size implies less managerial complexity.) I found that large promotions (that is, considerable jumps in both title and employer size) were relatively uncommon—less than 5%.

While step-downs generally detract from a CV, a lateral move is by no means a career killer. It may in fact prove beneficial in the long run if done wisely. For instance, a lateral move may be justified by the prospect of a promotion in the near future.

One employee I'll call Robert, for example, recently made a lateral move, from a managerial position at one industrial maintenance company to a consultative role at another. (All names have been changed for purposes of privacy.) But the new job offers the potential for entry into the executive ranks. His new boss is the VP for strategy, and Robert works with high-potential employees on projects that involve the COO and the CEO. He is now tied to the most important work and has become visible to top management. After 18 months, the company intends to reassign the high potentials, and Robert is in line for an executive post.

Lateral moves often enhance CVs when the new company conveys brand value. Robert's new firm has networks in many growing or high-profile industries like environmental protection and oil and gas—giving Robert a valuable set of contacts and a variety of learning opportunities. A lateral move into a different industry can broaden and deepen expertise, as well.

Lessons for executives.

Fast upward leaps may not secure long-term success; often, a slower ascent that includes a mixture of lateral and upward movement is what pays off. One multinational food company with more than 60,000 employees constructs a personalized, 10-year development plan for each high potential. A strong generalist view of the business (including knowledge of finance, marketing, and how to manage people) is the determining factor in making it to the top executive ranks. Many companies share this belief, valuing employees who switch between functional tracks and general management.

To be sure, those who remain in a single function may move faster in the first part of their career, but they soon reach a ceiling because they're too specialized. One of the top executives at the food company has been an employee for almost 20 years, having held one- to three-year stints in nine countries, worked in three functional areas, and switched several times between managerial and consultative roles. Although his moves always bumped him up in the company's job grade system, not all may have seemed like advancements on paper.

Also bear in mind that a move that's technically a promotion may turn out to be a detour. Another executive, Michael, worked in the corporate legal office of a multinational tech company with more 20,000 employees; when he was offered the chance to become the head of the legal department in one of the firm's seven business units, he jumped at the opportunity. He got a title change and new managerial responsibilities, and he reported directly to the business unit CEO. But it turned out to be a dead-end job, because Michael didn't work well with the chief executive. His compensation took a severe hit: Although his base salary stayed the same, he suffered a substantial cut both in his bonus and in his stock option plans.

It's easy to be distracted by a better title, a bigger pool of direct reports, or other trappings, so when making a switch, always consider what the next move might be and to what extent the current move will help or hinder your ability to achieve longer-term goals.

Fallacy 3: Big Fish Swim in Big Ponds

Big-name companies like Goldman Sachs and Morgan Stanley often appear to "swap" professionals. They have similar cultures, so people believe they recruit from their peers in order to get high-quality employees. They're also looking for valuable insider expertise.

But the data show that when executives leave well-known companies, they more typically trade down to smaller, less-recognized firms. In my data set, 64% of executives who left an admired company—as measured by its presence on Fortune's Most Admired or a similar list—transferred to a firm not included on the list. (Of course, one reason people trade down is that there are fewer and fewer positions available at big-name companies as they climb the ranks.)

Those who leave for lesser-known or less highly regarded companies often gain in terms of title or position. In other words, they cash in on the brand value of their former employer. On the flip side, those who transfer to organizations with stronger reputations seem more willing to take a step down in position—to pay a price to acquire some brand value.

Lessons for executives.

Obviously you should do your best to join well-regarded companies as early in your career as you can. Future employers and search firms tend to equate corporate brand names with knowledge and skills. Said one consultant at a large multinational, "You can tell what competencies senior executives have just by looking at which organization they belonged to." A headhunter at a smaller, boutique firm told me: "If you know that a person is with that company, you have already made a step in the right direction in terms of qualifying them."

You should transfer to a lesser company only if the career opportunity is very attractive, beyond a jump in title and salary; otherwise it can limit your prospects down the road. Back to Michael, described earlier—he joined a big law firm after passing the bar but left to follow his boss to a niche firm that specialized in legal advice to the maritime business. He received a 50% pay increase with the move.

Soon, however, he regretted his decision, and after only two years he wanted to move again. This time he had trouble finding a suitable job and realized that his stint at the niche firm had damaged his prospects. Michael said that potential employers "looked down on" him and saw him as unable to fit in at a large firm. He knew that the training and professional development he had received in the large firm from his boss had continued in the new position. But that didn't matter to recruiters—it was the firm name that counted. He eventually found a job in the public sector, but to this day he feels that his move limited his options.

Fallacy 4: Career and Industry Switchers Are Penalized

While you'd think that changing industries or careers (a function change, for instance) would set you back, switchers don't fare worse in terms of promotions than those who stick to one field or specialize. Changing to a new area is relatively common—29% of moves take people across industries and another 23% across different segments of the same industry (going from a consumer finance company to a bank, for example).

Why would a firm hire employees from a different business? In some cases, another industry might simply offer superior human capital. One consultant at a search firm specializing in the hotel, gaming, and restaurant industries told me that 40% of his work involved recruiting from outside that world. "I am looking for companies that continuously produce high quality. If the client wants somebody who has classic marketing skills, I go to Procter & Gamble. For a very aggressive P&L background, I may go to PepsiCo."

Another executive search consultant, this one in financial services, had a similar experience: The paucity of talent in the private equity arena made hiring overly expensive. Most industry candidates came from just two major investment banks, and those executives commanded outrageous compensation. By looking at adjacent industries—pension funds, for instance, or asset management—he could produce candidates who had, as he put it, "the right wiring and intellectual capability to learn the private equity product," at a cheaper price. He could hire an executive from a global asset firm for about \$800,000 to \$1 million. The same person coming from the private equity space would have cost two or three times as much, maybe more.

Even candidates who lack industry experience may match the hiring company's needs at other levels. An executive we'll call Steven made the switch from textiles to chemicals because he had a strong track record in sales and his new company had a sales-driven culture.

When hiring companies are not sufficiently attractive to job seekers, they often need to expand their searches. In one instance, the majority owner of another sales-driven company required that all professionals—even those entering at a managerial or executive level—spend some four to six months in the sales organization. That was unappealing to many applicants; half the candidates dropped out right after their interviews because the job didn't seem to match what they saw as their strengths. So to find the best people, the company had to broaden its searches.

Lessons for executives.

Look strategically for industries where your skills represent a genuine asset. Some specializations are very difficult to find and thus worth a premium to those seeking them. A former navy pilot, Marcus, got a job as a financial analyst with SunTrust at a 50% salary increase despite having no industry expertise, because the company was looking for knowledge of the defense sector. Three years later, he headed the department.

Consider, also, a transitional job. One manager I met recently moved from a law firm, where he was marketing director, to a consultancy specializing in relocation, expatriation, and cross-cultural training. His goal was to become a consultant—a change in both industry and function—but he knew it would be almost impossible to do both at once. So he accepted the marketing-manager position at the consulting firm. He even took a pay cut, but the job allows him to learn about cross-cultural management and, he hopes, ultimately achieve his career goals.

Every career is unique, and a move that's right for you might turn out to be disastrous for your colleague, even one whose résumé and career goals are similar to yours. The fallacies I've identified are based on the experiences of real executives making real choices—but it could be that, for instance, job-hopping is the quickest way to the top in your case. What's important is to look at each move with a critical eye, putting aside conventional wisdom and other people's assumptions to make the choice that fits your own ambitions.

Monika Hamori (monika.hamori@ie.edu) is a professor of human resource management at IE Business School in Madrid. She is the coauthor, with Peter Cappelli, of "New Road to the Top" (HBR, January 2005).



Why Top Young Managers Are in a Nonstop Job Hunt

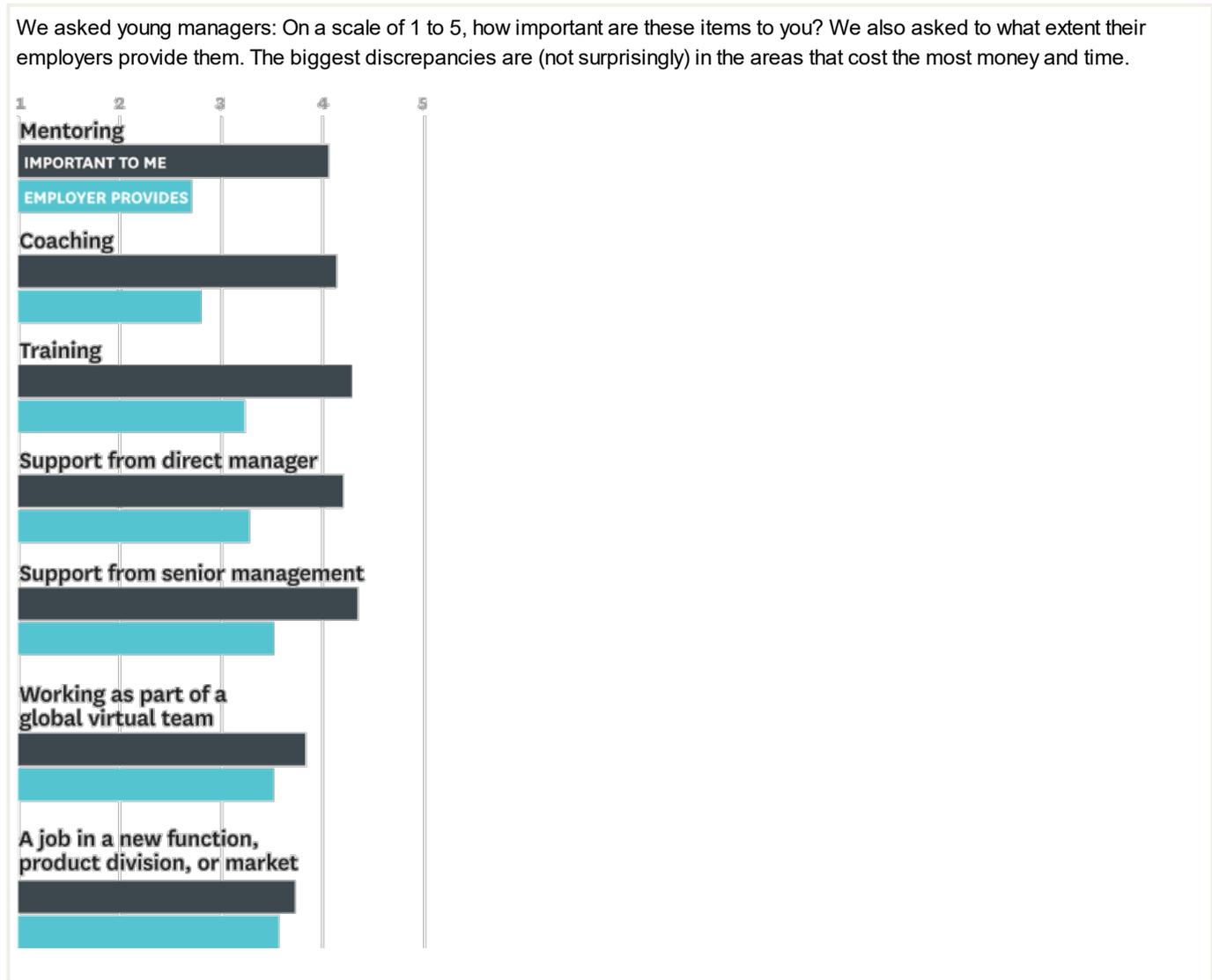
by Monika Hamori, Jie Cao, and Burak Koyuncu

You might suspect that your best young managers are looking for a better gig—and you’re probably right. Research shows that today’s most-sought-after early-career professionals are constantly networking and thinking about the next step, even if they seem fully engaged. And employee-development programs aren’t making them happy enough to stay.

We reached these conclusions after conducting face-to-face interviews and analyzing two large international databases created from online surveys of more than 1,200 employees. We found that young high achievers—30 years old, on average, and with strong academic records, degrees from elite institutions, and international internship experience—are antsy. Three-quarters sent out résumés, contacted search firms, and interviewed for jobs at least once a year during their first employment stint. Nearly 95% regularly engaged in related activities such as updating résumés and seeking information on prospective employers. They left their companies, on average, after 28 months.

And who can blame them? Comparing the peripatetic managers’ salary histories with those of peers who stayed put, we found that each change of employer created a measurable advantage in pay; in fact, a job change was the biggest single determinant of a pay increase. This represents a significant difference from the past. Job hopping has long been viewed as a shortcut to the top, but research showed that was a myth for earlier generations, who paid a price in terms of promotions and often saw their salaries suffer as well.

The Career-development Gap



Dissatisfaction with some employee-development efforts appears to fuel many early exits. We asked young managers what their employers do to help them grow in their jobs and what they’d like their employers to do, and found some large gaps. Workers

reported that companies generally satisfy their needs for on-the-job development and that they value these opportunities, which include high-visibility positions and significant increases in responsibility. But they're not getting much in the way of formal development, such as training, mentoring, and coaching—things they also value highly.

Why the disconnect? We think it's because formal training is costly and can take employees off the job for short periods of time. Employers are understandably reluctant to make big investments in workers who might not stay long. But this creates a vicious circle: Companies won't train workers because they might leave, and workers leave because they don't get training. By offering promising young managers a more balanced menu of development opportunities, employers might boost their inclination to stick around.

Monika Hamori is a professor, and Jie Cao is a doctoral student, at IE Business School, in Madrid. Burak Koyuncu is an assistant professor at Rouen Business School, in France.

6. E-business — Due: 1st class on e-business

References:

Fuller, Thomas (2020), “How Much Should the Public Know About Who Has the Coronavirus?,” *The New York Times*, March 28.

Lerner, Josh, and Jean Tirole (2002), “Some Simple Economics of Open Source,” *The Journal of Industrial Economics*, 50, 197-234.

Sang-Hun, Choe and Natasha Singer (2020), “As Coronavirus Surveillance Escalates, Personal Privacy Plummet,” *The New York Times*, March 23.

Satariano, Adam (2020), “How My Boss Monitors Me While I Work From Home,” *The New York Times*, May 6.

How Much Should the Public Know About Who Has the Coronavirus?

Amid calls for more transparency, a debate is raging among public health experts over how much data on the spread of the virus should be released.

People visit the Griffith Observatory in Los Angeles during the coronavirus pandemic. Philip Cheung for The New York Times



By **Thomas Fuller**

Published March 28, 2020 Updated March 30, 2020

SAN JOSE, Calif. — When the first case of the coronavirus in Silicon Valley was discovered in late January, health officials were faced with a barrage of questions: What city did the patient live in? Whom had he come in contact with? Which health clinic had he visited before he knew he was infected?

Dr. Sara Cody, the chief health officer for Santa Clara County, which has a population of two million across 15 cities, declined to give details.

“I can’t give the city,” she said, adding “we are not going to be giving out information about where he sought health care.”

As the coronavirus spreads across the United States the limited disclosure of data by officials would seem to be a footnote to the suffering and economic disruptions that the disease is causing.

But medical experts say that how much the public should know has become a critical question that will help determine how the United States confronts this outbreak and future ones.

Residents are clamoring to see whether the virus has been detected in their neighborhoods so they can take more steps to avoid any contact. American researchers are starved for data, unlike their colleagues in other countries who are harnessing rivers of information from their more centralized medical systems. And local politicians complain that they cannot provide basic information on the spread of the virus to their constituents.

In the perennial tug-of-war between privacy and transparency in the United States, privacy appears to be winning in the coronavirus pandemic.



A street in Palo Alto, Calif. Since the first coronavirus case in Silicon Valley was discovered in late January, health officials have declined to give details about cases to protect the privacy of patients. Glenn Chapman/Agence France-Presse — Getty Images

The bare-minimum approach to public disclosures in places like the San Francisco Bay Area is common across the United States. Armed with emergency powers in many areas, public health officers have vast discretion over what information they want, and do not want, to release to the public. Coronavirus cases in California are often listed by county, generally with very little additional information — such as gender, city of residence or age — provided.

Critics of the threadbare public reporting say it is striking that even in Silicon Valley, which is home to leading technology companies that thrive off the collection of data, residents are given very little information about the movement and dynamics of the virus.

California, which has more than 4,600 cases, is a microcosm for how inconsistent the distribution of information has been during the pandemic. Los Angeles County provides a rough age distribution of patients and breaks down the cases into more than 140 cities and communities. On Friday, for example, the county reported 21 cases in Beverly Hills, 28 in the city of Santa Monica and 49 in the neighborhood of Melrose.

Across the United States there is even less consistency. New York is listing cases by age bracket, gender and borough despite calls for more localized reporting. Connecticut lists data by town. Florida provides its residents with a wealth of data on the pandemic. The state's Department of Health has a detailed dashboard and reports showing the spread of the virus — rich with data on the cities affected, the number of people tested, the age brackets of patients, whether they are Florida residents, and the number of cases in nursing homes.

Health departments in the Bay Area make the case that releasing more granular data could heighten discrimination against certain communities where there might be clusters. The first cases in the Bay Area were among ethnic Chinese residents returning from trips to China.

“Pandemics increase paranoia and stigma,” said Dr. Rohan Radhakrishna, the deputy health officer of Contra Costa County, across the Bay from San Francisco, which provides only the total number of cases in the county on its website. “We must be extra cautious in protecting individuals and the community.”

In Santa Clara, health officials say they cannot disclose how many cases are found in each city because of the nation's strict medical privacy law, the Health Insurance Portability and Accountability Act, or HIPAA, signed by President Bill Clinton in 1996.

But that law was designed for the protection of personal data at doctors' offices and in hospitals and includes provisions for the release of otherwise protected information during emergencies.

Using the law as a justification for limiting the release of aggregate data about the coronavirus is “ridiculous,” according to Arthur L. Caplan, a professor of bioethics at the N.Y.U. School of Medicine in New York City.

Prof. Caplan is among many experts who say the coronavirus is likely to spur a reassessment of medical privacy laws. Already, the Trump Administration waived some provisions of the law this month.

“HIPAA was written for a time when there were paper charts,” Prof. Caplan said. The coronavirus, he said, “will cause us to rethink a lot of things.”

Latest Updates: Coronavirus Outbreak in the U.S.

- [W.H.O. gathers to push for global coordination amid a deepening rift between the U.S. and China.](#)
- [The first coronavirus vaccine tested in people appears to be safe and generates an immune response.](#)
- [Firms face tough choices: Today is the deadline to decide whether to return government loans.](#)

[See more updates](#)

Updated 29m ago

More live coverage: [Global](#) [Markets](#) [New York](#)

“We will also have to plan for better data exchange and testing,” he said.

The U.S. approach contrasts sharply with that of Singapore and Taiwan, whose fights against the virus have been praised as among the most effective. Both governments make public the suspected linkages of cases, anonymized by numbers. In Singapore the authorities sometimes list neighborhoods where patients lived, their workplaces and churches or mosques that they attended.

I. Glenn Cohen, an expert in bioethics at Harvard Law School, says the guiding principle during this crisis should be sharing more rather than less.

“Public health depends a lot on public trust,” he said. “If the public feels as though they are being misled or misinformed their willingness to make sacrifices — in this case social distancing — is reduced.”

“That’s a strong argument for sharing as much information as you can,” he said.

Experts also point out that it was the government’s suppression of information about the virus in China that allowed it to spread quickly before measures were taken to stem it.

Stanford University in California. The state is a microcosm for how inconsistent the distribution of information about coronavirus cases has been. Philip Pacheco/Getty Images

On Friday the health authorities in Santa Clara, which has more than 590 cases and is home to the headquarters of companies like Google and Apple, added a dashboard that charts the number of daily cases and other metrics.

But the county’s public information office says it will not publicly disclose the number of cases in each city because doing so could make individuals more easily identifiable.

In a sign of how contested the question of public disclosure is, disagreement exists even within the Santa Clara County government.

Dr. Jeffrey V. Smith, the county executive, who is both a medical doctor and a lawyer, argues that more precise geographical information about the spread does not help combat the virus because it is already widespread.

“Reporting positive tests with a census tract or a city name provides data that is not helpful,” Dr. Smith said. “In fact, such data has the risk of stigmatizing areas and regions of the country in a way that does not help.”

But David Cortese, a member of the county’s board of supervisors, says that the public has the right to know more and that a patient’s identity is unlikely to be revealed by giving a breakdown of cases by city.

“I think when people can’t get information they freak out, they think something is being hidden from them, conspiracy theories grow, suspicions grow,” he said. “I think it’s always better to be as truthful, calmly, and transparent with the public as you can be.”

As an example, Mr. Cortese says he is alarmed that health officers have not made more information public on the coronavirus-related death of a homeless man in the county. Given the medical vulnerabilities of that population, doctors and advocates of homeless people have called his office demanding to know in which encampment the man lived so that they could advise other homeless people in the area to be more vigilant. The county, which refused to disclose that information, said in a statement that health officials screened 60 members of the “specific community” and tested nine symptomatic individuals for the coronavirus. All nine tests were negative, the county said.

Mr. Cortese says it is obvious to him that more information on the spread of the pandemic should be shared.

“At the height of the information age in Silicon Valley we have stumbled and fallen flat in terms of our ability to use the tools and resources that we have to get necessary information out to the people we serve,” Mr. Cortese said.

Frustration over the dearth of data also extends to epidemiologists trying to understand the dynamics of the spread of the virus.

Joseph Lewnard, a professor of epidemiology at the University of California, Berkeley School of Public Health, says researchers are hamstrung in the United States by the lack of specific data on testing and on the symptoms patients show.

To make up for the lack of public data, researchers are scraping information on cases from news outlets and other media accounts, he said. They are mainly relying on data from South Korea, China and Italy to try to predict the spread of the virus.

“We are right now learning and trying to project what is happening here in the United States almost entirely based on observations from these other countries,” Prof. Lewnard said.

Moritz Kraemer, a scholar at Oxford University who is leading a team of researchers in mapping the global spread of the coronavirus, says China’s data “provided incredible detail,” including a patient’s age, sex, travel history and history of chronic disease, as well as where the case was reported, and the dates of the onset of symptoms, hospitalization and confirmation of infection.

The United States, he said, “has been slow in collecting data in a systematic way.”

Dr. C. Jason Wang, a researcher at Stanford University, who has studied how Taiwan handled the coronavirus outbreak, says some of the measures taken in Taiwan would most likely not be accepted in the United States given privacy concerns. The government, for example, merged the airport immigration database with the national medical database so that doctors could immediately see if a patient had traveled out of the country.

But Dr. Wang says the proactive approach that Taiwan took to the virus, including aggressive tracing of cases, has helped keep the total number of confirmed infections — 283 on Saturday — much lower than experts initially expected. By comparison, the borough of Queens in New York City, with one-tenth the population of Taiwan, has 10,000 cases.

Some of the information being released to the public in Taiwan and Singapore would most likely be uncontroversial in the United States, he said. Taiwanese authorities, for example, have pointed out linkages between anonymized cases, including family clusters, in an effort to warn the public how easily the virus is transmitted within households.

Prof. Caplan of the N.Y.U. School of Medicine says it is paradoxical that the United States is providing less precise information to its citizens on the outbreak than Singapore, which puts limits on the spread of information through internet controls.

“Here we expect to get information so we have our choices and we make our decisions,” he said. “Our notion is information is the oxygen for democracy. Wouldn’t we want to receive more information than them?”

SOME SIMPLE ECONOMICS OF OPEN SOURCE*

JOSH LERNER[†] AND JEAN TIROLE[‡]

There has been a recent surge of interest in open source software development, which involves developers at many different locations and organizations sharing code to develop and refine programs. To an economist, the behavior of individual programmers and commercial companies engaged in open source projects is initially startling. This paper makes a preliminary exploration of the economics of open source software. We highlight the extent to which labor economics, especially the literature on ‘career concerns’, and industrial organization theory can explain many of these projects’ features. We conclude by listing interesting research questions related to open source software.

I. INTRODUCTION

IN RECENT YEARS, there has been a surge of interest in open source software development. Interest in this process, which involves software developers at many different locations and organizations sharing code to develop and refine software programs, has been spurred by three factors:

- *The rapid diffusion of open source software.* A number of open source products, such as the Apache web server, dominate their product categories. In the personal computer operating system market, International Data Corporation estimates that the open source program Linux has between seven to twenty-one million users worldwide, with a 200% annual growth rate. Many observers believe it represents a leading potential challenger to Microsoft Windows in this important market segment.

*The assistance of the Harvard Business School’s California Research Center, and Chris Darwall in particular, was instrumental in the development of the case studies and is gratefully acknowledged. We also thank a number of practitioners—especially Eric Allman, Mike Balma, Brian Behlendorf, Keith Bostic, Tim O’Reilly, and Ben Passarelli—for their willingness to generously spend time discussing the open source movement. George Baker, Jacques Crémer, Rob Merges, Bernie Reddy, Pierre Régibeau, Bernard Salanié, many open source participants, seminar participants at the American Economics Association annual meetings, European Economic Association Bolzano meetings, and Harvard, and three anonymous referees provided helpful comments. Harvard Business School’s Division of Research provided financial support. The Institut D’Economie Industrielle receives research grants from a number of corporate sponsors, including French Telecom and the Microsoft Corporation. All opinions and errors, however, remain our own.

[†] Authors’ affiliations: Harvard Business School, Morgan Hall, Room 395, Boston, Massachusetts, 02163, USA.
email: jlerner@hbs.edu

[‡] Institut d’Economie Industrielle, Manufacture des Tabacs, Bureau MF529 - Bat. F 21 Allée de Brienne, 31000 Toulouse, France.
email: tirole@cict.fr

- *The significant capital investments in open source projects.* Over the past two years, numerous major corporations, including Hewlett Packard, IBM, and Sun, have launched projects to develop and use open source software. Meanwhile, a number of companies specializing in commercializing Linux, such as Red Hat and VA Linux, have completed initial public offerings, and other open source companies such as Cobalt Networks, Collab.Net, Scriptics, and Sendmail have received venture capital financing.
- *The new organization structure.* The collaborative nature of open source software development has been hailed in the business and technical press as an important organizational innovation.

Yet to an economist, the behavior of individual programmers and commercial companies engaged in open source processes is startling. Consider these quotations by two leaders of the open source community:

The idea that the proprietary software social system—the system that says you are not allowed to share or change software—is unsocial, that it is unethical, that it is simply wrong may come as a surprise to some people. But what else can we say about a system based on dividing the public and keeping users helpless? [Stallman, 1999]

The ‘utility function’ Linux hackers is maximizing is not classically economic, but is the intangible of their own ego satisfaction and reputation among other hackers. [Parenthetical comment deleted] Voluntary cultures that work this way are actually not uncommon; one other in which I have long participated is science fiction fandom, which unlike hackerdom explicitly recognizes ‘egoboo’ (the enhancement of one’s reputation among other fans) [Raymond, 1999b].

It is not initially clear how these claims relate to the traditional view of the innovative process in the economics literature. Why should thousands of top-notch programmers contribute freely to the provision of a public good? Any explanation based on altruism¹ only goes so far. While users in less developed countries undoubtedly benefit from access to free software, many beneficiaries are well-to-do individuals or Fortune 500 companies. Furthermore, altruism has not played a major role in other industries, so it would have to be explained why individuals in the software industry are more altruistic than others.

This paper seeks to make a preliminary exploration of the economics of open source software. Reflecting the early stage of the field’s develop-

¹ The media like to portray the open source community as wanting to help mankind, as it makes a good story. Many open source advocates put limited emphasis on this explanation.

ment, we do not seek to develop new theoretical frameworks or to statistically analyze large samples. Rather, we focus on four ‘mini-cases’ of particular projects: Apache, Linux, Perl, and Sendmail.² We seek to draw some initial conclusions about the key economic patterns that underlie the open source development of software. We find that much can be explained by reference to economic frameworks. We highlight the extent to which labor economics, in particular the literature on ‘career concerns’, and industrial organization theory can explain many of the features of open source projects.

At the same time, we acknowledge that aspects of the future of open source development process remain somewhat difficult to predict with ‘off-the-shelf’ economic models. In the final section of this paper, we highlight a number of puzzles that the movement poses. It is our hope that this paper will have itself an ‘open source’ nature: that it will stimulate research by other economic researchers as well.

Finally, it is important to acknowledge the relationship with the earlier literature on technological innovation and scientific discovery. The open source development process is somewhat reminiscent of the type of ‘user-driven innovation’ seen in many other industries. Among other examples, Rosenberg’s [1976] studies of the machine tool industry and von Hippel’s [1988] of scientific instruments have highlighted the role that sophisticated users can play in accelerating technological progress. In many instances, solutions developed by particular users for individual problems have become more general solutions for wide classes of users. Similarly, user groups have played an important role in stimulating innovation in other settings: certainly, this has been the case from the earliest days of the computer industry [e.g., Caminer, *et al.*, 1996].

A second strand of related literature examines the adoption of the scientific institutions (‘open science’, in Dasgupta and David’s [1994] terminology) within for-profit organizations. Henderson and Cockburn [1994] and Gambardella [1995] have highlighted that the explosion of knowledge in biology and biochemistry in the 1970s triggered changes in the management of R&D in major pharmaceutical firms. In particular, a number of firms encouraged researchers to pursue basic research, in addition to the applied projects that typically characterized these organizations. These firms that did so enjoyed substantially higher R&D productivity than their peers, apparently because the research scientists allowed them to more accurately identify promising scientific developments (in other words, their ‘absorptive capacity’ was enhanced) and because the interaction with cutting-edge research made these firms more attractive to top scientists. At the same time, the encouragement of ‘open

²These are summarized in Darwall and Lerner [2000].

science' processes has not been painless. Cockburn, Henderson, and Stern [1999] highlight the extent to which encouraging employees to pursue both basic and applied research led to substantial challenges in designing incentive schemes, because of the very different outputs of each activity and means through which performance is measured.³

But as we shall argue below, certain aspects of the open source process—especially the extent to which contributors' work is recognized and rewarded—are quite distinct from earlier settings. This study focuses on understanding this contemporaneous phenomenon, rather than seeking to make a general evaluation of the various cooperative schemes employed over time.

II. THE NATURE OF OPEN SOURCE SOFTWARE

While media attention to the phenomenon of open source software has been recent, the basic behaviors are much older in their origins. There has long been a tradition of sharing and cooperation in software development. But in recent years, both the scale and formalization of the activity have expanded dramatically with the widespread diffusion of the Internet.⁴ In the discussion below, we will highlight three distinct eras of cooperative software development.

II(i). *The First Era: Early 1960s to the Early 1980s*

Many of the key aspects of the computer operating systems and the Internet were developed in academic settings such as Berkeley and MIT during the 1960s and 1970s, as well as in central corporate research facilities where researchers had a great deal of autonomy (such as Bell Labs and Xerox's Palo Alto Research Center). In these years, the sharing by programmers in different organizations of basic operating code of computer programs—the source code—was commonplace.⁵

³ It should be noted that these changes are far from universal. In particular, many information technology and manufacturing firms appear to be moving to less of an emphasis on basic science in their research facilities (for a discussion, see Rosenbloom and Spencer [1996]).

⁴ This history is of necessity highly abbreviated and we do not offer a complete explanation of the origins of open source software. For more detailed treatments, see Browne [1999], DiBona, Ockman, and Stone [1999], Gomulkiewicz [1999], Levy [1984], Raymond [1999a], and Wayner [2000].

⁵ Programmers write source code using languages such as Basic, C, and Java. By way of contrast, most commercial software vendors only provide users with object, or binary, code. This is the sequence of 0s and 1s that directly communicates with the computer, but which is difficult for programmers to interpret or modify. When the source code is made available to other firms by commercial developers, it is typically licensed under very restrictive conditions.

Many of the cooperative development efforts in the 1970s focused on the development of an operating system that could run on multiple computer platforms. The most successful examples, such as Unix and the C language used for developing Unix applications, were originally developed at AT&T's Bell Laboratories. The software was then installed across institutions, being transferred freely or for a nominal charge. Many of the sites where the software was installed made further innovations, which were in turn shared with others. The process of sharing code was greatly accelerated with the diffusion of Usenet, a computer network begun in 1979 to link together the Unix programming community. As the number of sites grew rapidly (e.g., from 3 in 1979 to 400 in 1982), the ability of programmers in university and corporate settings to rapidly share technologies was considerably enhanced.

These cooperative software development projects were undertaken on a highly informal basis. Typically no effort to delineate property rights or to restrict reuse of the software were made. This informality proved to be problematic in the early 1980s, when AT&T began enforcing its (purported) intellectual property rights related to Unix.

II(ii). *The Second Era: Early 1980s to the Early 1990s*

In response to these threats of litigation, the first efforts to formalize the ground rules behind the cooperative software development process emerged. This ushered in the second era of cooperative software development. The critical institution during this period was the Free Software Foundation, begun by Richard Stallman of the MIT Artificial Intelligence Laboratory in 1983. The foundation sought to develop and disseminate a wide variety of software without cost.

One important innovation introduced by the Free Software Foundation was a formal licensing procedure that aimed to preclude the assertion of patent rights concerning cooperatively developed software (as many believed that AT&T had done in the case of Unix). In exchange for being able to modify and distribute the GNU software (as it was known), software developers had to agree to make the source code freely available (or at a nominal cost). As part of the General Public License (GPL, also known as 'copylefting'), the user had to also agree not to impose licensing restrictions on others. Furthermore, all enhancements to the code—and even code that intermingled the cooperatively developed software with that developed separately—had to be licensed on the same terms. It is these contractual terms that distinguish open source software from shareware (where the binary files but not the underlying source code are made freely available, possibly for a trial period only) and public-domain

software (where no restrictions are placed on subsequent users of the source code).⁶

This project, as well as contemporaneous efforts, also developed a number of important organizational features. In particular, these projects employed a model where contributions from many developers were accepted (and frequently publicly disseminated or posted). The official version of the program, however, was managed or controlled by a smaller subset of individuals closely involved with the project, or in some cases, an individual leader. In some cases, the project's founder (or his designated successor) served as the leader; in others, leadership rotated between various key contributors.

II(iii). *The Third Era: Early 1990s to Today*

The widespread diffusion of Internet access in the early 1990s led to a dramatic acceleration of open source activity. The volume of contributions and diversity of contributors expanded sharply, and numerous new open source projects emerged, most notably Linux (a UNIX operating system developed by Linus Torvalds in 1991). As discussed in detail below, interactions between commercial companies and the open source community also became commonplace in the 1990s.

Another innovation during this period was the proliferation of alternative approaches to licensing cooperatively developed software. During the 1980s, the GPL was the dominant licensing arrangement for cooperatively developed software. This changed considerably during the 1990s. In particular, Debian, an organization set up to disseminate Linux, developed the 'Debian Free Software Guidelines' in 1995. These guidelines allowed licensees greater flexibility in using the program, including the right to bundle the cooperatively developed software with proprietary code. These provisions were adopted in early 1997 by a number of individuals involved in cooperative software development, and were subsequently known as the 'Open Source Definition'. As the authors explained:

License Must Not Contaminate Other Software. The license must not place restrictions on other software that is distributed along with the licensed software. For example, the license must not insist that

⁶It should be noted, however, that some projects, such as the Berkeley Software Distribution (BSD) effort, did take alternative approaches during the 1980s. The BSD license also allows anyone to freely copy and modify the source code (as long as credit was given to the University of California at Berkeley for the software developed there, a requirement no longer in place). It is much less constraining than the GPL: anyone can modify the program and redistribute it for a fee without making the source code freely available. In this way, it is a continuation of the university-based tradition of the 1960s and 1970s.

all other programs distributed on the same medium must be open-source software. Rationale: Distributors of open-source software have the right to make their own choices about their own software [Open Source Initiative, 1999].

These new guidelines did not require open source projects to be 'viral': they need not 'infect' all code that was compiled with the software with the requirement that it be covered under the license agreement as well. At the same time, they also accommodated more restrictive licenses, such as the General Public License.

The past few years have seen unprecedented growth of open source software. At the same time, the movement has faced a number of challenges. We will highlight two of these here: the 'forking' of projects (the development of competing variations) and the development of products for high-end users.

One issue that has emerged in a number of open source projects is the potential for programs splintering into various variants. In some cases, passionate disputes over product design have led to the splintering of open source projects into different variants. Examples of such splintering are the Berkeley Unix program and Sendmail during the late 1980s.

Another challenge has been the apparently lesser emphasis on documentation and support, user interfaces,⁷ and backward compatibility found in at least some open source projects. The relative technological features of software developed in open source and traditional environments are a matter of passionate discussion. Some members of the community believe that this production method dominates traditional software development in all respects. But many open source advocates argue that open source software tends to be geared to the more sophisticated users.⁸ This point is made colorfully by one open source developer:

[I]n every release cycle Microsoft always listens to its *most ignorant customers*. This is the key to dumbing down each release cycle of software for further assaulting the non personal-computing population. Linux and OS/2 developers, on the other hand, tend to listen to their *smartest* customers . . . The good that Microsoft does in bringing computers to non-users is outdone by the curse that they bring on experienced users [Nadeau, 1999].

⁷Two main open source projects (GNOME and KDE) are meant to remedy Linux's limitations on desktop computers (by developing mouse and windows interfaces).

⁸For example, Torvalds [1999] argues that the Linux model works best with developer-type software. Ghosh [1999] views the open source process as a large repeated game process of give-and-take among developer-users (the 'cooking pot' model).

Certainly, the greatest diffusion of open source projects appears to be in settings where the end users are sophisticated, such as the Apache server installed by systems administrators. In these cases, users are apparently more willing to tolerate the lack of detailed documentation or easy-to-understand user interfaces in exchange for the cost savings and the possibility of modifying the source code themselves. In several projects, such as Sendmail, project administrators chose to abandon backward compatibility in the interests of preserving program simplicity.⁹ One of the rationales for this decision was that administrators using the Sendmail system were responsive to announcements that these changes would be taking place, and rapidly upgraded their systems. In a number of commercial software projects, it has been noted, these types of rapid responses are not as common. Once again, this reflects the greater sophistication and awareness of the users of open source software.

The debate about the ability of open source software to accommodate high-end users' needs has direct implications for the choice of license. The recent popularity of more liberal licenses and the concomitant decline of the GNU license are related to the rise in the 'pragmatists' influence. These individuals believe that allowing proprietary code and for-profit activities in segments that would otherwise be poorly served by the open-source community will provide the movement with its best chance for success.

II(iv). *Who Contributes?*

Computer system administrators, database administrators, computer programmers, and other computer scientists and engineers represented about 2.1 million jobs in the United States in 1998. (Unless otherwise noted, the information in this paragraph is from U.S. Department of Labor [2000].) A large number of these workers—estimated at between five and ten percent—are either self-employed or retained on a project-by-project basis by employers. Computer-related positions are projected by the federal government to be among the fastest-growing professions in the next decade.

The distribution of contributors to open source projects appears to be quite skewed. This is highlighted by an analysis of 25 million lines of open source code, constituting 3149 distinct projects [Ghosh and Prakash, 2000]. The distribution of contributions is shown in Figure 1. More than three-quarters of the nearly 13 thousand contributors made only one contribution; only one in twenty-five had more than five contributions. Yet the top decile of contributors accounted for fully 72% of the code

⁹To be certain, backward compatibility efforts may sometimes be exerted by status-seeking open source programmers. For example, Linux has been made to run on Atari machines, a pure bravado effort since no one uses Ataris anymore.

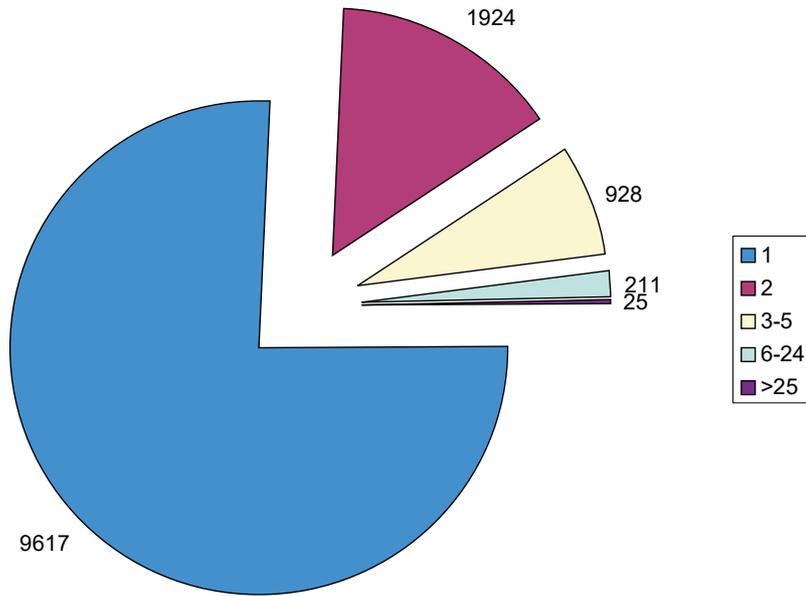


Figure 1
Distribution of Contributions Made, by Number of Participants

contributed to the open source projects, and the top two deciles for 81% (see Figure 2). This distribution would be even more skewed if those who simply reported errors, or ‘bugs’, were considered: for every individual who contributes code, five will simply report errors [Valloppillil, 1998]. To

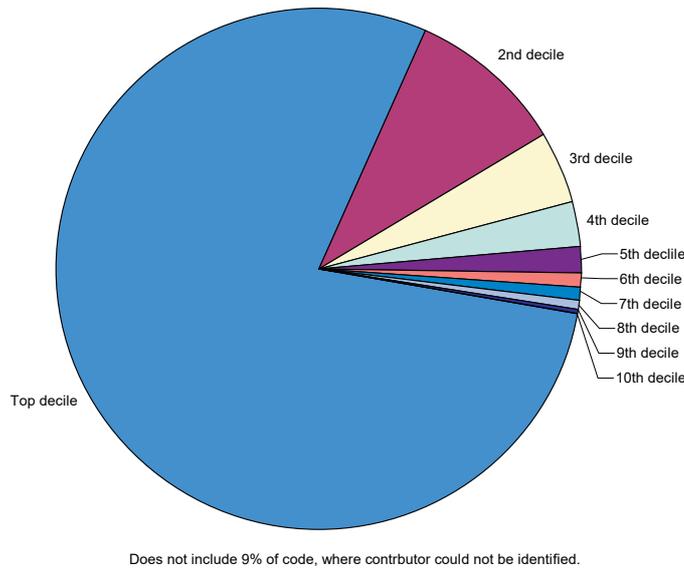


Figure 2
Distribution of Code Contributed, by Decile

what extent this distribution is unique to open source software is unclear: the same skewness of output is also observed among programmers employed in commercial software development facilities [e.g., see Brooks, 1975, and Cusumano, 1991], but it is unclear whether these distributions are similar in their properties.

The overall picture that we drew from our interviews and from the responses we received in reaction to the first draft of the paper is that the open source process is quite élitist. Important contributors are few and ascend to the ‘core group’ status, the ultimate recognition by one’s peers. The élitist view is also supported by Mockus, *et al’s* [1999] study of contributions to Apache. For Apache, the (core) ‘developers mailing list’ is considered as the key list of problems to be solved, while other lists play a smaller role. The top 15 developers contribute 83% to 91% of changes (problem reports by way of contrast offer a much less elitist pattern).

Some evidence consistent with the suggestion that contributions to open source projects are being driven by signaling concerns can be found in the analysis of contributors to a long-standing archive of Linux postings maintained at the University of North Carolina by Dempsey, *et al.* [1999]. These authors examine the suffix of the contributors’ e-mail addresses. While the location of many contributors cannot be precisely identified (for instance, contributors at ‘.com’ entities may be located anywhere in the world), the results are nonetheless suggestive. As Figure 3 depicts, 12% of the contributors are from entities with a suffix ‘.edu’ (typically, U.S.

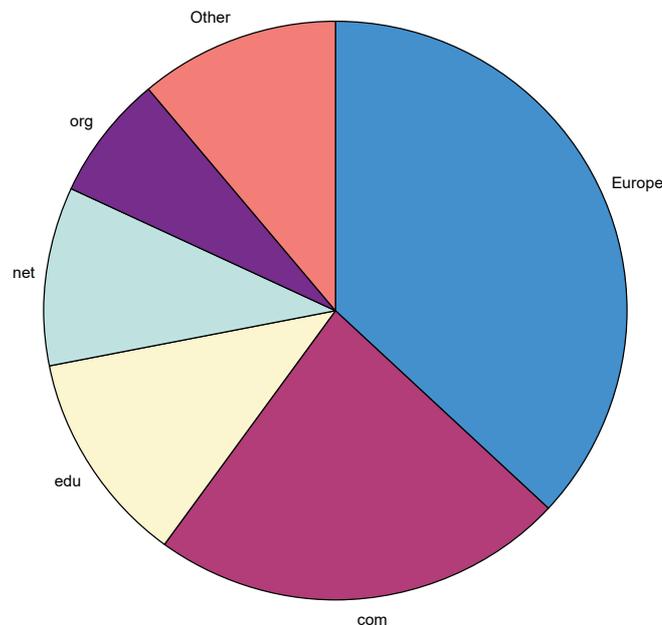


Figure 3
Suffix of Linux Contributors

educational institutions), 7% from '.orgs' (traditionally reserved from U.S. non-profits), fully 37% are from Europe (e.g., with suffixes such as '.de' and '.uk'), and 11% have other suffixes, many of which represent other foreign countries. This suggests that many of the contributions are coming from individuals outside the major software centers.

III. THE ORIGINS OF FOUR PROGRAMS

Each of the four case studies was developed through the review of printed materials and interviews (as well as those posted on various web sites) and face-to-face meetings with one or more key participants in the development effort. In addition, we held a number of conversations with knowledgeable observers of the open source movement. In Sections IV and V, we will frequently draw on examples from the four cases. Nonetheless, we felt it would be helpful to first provide a brief overview of the development projects.

III(i). *Apache*

The development of Apache began in 1994. Brian Behlendorf, then 21, had the responsibility for operating one of the first commercial Internet servers in the country, that powering *Wired* magazine's HotWired web site. This server, like most others in the country, was at the time running the Unix-based software written at the National Center for Supercomputer Applications (NCSA) at the University of Illinois. (The only competitive product at the time was the server developed at the joint European particle physics research facility, CERN.) The NCSA had distributed its source code freely and had a development group actively involved in refining the code in consultation with the pioneering users. As Behlendorf and other users wrote emendations, or 'patches', for the NCSA server, they would post them as well to mailing lists of individuals interested in Internet technology.

Behlendorf and a number of other users, however, encountered increasing frustrations in getting the NCSA staff to respond to their suggestions. (During this time, a number of the NCSA staff had departed to begin Netscape, and the University was in the process of negotiating a series of licenses of its software with commercial companies.) As a result, he and six other pioneering developers decided to establish a mailing list to collect and integrate the patches to the NCSA server software. They agreed that the process would be a collegial one. While a large number of individuals would be able to suggest changes, only a smaller set would be able to actually make changes to the physical code. In August 1995, the group released Apache 0.8, which represented a substantial departure from earlier approaches. A particular area of revision was the Application

Program Interface (API), which allowed the development of Apache features to be very ‘modular’. This step enabled programmers to make contributions to particular areas without affecting other aspects of the programs.

At the time that Apache was introduced, there was little in the way of competitive products: in fact, the absence of a good commercial alternative was a powerful motivation for the launching of the project. A variety of commercial software vendors, most notably Microsoft and Netscape, have subsequently targeted server software. Despite this competition, Apache has retained its dominant position. The November 2000 Netcraft survey [2000] of nearly 24 million Internet domains found that Apache had a dominant position: 59.7% of the sites used this server software. The closest competitors, Microsoft’s IIS and Netscape’s Enterprise software, were at 20.2% and 6.7% respectively.¹⁰

In 1999, the Apache Software Foundation was established to oversee the development and diffusion of the program. The current status of Apache, as well as the other open source projects that we focused on, is summarized in Table I.

III(ii). *Linux*

Linux, an amalgam of ‘Linus’ and ‘Unix’, was created by Linus Torvalds in 1991. Unlike the other case studies considered here, Torvalds was motivated to pursue this project by intellectual curiosity, rather than by a pressing practical need. A 21-year-old graduate student, he sought to build the ‘kernel’—or core element—of a truly open source operating system.

Torvalds based his system on Minix, a public domain Unix system for personal computers. After approximately six months of development, a friend allowed him to post the operating system on a university server. He began encouraging contributions in a series of postings to on-line bulletin boards, such as one that posed the question ‘are you without a project and just dying to cut your teeth on an [operating system] you can try to modify to your needs?’

Torvalds initially distributed Linux under a licensing agreement that restricted any payment for the program, as well as requiring that all programs distributed or used with Linux be freely available. After half a year, however, he relaxed these restrictions. The number of users grew rapidly, from about one hundred after one year to half-a-million in 1994.

¹⁰A complication is introduced by the fact that firewall-protected servers may be quite different in nature. For instance, a survey of both protected and unprotected servers in the summer of 1996 by Zoma Research concluded that open source server programs, including Apache, accounted for only 7% of all installations, far less than the contemporaneous Netcraft estimate.

TABLE I
THE OPEN SOURCE PROGRAMS STUDIED

Program	Apache	Perl	Sendmail
Nature of program:	World wide web (HTTP) server	System administration and programming language	Internet mail transfer agent
Year of introduction:	1994	1987	1979 (predecessor program)
Governing body:	Apache Software Foundation	Selected programmers (among the 'perl-5-porters' (formerly, The Perl Institute)	Sendmail Consortium
Competitors:	Internet Information Server (Microsoft) Various servers (Netscape)	Java (Sun) Python (open source program) Visual Basic, ActiveX (Microsoft)	Exchange (Microsoft) IMail (Ipswich) Post.Office (Software.com)
Market penetration:	55% (September 1999) (of publicly observable sites only)	Estimated to have 1 million users	Handles ~80% of Internet e-mail traffic
Web site:	www.apache.org	www.perl.org	www.sendmail.com

From the beginning, Torvalds retained clear leadership of the Linux project. He rapidly moved to writing less code and coordinating the software development project, assessing contributions and arbitrating disputes. Over time, a set of lieutenants have assumed responsibility for most of the decision-making, but Torvalds still retains authority for making the ultimate decisions. While employed at California-based semiconductor manufacturer Transmeta, Torvalds continues to devote about half his time to the Linux project.

While the origin of Linux was largely driven by intellectual curiosity on the part of Torvalds and his peers, the program has evolved into one that represents a significant competitor to Microsoft's Windows operating system. While the number of Linux users is difficult to determine because of the numerous channels through which the program is distributed, estimates range from 7 to 16 million users worldwide.

Reflecting its widespread diffusion, Linux has attracted a large share of the commercial investment in open source projects. A number of firms dedicated to supporting Linux have been established: pioneers included VA Linux, founded in 1993, and Red Hat, established in 1995. These commercial firms sell Linux software 'packages', which are often far easier to install and operate than free versions available, provide technical support to end users and computer resellers, and sell complementary proprietary products. In addition, a number of established computer hardware and software firms have made extensive investments in Linux development.

III(iii). *Perl*

Perl, or the Practical Extraction and Reporting Language, was created by Larry Wall in 1987. Wall, a programmer with Burroughs (a computer mainframe manufacturer now part of Unisys) had already written a number of widely adopted software programs. These included a program for reading postings on on-line newsgroups and a program that enabled users to readily update old source code with new patches.

The specific genesis of Perl was the large number of repetitive system administration tasks that Wall was asked to undertake while at Burroughs. In particular, Wall was required to synchronize and generate reports on two Unix-based computers as part of a project that Burroughs was undertaking for the U.S. National Security Agency. He realized that there was a need for a program language that was somewhere between the Unix shell language and the C language (suitable for developing complex programming applications). The Perl language sought to enable programmers rapidly to undertake a wide variety of tasks, particularly relating to system administration. The program was first introduced in 1987 via the Internet. It has become widely accepted as a language for

developing scripts for Apache web servers, and is incorporated in a number of other programs.

Perl is administered on a rotating basis: the ten to twenty programmers (the number fluctuates over time) who have been most actively involved in the program take turns managing different aspects of the project. Wall himself has joined the staff of O'Reilly & Associates, a publisher specializing in manuals documenting open source programs. While he is no longer actively contributing to the programming, he remains active in managing the project.

As in the case of Apache, Perl's success has attracted competition from commercial developers. In particular, Sun's Java and Microsoft's ActiveX, both of which were introduced well after the diffusion of Perl, incorporate many of the same features. Rough estimates suggest that the number of Perl users is about one million. Some observers believe (see, for instance, the conversations archived at <http://www.mail-archive.com/advocacy%40perl.org>) that the growth usage of Perl has largely stabilized, and that many of the new users are turning to Java. As is often the case in this sector, confirming these claims is exceedingly difficult.

Two efforts to establish a Perl-related foundation have foundered. For instance, the Perl Institute had been intended to ensure that less glamorous tasks, such as documentation, were undertaken, in order to enhance the long-run growth of Perl. The failure of these efforts, however, may have reflected more about the specifics of the individual personalities involved than the prospects of the program itself.

III(iv). *Sendmail*

Sendmail was originally developed in the late 1970s by Eric Allman, a graduate student in computer science at the University of California at Berkeley. As part of his responsibilities, Allman worked on a variety of software development and system administration tasks at Berkeley.

One of the major challenges that Allman faced was the incompatibility of the two major computer networks on campus. The approximately one dozen Unix-based computers had been originally connected through 'BerkNet', a locally developed program that provided continuous inter-connection. These computers, in turn, connected to those on other campuses through telephone lines, using the UUCP protocol (Unix-to-Unix Copy Protocol). Finally, the Arpanet, the direct predecessor of the Internet, was introduced on the Berkeley campus around this time. Each of the networks used a different communications protocol: for instance, each person had multiple e-mail addresses, depending on the network from which the message was sent. To cope with this problem, Allman developed in 1979 a program called 'Delivermail', which provided a way to greatly

simplify the addressing problem. In an emendated form that allowed it to address a large number of domains, it was released two years later as ‘Sendmail’.

Sendmail was soon adopted as the standard method of routing e-mail on the Arpanet. As the network grew, however, its limitations became increasingly apparent. A variety of enhanced versions of Sendmail were released in the 1980s and early 1990s which were incompatible with each other—in the argot of the open source community, the development of the program ‘forked’. In 1993, Allman, who had returned to working at Berkeley after being employed at a number of software firms, undertook a wholesale rewrite of Sendmail. The development was sufficiently successful that the incompatible versions were largely abandoned in favor of the new version. While a variety of competitive products had appeared, such as Software.com’s Post Office, Microsoft’s Exchange, and Ipswitch’s Imail, the open source program appeared to have a dominant competitive position. Observers have attributed this to the presence of an installed base of users and the ease of customizing the program. The program was estimated to handle about 75% of all Internet e-mail traffic in 2000.

In 1997, Allman established Sendmail, Inc. The company, which has been financed by a leading venture capital group, Benchmark Capital, is seeking to sell Sendmail-related software enhancements (such as more user-friendly interfaces) and services. At the same time, the company seeks to encourage the continuing development of the software on an open source basis. For instance, Sendmail, Inc. employs two engineers who work almost full time on contributions to the open source program, which is run by the non-profit Sendmail Consortium.

IV. WHAT DOES ECONOMIC THEORY TELL US ABOUT OPEN SOURCE?

This section and the next use economic theory to shed light on the three key questions: Why do people participate?¹¹ Why are there open source projects in the first place? And how do commercial vendors react to the open source movement?

IV(i). *What Motivates Programmers?*

A programmer participates in a project, whether commercial or open source, only if she derives a net benefit (broadly defined) from engaging in

¹¹We focus primarily on programmers’ contributions to code. A related field of study concerns field support, which is usually also provided free of charge in the open source community. Lakhani and von Hippel [2000] provide empirical evidence for field support in the Apache project. They show that providers of help often gain learning for themselves, and that the cost of delivering help is therefore usually low.

the activity. The net benefit is equal to the immediate payoff (current benefit minus current cost) plus the delayed payoff (delayed benefit minus delayed cost).

A programmer working on an open source software development project incurs a variety of benefits and costs. The programmer incurs an opportunity cost of her time. While she is working on this project, she is unable to engage in another programming activity. This opportunity cost exists at the extensive and intensive margins. First, a programmer who would work as an independent on open source projects would forgo the monetary compensation she would receive if she were working for a commercial firm or a university. Second, and more to the point, for a programmer with an affiliation with a commercial company, a university or research lab, the opportunity cost is the cost of not focusing on her primary mission. For example, the academic's research output may sag, and the student's progress towards a degree slow down; these involve delayed costs. The size of this opportunity cost of not focusing on the primary mission of course depends on the extent of monitoring by the employer and more generally, the pressure on the job.

Two immediate benefits may counter this cost. First, the programmer, when fixing a bug or customizing an open source program, may actually improve rather than reduce her performance in the mission endowed upon her by her employer. This is particularly relevant for system administrators looking for specific solutions for their company. Second, the programmer compares how enjoyable the mission set by the employer and the open source alternative are. A 'cool' open source project may be more fun than a routine task.

The delayed reward covers two distinct, although hard-to-distinguish, incentives. The *career concern incentive* refers to future job offers, shares in commercial open source-based companies,¹² or future access to the venture capital market.¹³ The *ego gratification incentive* stems from a desire for peer recognition. Probably most programmers respond to both

¹² Linus Torvalds and others have been awarded shares in Linux-based companies that went public. Most certainly, these rewards were unexpected and did not affect the motivation of open source programmers. If this practice becomes 'institutionalized', such rewards will in the future be expected and therefore impact the motivation of open source leaders. More generally, leaders of open source movements may initially not have been motivated by ego gratification and career concerns. Like Behlendorf, Wall, and Allman, the 'bug fixing' motivation may have originally been paramount. The private benefits of leadership may have grown in importance as the sector matured.

¹³ Success at a commercial software firm is likely to be a function of many attributes. Some of these (e.g., programming talent) can be signaled through participation in open source projects. Other important attributes, however, are not readily signaled through these projects. For instance, commercial projects employing a top-down architecture require that programmers work effectively in teams, while many open source projects are initiated by relatively modest pieces of code, small enough to be written by a single individual.

incentives. There are some differences between the two. The programmer mainly preoccupied by peer recognition may shun future monetary rewards, and may also want to signal her talent to a slightly different audience than those motivated by career concerns. From an economic perspective, however, the incentives are similar in most respects. We will group the career concern incentive and the ego gratification incentive under a single heading: the *signaling incentive*.

Economic theory [e.g., Holmström, 1999] suggests that this signaling incentive is stronger,

- (i) the more visible the performance to the relevant audience (peers, labor market, venture capital community),
- (ii) the higher the impact of effort on performance, and
- (iii) the more informative the performance about talent.

The first condition gives rise to what economists call ‘strategic complementarities’. To have an ‘audience’, programmers will want to work on software projects that will attract a large number of other programmers. This suggests the possibility of multiple equilibria. The same project may attract few programmers because programmers expect that other programmers will not be interested; or it may flourish as programmers (rationally) have faith in the project.

The same point applies to forking in a given open source project. Open source processes are in this respect quite similar to academic research. The latter is well known to exhibit fads: see the many historical examples of simultaneous discoveries discussed by Merton [1973]. Fields are completely neglected for years, while others with apparently no superior intrinsic interest attract large numbers of researchers. Fads in academia are frowned upon for their inefficient impact on the allocation of research. It should not be ignored, however, that fads also have benefits. A fad can create a strong signaling incentive: researchers working in a popular area may be highly motivated to produce a high-quality work, since they can be confident that a large audience will examine their work.¹⁴

Turning to the *leadership* more specifically, it may still be a puzzle that the leader initially turns over valuable code to the community.¹⁵ Despite the substantial status and career-concerns benefits of being a leader of an

¹⁴ Dasgupta and David [1994] suggest an alternative explanation for these patterns: the need to impress less-informed patrons who are likely to be impressed by the academic’s undertaking research in a ‘hot’ area. These patterns probably are driven by academic career concerns. New fields tend to be relatively more attractive to younger researchers, since older researchers have already invested in established fields and therefore have lower marginal costs of continuing in these fields. At the same time, younger researchers need to impress senior colleagues who will evaluate them for promotion. Thus, they need the presence of some of their seniors in the new fields.

¹⁵ Section V will discuss *companies’* incentives to release code.

important open source project, it would seem that most should not resist the large monetary gains from taking a promising technology private. We can only conjecture as to why this is not the case. One possibility is that taking the technology private may meet layers of resistance within the leader's corporation. To the extent that the innovation was made while working in-house, the programmer must secure a license from the employer;¹⁶ and her division, which does not want to lose a key programmer, may not be supportive of her demand. Another possibility is that the open source process may be a more credible way of harnessing energies when, say, fighting against a dominant player in the industry.

IV(ii). *Comparison Between Open Source and Closed Source Programming Incentives.*

To compare programmers' incentives in the open source and proprietary settings, we need to examine how the fundamental features of the two environments shape the incentives just reviewed. We will first consider the relative short-term rewards, and then turn to the deferred compensation.

Commercial projects have an edge on the current-compensation dimension because the proprietary nature of the code generates income. This makes it privately worthwhile for private companies to offer salaries.¹⁷ This contention is the old argument in economics that the prospect of profit encourages investment, which is used, for instance, to justify the awarding of patents to encourage invention.

By way of contrast, an open source project may well lower the cost for the programmer, for two reasons:

- (i) *'Alumni effect'*: Because the code is freely available to all, it can be used in schools and universities for learning purposes; so it is already familiar to programmers. This reduces their cost of programming for UNIX, for example.¹⁸
- (ii) *Customization and bug-fixing benefits*: The cost of contributing to an open source project can be offset if the activity brings about a private

¹⁶Open source projects may be seen as imposing less of a competitive threat to the firm. As a result, the firm may be less inclined to enforce its property rights on innovations turned open source. Alternatively, the firm may be unaware that the open source project is progressing.

¹⁷To be certain, commercial firms (e.g., Netscape, Sun, O'Reilly, Transmeta) supporting open source projects are also able to compensate programmers, because they indirectly benefit financially from these projects. Similarly, the government and not-for-profit corporations have done some subsidizing of open source projects. Still, there should be an edge for commercial companies.

¹⁸While we are here interested in private incentives to participate, note that this complementarity between apprenticeship and projects is socially beneficial. The social benefits might not increase linearly with open source market share, however, since the competing open source projects may end up competing for attention in the same common pool of students.

benefit (bug fixing, customization) for the programmer and her firm. Note again that this factor of cost reduction is directly linked to the openness of the source code.¹⁹

Let us now turn to the delayed reward (signaling incentive) component. In this respect too, the open source process has some benefits over the closed source approach. As we noted, signaling incentives are stronger, the more visible the performance and the more attributable the performance to a given individual. Signaling incentives therefore may be stronger in the open source mode for three reasons:

- (i) *Better performance measurement*: Outsiders can only observe inexactly the functionality and/or quality of individual elements of a typical commercially developed program, as they are unable to observe the proprietary source code. By way of contrast, in an open source project, the outsiders are able to see not only what the contribution of each individual was and whether that component ‘worked’, but also whether the task was hard, if the problem was addressed in a clever way, whether the code can be useful for other programming tasks in the future, and so forth.
- (ii) *Full initiative*: The open source programmer is her own boss and takes full responsibility for the success of a subproject. In a hierarchical commercial firm, however, the programmer’s performance depends on her supervisor’s interference, advice, etc. Economic theory would predict that the programmer’s performance is more precisely measured in the former case.²⁰
- (iii) *Greater fluidity*: It may be argued that the labor market is more fluid in an open source environment. Programmers are likely to have less idiosyncratic, or firm-specific, human capital that limits shifting one’s efforts to a new program or work environment. (Since many elements of the source code are shared across open source projects, more of the knowledge they have accumulated can be transferred to the new environment).

These theoretical arguments also provide insights as to *who* is more likely to contribute and *what tasks* are best suited to open source projects.

¹⁹To be certain, commercial companies leave Application Programming Interfaces for other people to provide add-ons, but this is still quite different from opening the source code.

²⁰On the relationship between empowerment and career concerns, see Ortega [2000]. In Cassiman’s [1998] analysis of research corporations (for-profit centers bringing together firms with similar research goals), free riding by parent companies boosts the researchers’ autonomy and helps attracting better talents. Cassiman argues that it is difficult to sustain a reputation for respecting the autonomy of researchers within firms. Cassiman’s analysis looks at real control, while our argument here results from the absence of formal control over the OS programmer’s activity.

Sophisticated users derive direct benefits when they customize or fix a bug in open source software.²¹ A second category of potential contributors consists of individuals with strong signaling incentives; these may use open source software as a port of entry. For instance, open source processes may give a talented system administrator at a small academic institution (who is also a user!) a unique opportunity to signal her talent to peers, prospective employers, and the venture capital community.²²

As to the tasks that may appeal to the open source community, one would expect that tasks such as those related to the operating systems and programming languages, whose natural audience is the community of programmers, would give rise to strong signaling incentives. (For instance, the use of Perl is largely restricted to system administrators.) By way of contrast, tasks aiming at helping the much-less-sophisticated end user—e.g., design of easy-to-use interfaces, technical support, and ensuring backward compatibility—usually provide lower signaling incentives.²³

IV(iii). *Evidence on Individual Incentives*

A considerable amount of evidence is consistent with an economic perspective.

First, user benefits are key to a number of open source projects. One of

²¹ A standard argument in favor of open source processes is their massive parallel debugging. Typically, commercial software firms can only ask users to point at problems: beta testers do not fix the bugs, they just report them. It is also interesting to note that many commercial companies do not discourage their employees from working on open source projects. In many cases where companies encourage such involvement, programmers use open source tools to fix problems. Johnson [1999] builds a model of open source production by a community of user-developers. There is one software program or module to be developed, which is a public good for the potential developers. Each of the potential developers has a private cost of working on the project and a private value of using it; both of which are private information. Johnson shows that the probability that the innovation is made need not increase with the number of developers, as free-riding is stronger when the number of potential developers increases.

²² An argument often heard in the open source community is that people participate in open source projects because programming is fun and because they want to be 'part of a team'. While this argument may contain a grain of truth, it is somewhat puzzling as it stands; for example, it is not clear why programmers who are part of a commercial team could not enjoy the same intellectual challenges and the same team interaction as those engaged in open source development. (To be sure, it may be challenging for programmers to readily switch employers if their peers in the commercial entity are not congenial.) The argument may reflect the ability of programmers to use participation in open source projects to overcome the barriers that make signaling in other ways problematic.

²³ Valloppillil [1998] further argues that reaching commercial grade quality often involves unglamorous work on power management, management infrastructure, wizards, etc., that makes it unlikely to attract open source developers. Valloppillil's argument seems a fair description of past developments in open source software. Some open source proponents do not confer much predictive power on his argument, though; they predict, for example, that open source user interfaces such as GNOME and KDE will achieve commercial grade quality.

the origins of the free software movement was Stallman's inability to improve a printer program because Xerox refused to release the source code. In three of the four scenarios described in Section III, the project founders were motivated by information technology problems that they had encountered in their day-to-day work. For instance, in the case of Apache, the initial set of contributors was almost entirely system administrators who were struggling with the same types of problems as Behlendorf. In each case, the initial release was 'runnable and testable': it provided a potential, even if imperfect, solution to a problem that was vexing considerable numbers of data processing professionals.

Second, it is clear that giving credit to authors is essential in the open source movement. This principle is included as part of the nine key requirements in the 'Open Source Definition' [Open Source Initiative, 1999]. This point is also emphasized by Raymond [1999b], who points out 'surreptitiously filing someone's name off a project is, in cultural context, one of the ultimate crimes'.

More generally, the reputational benefits that accrue from successful contributions to open source projects appear to have real effects on the developers. This is acknowledged within the open source community itself. For instance, according to Raymond [1999b], the primary benefits that accrue to successful contributors of open source projects 'good reputation among one's peers, attention and cooperation from others, . . . [and] higher status [in the] . . . exchange economy'. Thus, while some of benefits conferred from participation in open source projects may be less concrete in nature, there also appear to be quite tangible—if delayed—rewards.

The Apache project provides a good illustration of these observations. The project makes a point of recognizing all contributors on its web site, even those who simply identify a problem without proposing a solution. Similarly, the organization highlights its most committed contributors, who have the ultimate control over the project's evolution. Moreover, it appears that many of the skilled Apache programmers have benefited materially from their association with the organization. Numerous contributors have been hired into Apache development groups within companies such as IBM, become involved in process-oriented companies such as Collab.Net which seek to make open source projects more feasible (see below), or else moved into other Internet tools companies in ways that were facilitated by their expertise and relationships built up during their involvement in the open source movement. Meanwhile, many of the new contributors are already employed by corporations, and working on Apache development as part of their regular assignments.

There is also substantial evidence that open source work may be a good stepping stone for securing access to venture capital. For example, the founders of Sun, Netscape, and Red Hat had signaled their talent in the open source world. In Table II, we summarize some of the subsequent

TABLE II
 COMMERCIAL ROLES PLAYED BY SELECTED INDIVIDUALS ACTIVE IN OPEN SOURCE MOVEMENT

Individual	Role and Company
Eric Allman	Chief Technical Officer, Sendmail, Inc. (support for open source software product)
Brain Behlendorf	Founder, President, and Chief Technical Officer, Collab.Net (management of open source projects)
Keith Bostic	Founder and President, Sleepycat Software
L. Peter Deutsch	Founder, Aladdin Enterprises (support for open source software product)
William Joy	Founder and Chief Scientist, Sun Microsystems (workstation and software manufacture)
Michael Tiemann	Founder, Cygnus Solutions (open source support)
Linus Torvalds	Employee, Transmeta Corporation (chip design company)
Paul Vixie	President, Vixie Enterprises (engineering and consulting services)
Larry Wall	Employee, O'Reilly & Associates (software documentation publisher)

commercial roles played by individuals active in the open source movement.

IV(iv). *Organization and Governance*

Favorable characteristics for open source production are (a) its modularity (the overall project is divided into much smaller and well-defined tasks ('modules') that individuals can tackle independently from other tasks) and (b) the existence of fun challenges to pursue.²⁴ A successful open source project also requires a credible leader or leadership, and an organization consistent with the nature of the process. Although the leader is often at the origin a user who attempts to solve a particular program, the leader over time performs less and less programming. The leader must provide a 'vision', attract other programmers, and, last but not least, 'keep the project together' (prevent it from forking or being abandoned).

Initial Characteristics

The success of an open source project is dependent on the ability to break the project into distinct components. Without an ability to parcel out work in different areas to programming teams who need little contact with one another, the effort is likely to be unmanageable. Some observers argue that the underlying Unix architecture lent itself well to the ability to break development tasks into distinct components. It may be that as new open source projects move beyond their Unix origins and encounter new programming challenges, the ability to break projects into distinct units will be less possible. But recent developments in computer science and programming languages (e.g., the development of object-oriented programming) have encouraged further modularization, and may facilitate future open source projects.

The initial leader must also assemble a critical mass of code to which the programming community can react. Enough work must be done to show that the project is doable and has merit. At the same time, to attract additional programmers, it may be important that the leader does not perform too much of the job on his own and leaves challenging programming problems to others.²⁵ Indeed, programmers will initially be reluctant to join a project unless they identify an exciting challenge.

²⁴ Open source projects have trouble attracting people initially unless they leave fun challenges 'up for grabs'. On the other hand, the more programmers an open source project attracts, the more quickly the fun activities are completed. The reason why the projects need not burn out once they grow in ranks is that the 'fixed cost' that individual programmers incur when they first contribute to the project is sunk and so the marginal cost of continuing to contribute is smaller than the initial cost of contributing.

²⁵ For example, Valloppillil's [1998] discussion of the Mozilla release.

Another reason why programmers are easier to attract at an early stage is that, if successful, the project will keep attracting a large number of programmers in the future, making early contributions very visible.

Consistent with this argument, it is interesting to note that each of the four cases described above appeared to pose challenging programming problems.²⁶ When the initial release of each of these open source programs was made, considerable programming problems were unresolved. The promise that the project was not near a 'dead end', but rather would continue to attract ongoing participation from programmers in the years to come, appears to be an important aspect of its appeal.

In this respect, Linux is perhaps the quintessential example. The initial Linux operating system was quite minimal, on the order of a few tens of thousands of lines of code. In Torvalds' initial postings in which he sought to generate interest in Linux, he explicitly highlighted the extent to which the version would require creative programming in order to achieve full functionality. Similarly, Larry Wall attributes the much of the success of Perl to the fact that it 'put the focus on the creativity of the programmer'. Because it has a very limited number of rules, the program has evolved in a variety of directions that were largely unanticipated when Wall initiated the project.

Leadership

Another important determinant of project success appears to be the nature of its leadership. In some respects, the governance structures of open source projects are quite different. In a number of instances, such as Linux, there is an undisputed leader. While certain aspects are delegated to others, a strong centralization of authority characterizes these projects. In other cases, such as Apache, a committee will resolve the disputes by voting or a consensus process.

At the same time, leaders of open source projects share some common features. Most leaders are the programmers who developed the initial code for the project (or made another important contribution early in the project's development). While many make fewer programming contributions, having moved on to broader project management tasks, the individuals that we talked to believed that the initial experience was important in establishing credibility to manage the project. The splintering of the Berkeley-derived Unix development programs has been attributed in part to the absence of a single credible leader.

But what does the leadership of an open source project do? It might appear at first sight that the unconstrained, quasi-anarchistic nature of the

²⁶ It should be cautioned that these observations are based on a small sample of successful projects. Observing which projects succeed or fail and the reasons for these divergent outcomes in an informal setting such as this one is quite challenging.

open source process leaves little scope for a leadership. This, however, is incorrect. While the leader has no 'formal authority' (she is unable to instruct anyone to do anything), she has substantial 'real authority' in successful open source projects.²⁷ That is, her 'recommendations', broadly viewed, tend to be followed by the vast majority of programmers working on the project. These recommendations include the initial 'vision' (agenda for work, milestones), the subsequent updating of goals as the project evolves, the appointment of key leaders, the cajoling of programmers so as to avoid attrition or forking, and the overall assessment of what has been and should be achieved. (Even though participants are free to take the project where they want as long as they release the modified code, acceptance by the leadership of a modification or addition provides some certification as to its quality and its integration/compatibility with the overall project. The certification of quality is quite crucial to the open source project because the absence of liability raises concerns among users that are stronger than for commercial software, for which the vendor is liable).

The key to a successful leadership is the programmers' trust in the leadership: that is, they must believe that the leader's objectives are sufficiently congruent with theirs and not polluted by ego-driven, commercial, or political biases. In the end, the leader's recommendations are only meant to convey her information to the community of participants. The recommendations receive support from the community only if they are likely to benefit the programmers, that is only if the leadership's goals are believed to be aligned with the programmers' interests.

For instance, the leadership must be willing to accept meritorious improvements, even though they do not fit within the leader's original blueprint. Trust in the leadership is also key to the prevention of forking. While there are natural forces against forking (the loss of economies of scale due to the creation of smaller communities, the hesitations of programmers in complementary segments to port to multiple versions, and the stigma attached to the existence of a conflict), other factors may encourage forking. User-developers may have conflicting interests as to the evolution of the technology. Ego (signaling) concerns may also prevent a faction from admitting that another approach is more promising, or simply from accepting that it may socially be preferable to have one group join the other's efforts even if no clear winner has emerged. The presence of a charismatic (i.e., trusted) leader is likely to substantially reduce the probability of forking in two ways. First, indecisive programmers are likely to rally behind the leadership's preferred alternative. Second, the dissenting faction may not have an obvious leader of its own.

²⁷ The terminology and the conceptual framework are here borrowed from Aghion-Tirole [1997].

A good leadership should also clearly communicate its goals and evaluation procedures. Indeed, the open source organizations go to considerable efforts to make the nature of their decision making process transparent: the process by which the operating committee reviews new software proposals is frequently posted and all postings archived. For instance, on the Apache web site, it is explained how proposed changes to the program are reviewed by the program's governing body, whose membership is largely based on contributions to the project. (Any significant change requires at least three 'yes' votes—and no vetoes—by these key decision-makers.)

V. COMMERCIAL SOFTWARE COMPANIES' REACTIONS TO THE OPEN SOURCE MOVEMENT

This section examines the interface between open and closed source software development. Challenged by the successes of the open source movement, the commercial software corporations may employ one of the following two strategies. The first is to emulate some incentive features of open source processes in a distinctively closed source environment. Another is to try to mix open and closed source processes to get the best of both worlds.

V(i). *Why Don't Corporations Duplicate the Open Source Incentives?*

As we already noted, owners of proprietary code are not able to enjoy the benefits of getting free programmer training in schools and universities (the alumni effect); nor can they easily allow users to modify their code and customize it without jeopardizing intellectual property rights.

Similarly, and for the reasons developed in Section IV, commercial companies will never be able to fully duplicate the visibility of performance reached in the open source world. At most can they duplicate to some extent some of the signaling incentives of the open source world. Indeed, a number of commercial software companies (e.g., video game companies, Qualcomm for the Eudora email program) list people who have developed the software. It is an interesting question why others do not. To be certain, commercial companies do not like their key employees to become highly visible, lest they be hired away by competitors.²⁸ But, to a large extent, firms also realize that this very visibility enables them to attract talented individuals and provides a powerful incentive to existing employees.²⁹

²⁸ For instance, concerns about the 'poaching' of key employees was one of the reasons cited for Steve Jobs' recent decision to cease giving credit to key programmers in Apple products [Claymon, 1999].

²⁹ For the economic analysis of employee visibility, see Gibbons [1997] and Gibbons and Waldman's [1999] review essays. Ronde [1999] models the firms' incentives to 'hide' their workers from the competition in order to preserve their trade secrets.

To be certain, team leaders in commercial software build reputations and get identified with proprietary software just as they can on open source projects; but the ability of reputations to spread beyond the leaders is more limited, due to the non-verifiability of claims about who did what.³⁰

Another area in which software companies might try to emulate open source development is the promotion of widespread code sharing within the company. This may enable them to reduce code duplication and to broaden a programmer's audience. Interestingly, existing organizational forms may preclude the adoption of open source systems within commercial software firms. An internal Microsoft document on open source [Valloppillil, 1998] describes a number of pressures that limit the implementation of features of open source development within Microsoft. Most importantly, each software development group appears to be largely autonomous. Software routines developed by one group are not shared with others. In some instances, the groups seek to prevent being broken up by not documenting a large number of program features. These organizational attributes, the document suggests, lead to very complex and interdependent programs that do not lend themselves to development in a 'compartmentalized' manner nor to widespread sharing of source code.³¹

V(ii). The Commercial Software Companies' Open Source Strategies

As should be expected, many commercial companies have undertaken strategies to capitalize on the open source movement. In a nutshell, they expect to benefit from their expertise in some segment whose demand is boosted by the success of a complementary open source program. While improvements in the open source software are not appropriable, commercial companies can benefit indirectly in a complementary proprietary segment.³²

Living symbiotically off an open source project

One such strategy is straightforward. It consists of commercially providing complementary services and products that are not supplied efficiently by

³⁰ Commercial vendors try to address this problem in various ways. For example, Microsoft developers now have the right to present their work to their users. Promotions to 'distinguished engineer' or to a higher rank more generally as well as the granting of stock options as a recognition of contributions also make the individual performance more visible to the outside world.

³¹ Cusamano and Selby (1995), however, document a number of management institutions at Microsoft that attempt to limit these pressures.

³² Another motivation for commercial companies to interface with the open source world may be public relations. Furthermore, firms may temporarily encourage programmers to participate in open source projects to learn about the strengths and weaknesses of this development approach.

the open source community. Red Hat and VA Linux for example, exemplify this 'reactive' strategy.³³

In principle, a 'reactive' commercial company may want to encourage and subsidize the open source movement, for example by allocating a few programmers to the open source project.³⁴ Red Hat will make more money on support if Linux is successful. Similarly, if logic semiconductors and operating systems for personal computers are complements, one can show by a revealed preference argument that Intel's profits will increase if Linux (which unlike Windows is free) takes over the PC operating system market. Sun may benefit if Microsoft's position is weakened; Oracle may wish to port its database products to a Linux environment in order to lessen its dependence on Sun's Solaris operating system; and so forth. Because firms do not capture all the benefits of the investments, however, the free-rider problem often discussed in the economics of innovation should apply here as well. Subsidies by commercial companies for open source projects should remain limited unless the potential beneficiaries succeed in organizing a consortium (which will limit the free-riding problem).

Code Release

A second strategy is to take a more proactive role in the development of open source software. Companies can release existing proprietary code and create some governance structure for the resulting open source process. For example, Hewlett-Packard recently released its Spectrum Object Model-Linker to the open source community in order to help the Linux community port Linux to Hewlett Packard's RISC architecture.³⁵ This is similar to the strategy of giving away the razor (the released code) to sell more razor blades (the related consulting services that HP will provide).

When can it be advantageous for a commercial company to release proprietary code under an open source license? The first condition is, as we have noted, that the company expects to thereby boost its profit on a complementary segment. A second is that the increase in profit in the proprietary complementary segment offsets any profit that would have been made in the primary segment, had it not been converted to open source. Thus, the temptation to go open source is particularly strong when

³³ Red Hat provides support for Linux-based products, while VA Linux provides hardware products optimized for the Linux environment. In December 1999, their market capitalizations were \$17 and \$10 billion respectively, though they have subsequently declined significantly.

³⁴ Of course, these programmers also increase the company's ability to learn from scientific and technical discoveries elsewhere and help the company with the development of the proprietary segment.

³⁵ Companies could even (though probably less likely) encourage *ex nihilo* development of new pieces of open source software.

the company is too small to compete commercially in the primary segment or when it is lagging behind the leader and about to become extinct in that segment.^{36,37}

Various efforts by corporations selling proprietary software products to develop additional products through an open source approach have been undertaken. One of the most visible of these efforts was Netscape's 1998 decision to make 'Mozilla', a portion of its browser source code, freely available. This effort encountered severe difficulties in its first year, only receiving approximately two dozen postings by outside developers. Much of the problems appeared to stem from the insufficiently modular nature of the software: reflecting its origins as a proprietary commercial product, the different portions of the program were highly interdependent and interwoven. Netscape eventually realized it needed to undertake a major restructuring of the program, in order to enhance the ability of open source programmers to contribute to individual sections. It is also likely that Netscape raised some suspicions by not initially adopting the right governance structure. Leadership by a commercial entity may not internalize enough of the objectives of the open source community. In particular, a corporation may not be able to credibly commit to keeping all source code in the public domain and to adequately highlighting important contributions.³⁸

For instance, in the Mozilla project, Netscape's unwillingness to make large amounts of browser code public was seen as an indication of its questionable commitment to the open source process. In addition, Netscape's initial insistence on the licensing terms that allowed the corporation to relicense the software developed in the open source project on a proprietary basis was viewed as problematic [Hamerly, Paquin and Walton, 1999]. (The argument is here the mirror image of the standard argument in industrial economics that a firm may want to license its technology to several licensees in order to commit not to expropriate producers of complementary goods and services in the future: see Shepard

³⁶ See, for example, the discussion of SGI's open source strategy in Taschek [1999].

³⁷ It should also be noted that many small developers are uncomfortable doing business with leading software firms, feeling them to be exploitative, and that these barriers may be overcome by the adoption of open source practices by the large firms. A rationalization of this story is that, along the lines of Farrell and Katz [2000], the commercial platform owner has an incentive to introduce substitutes in a complementary segment, in order to force prices down in that segment and to raise the demand for licenses to the software platform. When, however, the platform is available through (say) a BSD-style license, the platform owner has no such incentives, as he cannot raise the platform's price. Vertical relationships between small and large firms in the software industry are not fully understood, and would reward further study.

³⁸ An interesting question is why corporations do not replicate the modular structure of open source software in commercial products more generally. One possibility may be that modular code, whatever its virtues for a team of programmers working independently, is not necessarily better for a team of programmers and managers working together.

[1987] and Farrell and Gallini [1988].) Netscape initially proposed the 'Netscape Public License', a cousin to the BSD license that allowed Netscape to take pieces of the open source code and turn them back into a proprietary project again. The licensing terms, however, may not have been the hindering factor, since the terms of the final license are even stricter than those of the GPL. Under this new license (the 'Mozilla Public License'), Netscape cannot relicense the modifications to the code.

Intermediaries

In this light, it is tempting to interpret the creation of organizations such as Collab.Net as efforts to certify corporate open source development programs, just as investment banks and venture capitalists play a certification role for new firms. Collab.Net, a new venture funded by the venture capital group Benchmark Partners, will organize open source projects for corporations who wish to develop part of their software in this manner. Collab.Net will receive fees for its online marketplace (SourceXchange, through which corporations will contact open source developers), for preparing contracts, for helping select and monitor developers, and for settling disputes. Hewlett Packard released the core of its E-speak technology (which enable brokering capabilities) to open source³⁹ and posted six projects related to this technology.

Hewlett Packard's management of the open source process seems consistent with Dessein [1999]. Dessein shows that a principal with formal control rights over an agent's activity in general gains by delegating his control rights to an intermediary with preferences or incentives that are intermediate between his and the agent's. The partial alignment of the intermediary's preferences with the agent's fosters trust and boosts the agent's initiative, ultimately offsetting the partial loss of control for the principal. In the case of Collab.Net, the congruence with the open source developers is obtained through the employment of visible open source developers (for example, the president and chief technical officer is Brian Behlendorf, one of the cofounders of the Apache project) and the involvement of O'Reilly, a technical book publisher with strong ties to the open source community.

VI. FOUR OPEN ECONOMIC QUESTIONS ABOUT OPEN SOURCE

There are many other issues posed by open source development that require further thought. This section will highlight a number of these as suggestions for future work.

³⁹Some of the E-speak code remains proprietary to Hewlett Packard; so will some applications and utilities developed in the future. It should also be noted that HP can profit by providing services to E-speak users, which, while not proprietary, should be an arena in which HP has a natural advantage.

VI(i). *Which Technological Characteristics are Conducive to a Smooth Open Source Development?*

This paper has identified a number of attributes that make a project a good or poor candidate for open source development. But it has stopped short of providing a comprehensive picture of determinants of a smooth open source development. Let us mention a few topics that are worth further investigation:

- *Role of applications and related programs.* Open source projects differ in the functionalities they offer and in the number of add-ons that are required to make them attractive. As the open source movement comes to maturity, it will confront some of the same problems as commercial software does, namely the synchronization of upgrades and the efficient level of backward compatibility. A user who upgrades a program (which is very cheap in the open source case) will want either the new program to be backward compatible or applications to have themselves been upgraded to the new version.⁴⁰ We know from commercial software that both approaches to compatibility are costly; for example, Windows programmers devote a lot of time to backward compatibility issues, and encouraging application development requires fixing applications programming interfaces about three years before the commercial release of the operating system. A reasonable conjecture could be that open source programming would be appropriate when there are fewer applications or when IT professionals can easily adjust the code so as to ensure compatibility themselves.
- *Influence of competitive environment.* Based on very casual observation, it seems that open source projects sometimes gain momentum when facing a battle against a dominant firm, although our examples show open source projects can do well even in the absence of competition.⁴¹ To understand why this might be the case (assuming this is an empirical fact, which remains to be established!), it would be useful to go back to the economics of cooperative joint ventures. The latter are known to work better when the members have similar objectives.⁴² The existence of a dominant competitor in this respect tends to align the goals of the members, and the

⁴⁰The former solution may be particularly desirable if the user has customized last generation's applications.

⁴¹Wayner [2000] argues that the open source movement is not about battling Microsoft or other leviathans and notes that in the early days of computing (say, until the late seventies) code sharing was the only way to go as 'the computers were new, complicated, and temperamental. Cooperation was the only way that anyone could accomplish anything'. This argument is consistent with the hypothesis stated below, according to which the key factor behind cooperation is the alignment of objectives and this alignment may come from the need to get a new technology of the ground, from the presence of a dominant firm, or from other causes.

⁴²See, e.g., Hansmann [1996].

best way to fight an uphill battle against the dominant player is to remain united. To be certain, open source software development works differently from joint venture production, but it also relies on cooperation within a heterogeneous group; the analogy is well worth pursuing.

- *Project lifespan.* One of the arguments offered by open source advocates is that because their source code is publicly available, and at least some contributions will continue to be made, its software will have a longer duration. (Many software products by commercial vendors are abandoned or no longer upgraded after the developer is acquired or liquidated, or even when the company develops a new product to replace the old program.) But another argument is that the nature of incentives being offered open source developers—which as discussed above, lead them to work on highly visible projects—might lead to a ‘too early’ abandonment of projects that experience a relative loss in popularity. An example is the XEmacs project, an open source project to create a graphical environment with multiple ‘windows’ that originated at Stanford. Once this development effort encountered an initial decline in popularity, many of the open source developers appeared to move onto alternative projects.

VI(ii). *Optimal Licensing*

Our discussion of open source licensing has been unsatisfactory. Some licenses (e.g., BSD and its close cousin the Apache license) are relatively permissive, while others (e.g., GPL) force the user to distribute any changes or improvements (share them) if they distribute the software at all.

Little is known about the trade-off between encouraging add-ons that would not be properly supplied by the open source movement and preventing commercial vendors (including open source participants) from free riding on the movement or even ‘hijacking it’. An open source project may be ‘hijacked’ by a participant who builds a valuable module and then offers proprietary APIs to which application developers start writing. The innovator has then built a platform that appropriates some of the benefits of the project. To be certain, open source participants might then be outraged, but it is unclear whether this would suffice to prevent the hijacking. The open source community would then be as powerless as the commercial owner of a platform above which a ‘middleware’ producer superimposes a new platform.⁴³

⁴³ The increasing number of software patents being granted by the U.S. Patent and Trademark Office provide another avenue through which such a ‘hijacking’ might occur. In a number of cases, industry observers have alleged that patent examiners—not being very familiar with the unpatented ‘prior art’ of earlier software code—have granted unreasonably broad patents, in some cases giving the applicant rights to software that was originally developed through open source processes.

The exact meaning of the ‘viral’ provisions in the GPL license, say, or more generally the implications of open source licenses have not yet been tested in court. Several issues may arise in such litigation: for instance, who has standing for representing the project if the community is fragmented, and how a remedy would be implemented (e.g., the awarding of damages for breach of copyright agreement may require incorporating the beneficiaries).

VI(iii). *Coexistence of Commercial and Open Source Software*

On a related note, the existence of commercial entities living symbiotically off the efforts of open source programmers as well as participating in open source projects raises new questions.

The flexible open source licenses allow for the coexistence of open and closed source code. While it represents in our view (and in that of many open source participants) a reasonable compromise, it is not without hazards.

The coexistence of commercial activities may alter the programmers’ incentives. Programmers working on an open source project may be tempted to stop interacting and contributing freely if they think they have an idea for a module that might yield a huge commercial payoff. Too many programmers may start focusing on the commercial side, making the open source process less exciting. Although they refer to a different environment, the concerns that arise about academics’ involvement in start-up firms, consulting projects, and patenting may be relevant here as well. While it is too early to tell, some of these same issues may appear in the open source world.⁴⁴

VI(iv). *Can the Open Source Process be Transposed to Other Industries?*

An interesting final question is whether the open source model can be transposed to other industries. Could automobile components be developed in an open source mode, with GM and Toyota performing an assembler function similar to that of Red Hat for Linux? Many industries involve forms of cooperation between commercial entities in the form of for-profit or not-for-profit joint ventures. Others exhibit user-driven innovation or open science cultures. Thus, a number of ingredients of open source software are not specific to the software industry. Yet no

⁴⁴ A related phenomenon that would reward academic scrutiny is ‘shareware’. Many of packages employed by researchers (including several used by economists, such as MATLAB, SAS, and SPSS) have grown by accepting modules contributed by users. The commercial vendors co-exist with the academic user community in a positive symbiotic relationship. These patterns provide a useful parallel to open source.

other industry has yet produced anything quite like open source development. An important research question is whether other industries ever will.

Although some aspects of open source software collaboration (such as electronic information exchange across the world) could easily be duplicated, other aspects would be harder to emulate. Consider, for example, the case of biotechnology. It may be impossible to break up large projects into small manageable and independent modules and there may not be sufficient sophisticated users who can customize the molecules to their own needs. The tasks that are involved in making the product available to the end user involve much more than consumer support and even friendlier user interfaces. Finally, the costs of designing, testing, and seeking regulatory approval for a new drug are enormous.

More generally, in many industries the development of individual components require large team work and substantial capital costs, as opposed to (for some software programs) individual contributions and no capital investment (besides the computer the programmer already has). Another obstacle is that in mass-market industries users are numerous and rather unsophisticated, and so deliver few services of peer recognition and ego gratification. This suggests that the open source model may not easily be transposed to other industries, but further investigation is warranted.

Our ability to answer confidently these and related questions is likely to increase as the open source movement itself grows and evolves. At the same time, it is heartening to us how much of open source activities can be understood within existing economic frameworks, despite the presence of claims to the contrary. The literatures on 'career concerns' and on competitive strategies provide lenses through which the structure of open source projects, the role of contributors, and the movement's ongoing evolution can be viewed.

REFERENCES

- Aghion, P. and Tirole, J., 1997, 'Formal and Real Authority in Organizations', *Journal of Political Economy*, 105, pp. 1–29.
- Brooks, F., 1975, *The Mythical Man Month: Essays on Software Engineering* (Addison-Wesley, Reading, Mass).
- Browne, C. B., 1999, 'Linux and Decentralized Development', http://www.firstmonday.dk/issues/issue3_3/browne/index.html (accessed December 17, 1999).
- Caminer, D., Aris, J., Hermon, P. and Land, F., 1996, *User Driven Innovation: The World's First Business Computer* (McGraw-Hill, New York).
- Cassiman, B., 1998, 'The Organization of Research Corporations and Researcher Ability', Unpublished working paper, University Pompeu Fabra.
- Claymon, D., 1999, 'Apple in Tiff with Programmers over Signature Work', *San Jose Mercury News*, December 2.

- Cockburn, I., Henderson, R. and Stern, S., 1999, 'Balancing Incentives: The Tension Between Basic and Applied Research', Working Paper 6882, National Bureau of Economic Research.
- Cusumano, M. A., 1991, *Japan's Software Factories: A Challenge to U.S. Management* (Oxford University Press, New York, NY).
- Cusumano, M. A. and Selby, R. W., 1995, *Microsoft Secrets: How the World's Most Powerful Software Company Creates Technology, Shapes Markets, and Manages People* (Free Press, New York).
- Darwall, C. and Lerner, J., 2000, 'A Note on Open Source Software', Harvard Business School Note 9-201-078.
- Dasgupta, P. and David, P., 1994, 'Towards a New Economics of Science', *Research Policy*, 23, pp. 487–521.
- Dempsey, B. J., Weiss, D., Jones, P. and Grenberg, J., 1999, 'A Quantitative Profile of a Community of Open Source Linux Developers', Unpublished working paper, School of Information and Library Science, University of North Carolina at Chapel Hill, <http://metalab.unc.edu/osrt/develop.html> (accessed November 1, 1999).
- Dessein, W., 1999, 'Authority and Communication in Organizations', Unpublished working paper, Université Libre de Bruxelles.
- DiBona, C., Ockman, S. and Stone, M. eds, 1999, *Open Sources: Voices from the Open Source Revolution* (O'Reilly, Sebastopol, California).
- Edwards, K., 2000, 'When Beggars Become Choosers', http://www.firstmonday.org/issues/issue5_10/edwards/index.html (accessed Oct. 5, 2000).
- Farrell, J. and Gallini, N., 1988, 'Second Sourcing as a Commitment: Monopoly Incentives to Attract Competition', *Quarterly Journal of Economics*, 103, pp. 673–694.
- Farrell, J. and Katz, M. L., 2000, 'Innovation, Rent Extraction, and Integration in Systems Markets', *Journal of Industrial Economics*, 48, pp. 413–432.
- Gambardella, A., 1995, *Science and Innovation: The US Pharmaceutical Industry During the 1980's* (Cambridge University Press, Cambridge, U.K.).
- Ghosh, R. A., 1999, 'Cooking Pot Markets: An Economic Model for the Trade in Free Goods and Services on the Internet', http://www.firstmonday.dk/issues/issue3_3/ghosh/index.html (accessed December 17, 1999).
- Ghosh, R. A. and Prakash, V. V., 'The Oribiten Free Software Survey: May 2000', <http://orniten.org/ofss/01.html> (accessed May 10, 2000).
- Gibbons, R., 1997, 'Incentives and Careers in Organizations', in D. Kreps and K. Wallis (eds), *Advances in Economic Theory and Econometrics*, vol. 2 (Cambridge University Press).
- Gibbons, R. and Waldman, M., 1999, 'Careers in Organizations: Theory and Evidence', in O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, vol. 3B, Chapter 36 (North Holland, New York).
- Gomulkiewicz, R. W., 1999, 'How Copyleft Uses License Rights to Succeed in the Open Source Software Revolution and the Implications for Article 2B', *Houston Law Review*, 36, pp. 179–194.
- Hammerly, J., Paquin, T. and Walton, S., 1999, 'Freeing the Source: The Story of Mozilla', in C. DiBona, S. Ockman and M. Stone (eds), *Open Sources: Voices from the Open Source Revolution* (O'Reilly, Sebastopol, California), pp. 197–206.
- Hansmann, H., 1996, *The Ownership of Enterprise*, (Belknap Harvard, New York).
- Henderson, R. and Cockburn, I., 1994, 'Measuring Competence? Exploring Firm Effects in Pharmaceutical Research', *Strategic Management Journal*, 15 (Winter Special Issue), pp. 63–84.

- Holmström, B., 1999, 'Managerial Incentive Problems: A Dynamic Perspective', *Review of Economic Studies*, 66, pp. 169–182.
- Johnson, J. P., 1999, 'Economics of Open-Source Software', Unpublished working paper, Massachusetts Institute of Technology.
- Lakhani, K. and von Hippel, E., 2000, 'How Open Source Software Works: 'Free' User-to-User Assistance', Working Paper 4117, Sloan School of Management, Massachusetts Institute of Technology.
- Levy, S., 1984, *Hackers: Heroes of the Computer Revolution* (Anchor Press/Doubleday, Garden City, NY).
- Merton, R. K., 1973, *The Sociology of Science: Theoretical and Empirical Investigations* (University of Chicago Press, Chicago, IL).
- Mockus, A., Fielding, R. and Herbsleb, J., 2000, 'A Case Study of Open Source Software Program Movement: The Apache Server', herbsleb@research.bell-labs.com (accessed December 30, 2000).
- Nadeau, T., 1999, 'Learning from Linux', <http://www.os2hq.com/archives/linmemo1.htm> (accessed November 12, 1999).
- Netcraft, 2000, 'The Netcraft Web Server Survey', <http://www.netcraft.com/survey> (accessed December 24, 2000).
- Open Source Initiative, 1999, 'Open Source Definition', <http://www.opensource.org/osd.html> (accessed November 14, 1999).
- Ortega, J., 2000, 'Power in the Firm and Managerial Career Concerns', Unpublished working paper, Universidad Carlos III de Madrid.
- Raymond, E. S., 1999a, 'The Cathedral and the Bazaar', <http://www.tuxedo.org/~esr/writings/cathedral-bazaar> (accessed November 9, 1999).
- Raymond, E. S., 1999b, 'Homesteading the Noosphere: An Introductory Contradiction', <http://www.tuxedo.org/~esr/writings/homesteading> (accessed November 11, 1999).
- Ronde, T., 1999, 'Trade Secrets and Information Sharing', Unpublished working paper, University of Mannheim.
- Rosenberg, N., 1976, *Perspectives on Technology* (Cambridge University Press, Cambridge).
- Rosenbloom, R. S. and Spencer, W. J., ed, 1996, *Engines of Innovation: U.S. Industrial Research at the End of an Era* (Harvard Business School Press, Boston).
- Shepard, A., 1987, 'Licensing to Enhance Demand for New Technologies', *Rand Journal of Economics*, 18, pp. 360–368.
- Stallman, R., 1999, 'The GNU Operating System and the Free Software Movement', in C. DiBona, S. Ockman and M. Stone (eds), *Open Sources: Voices from the Open Source Revolution* (O'Reilly, Sebastopol, California), pp. 53–70.
- Taschek, J., 1999, 'Vendor Seeks Salvation By Giving Away Technology', <http://www.zdnet.com/pcweek/stories/news/0,4153,404867,00.html> (accessed December 17, 1999).
- Torvalds, L., 1999, 'Interview with Linus Torvalds: What Motivates Free Software Developers?', http://www.firstmonday.dk/issues/issue3_3/torvalds/index.html (accessed December 17, 1999).
- U.S. Department of Labor, Bureau of Labor Statistics, 2000, *Occupational Outlook Handbook: 2000–2001 Edition* (Government Printing Office, Washington).
- Valloppillil, V., 1998, 'Open Source Software: A (New?) Development Methodology', [also referred to as The Halloween Document], Unpublished working paper, Microsoft Corporation, <http://www.opensource.org/halloween/halloween1.html> (accessed November 9, 1999).

von Hippel, E., 1988, *The Sources of Invention* (Oxford University Press, New York).

Wayner, P., 2000, *Free for All: How Linux and the Free Software Movement Undercut the High-Tech Titans* (Harper Collins, New York).

As Coronavirus Surveillance Escalates, Personal Privacy Plummet

Tracking entire populations to combat the pandemic now could open the doors to more invasive forms of government snooping later.

By **Natasha Singer and Choe Sang-Hun**

Published March 23, 2020 Updated April 17, 2020

In South Korea, government agencies are harnessing surveillance-camera footage, smartphone location data and credit card purchase records to help trace the recent movements of coronavirus patients and establish virus transmission chains.

In Lombardy, Italy, the authorities are analyzing location data transmitted by citizens' mobile phones to determine how many people are obeying a government lockdown order and the typical distances they move every day. About 40 percent are moving around "too much," an official recently said.

In Israel, the country's internal security agency is poised to start using a cache of mobile phone location data — originally intended for counterterrorism operations — to try to pinpoint citizens who may have been exposed to the virus.

As countries around the world race to contain the pandemic, many are deploying digital surveillance tools as a means to exert social control, even turning security agency technologies on their own civilians. Health and law enforcement authorities are understandably eager to employ every tool at their disposal to try to hinder the virus — even as the surveillance efforts threaten to alter the precarious balance between public safety and personal privacy on a global scale.

Yet ratcheting up surveillance to combat the pandemic now could permanently open the doors to more invasive forms of snooping later. It is a lesson Americans learned after the terrorist attacks of Sept. 11, 2001, civil liberties experts say.

Nearly two decades later, law enforcement agencies have access to higher-powered surveillance systems, like fine-grained location tracking and facial recognition — technologies that may be repurposed to further political agendas like anti-immigration policies. Civil liberties experts warn that the public has little recourse to challenge these digital exercises of state power.

"We could so easily end up in a situation where we empower local, state or federal government to take measures in response to this pandemic that fundamentally change the scope of American civil rights," said Albert Fox Cahn, the executive director of the Surveillance Technology Oversight Project, a nonprofit organization in Manhattan.

As an example, he pointed to a law enacted by New York State this month that gives Gov. Andrew M. Cuomo unlimited authority to rule by executive order during state crises like pandemics and hurricanes. The law allows him to issue emergency response directives that could overrule any local regulations.



In Lombardy, the Italian authorities are using cellphone location data to determine what percentage of people are obeying a lockdown order. Alessandro Grassani for The New York Times

Increased surveillance and health data disclosures have also drastically eroded people's ability to keep their health status private.

This month, Australia's health minister publicly chastised a doctor whom she accused of treating patients while experiencing symptoms of the virus — essentially outing him by naming the small clinic in Victoria where he worked with a handful of other physicians.

The health provider, who tested positive for the coronavirus, responded with a Facebook post saying the minister had incorrectly characterized his actions for political gain and demanded an apology.

“That could extend to anyone, to suddenly have the status of your health blasted out to thousands or potentially millions of people,” said Chris Gilliard, an independent privacy scholar based in the Detroit area. “It's a very strange thing to do because, in the alleged interest of public health, you are actually endangering people.”

But in emergencies like pandemics, privacy must be weighed against other considerations, like saving lives, said Mila Romanoff, data and governance lead for United Nations Global Pulse, a U.N. program that has studied using data to improve emergency responses to epidemics like Ebola and dengue fever.

“We need to have a framework that would allow companies and public authorities to cooperate, to enable proper response for the public good,” Ms. Romanoff said. To reduce the risk that coronavirus surveillance efforts might violate people's privacy, she said, governments and companies should limit the collection and use of data to only what is needed. “The challenge is,” she added, “how much data is enough?”

New software in China decides whether people should be quarantined or permitted to enter public places like subways. Green means a person is at liberty to go out. Raymond Zhong

The fast pace of the pandemic, however, is prompting governments to put in place a patchwork of digital surveillance measures in the name of their own interests, with little international coordination on how appropriate or effective they are.

Latest Updates: Economy

- [Wall Street is set for an upbeat start to the week.](#)
- [Here's the business news to watch this week.](#)
- [The Japanese economy enters its first recession since 2015.](#)

[See more updates](#)

Updated 26m ago

More live coverage: [Global](#) [U.S.](#) [New York](#)

In hundreds of cities in China, the government is requiring citizens to use software on their phones that automatically classifies each person with a color code — red, yellow or green — indicating contagion risk. The software determines which people should be quarantined or permitted to enter public places like subways. But officials have not explained how the system makes such decisions, and citizens have felt powerless to challenge it.

In Singapore, the Ministry of Health has posted information online about each coronavirus patient, often in stunning detail, including relationships to other patients. The idea is to warn individuals who may have crossed paths with them, as well as alert the public to potentially infected locations. “Case 219 is a 30-year-old male,” says one entry on the Health Ministry’s site, who worked at the “Sengkang Fire Station (50 Buangkok Drive),” is “in an isolation room at Sengkang General Hospital” and “is a family member of Case 236.”

In Singapore, the Ministry of Health has posted information online about each coronavirus patient, often in stunning detail. Adam Dean for The New York Times

On Friday, Singapore also introduced a smartphone app for citizens to help the authorities locate people who may have been exposed to the virus. The app, called TraceTogether, uses Bluetooth signals to detect mobile phones that are nearby. If an app user later tests positive for the virus, the health authorities may examine the data logs from the app to find people who crossed their paths. A government official said the app preserved privacy by not revealing users' identities to one another.

In Mexico, after public health officials notified Uber about a passenger infected with the virus, the company suspended the accounts of two drivers who had given him rides, along with more than 200 passengers who had ridden with those drivers.

In the United States, the White House recently spoke with Google, Facebook and other tech companies about potentially using aggregated location data captured from Americans' mobile phones for public health surveillance of the virus. Several members of Congress subsequently wrote a letter urging President Trump and Vice President Mike Pence to protect any virus-related data that companies collected from Americans.

The digital dictates may enable governments to exert more social control and enforce social distancing during the pandemic. They also raise questions about when surveillance may go too far.

In January, South Korean authorities began posting detailed location histories on each person who tested positive for the coronavirus. The site has included a wealth of information — such as details about when people left for work, whether they wore masks in the subway, the name of the stations where they changed trains, the massage parlors and karaoke bars they frequented and the names of the clinics where they were tested for the virus.

In South Korea's highly wired society, however, internet mobs exploited patient data disclosed by the government site to identify people by name and hound them.

Internet mobs used patient data disclosed by the South Korean government to identify people by name and hound them. Jean Chung for The New York Times

As other countries increase surveillance, South Korea had an unusual reaction. Concerned that privacy invasions might discourage citizens from getting tested for the virus, health officials announced this month that they would refine their data-sharing guidelines to minimize patient risk.

“We will balance the value of protecting individual human rights and privacy and the value of upholding public interest in preventing mass infections,” said Jung Eun-kyeong, the director of South Korea’s Centers for Disease Control and Prevention.

That is a tricky balance that some United States officials may need to consider.

In New York this month, Mayor Bill de Blasio posted details on Twitter about a lawyer in Westchester County who was the second person in the state to test positive for the virus — including the name of the man’s seven-person law firm and the names of the schools attended by two of his children. A few hours later, The New York Post identified the lawyer by name and was soon referring to him as “patient zero” in the coronavirus outbreak in New Rochelle.

In a response posted on Facebook, Adina Lewis Garbuz, a lawyer who is the wife of the man, Lawrence Garbuz, pleaded with the public to focus instead on the personal efforts the family had made to isolate themselves and notify people who came into contact with them.

“We would have preferred this all remain private,” Ms. Garbuz wrote in the Facebook post, “but since it is no longer, I wanted to at least share some truths and allay people’s fears.”

Aaron Krolik and Adam Satariano contributed research.

How My Boss Monitors Me While I Work From Home

As we shelter in place in the pandemic, more employers are using software to track our work — and us.

By Adam Satariano

Published May 6, 2020 Updated May 7, 2020

LONDON — On April 23, I started work at 8:49 a.m., reading and responding to emails, browsing the news and scrolling Twitter. At 9:14 a.m., I made changes to an upcoming story and read through interview notes. By 10:09 a.m., work momentum lost, I read about the Irish village where Matt Damon was living out the quarantine.

All of these details — from the websites I visited to my GPS coordinates — were available for my boss to review.

Here's why: With millions of us working from home in the coronavirus pandemic, companies are hunting for ways to ensure that we are doing what we are supposed to. Demand has surged for software that can monitor employees, with programs tracking the words we type, snapping pictures with our computer cameras and giving our managers rankings of who is spending too much time on Facebook and not enough on Excel.

The technology raises thorny privacy questions about where employers draw the line between maintaining productivity from a homebound work force and

creepy surveillance. To try to answer them, I turned the spylike software on myself.

Last month, I downloaded employee-monitoring software made by Hubstaff, an Indianapolis company. Every few minutes, it snapped a screenshot of the websites I browsed, the documents I was writing and the social media sites I visited. From my phone, it mapped where I went, including a two-hour bike ride that I took around Battersea Park with my kids in the middle of one workday. (Whoops.)

To complete the experiment, I gave my editor, Pui-Wing Tam, the keys to the Hubstaff program so she could track me. After three weeks of digital monitoring, the future of work surveillance seemed to both of us to be overly intrusive. As she put it, “Ick.”

Week 1

Adam: I downloaded Hubstaff to my laptop and phone with more than a touch of skepticism. I had heard of this type of tool being used by Wall Street firms for years, mainly in the name of security, with employees rarely having any say about how they were being watched.

Dave Nevogt, a founder and the chief executive of Hubstaff, who gave me a free trial to test its subscription software, said work-from-home orders in the coronavirus outbreak had made employee-monitoring software a hot ticket. Trials of Hubstaff software, which cost \$7 to \$20 a month per user, have tripled since March, he said.

“The world is changing,” Mr. Nevogt told me. Workers know they are being watched, so it does not violate privacy, he added.

One main feature of Hubstaff is an activity monitor that gives managers a snapshot of what an employee is doing. Broken down in 10-minute increments, the system tallies what percentage of time the worker has been typing or moving the computer mouse. That percentage acts as a productivity score.

I tried to embrace the feedback. Each day, an email was sent to me and Pui-Wing with an overview of my day: hours worked, the productivity score, and the websites and apps that I was using.

One day last month, when I was putting the finishing touches on an article, I spent 3 hours and 35 minutes editing the document, and an hour inside a file holding background research and interview notes. Another 90 minutes were spent on email.

This was one of my more productive days, but the software still tallied my deviations. It showed I was on Twitter for 35 minutes and lost 11 minutes browsing Spotify. Slack, the collaboration tool, swallowed 22 minutes. Other days, food was a common distraction, including one 10-minute hunt for takeout pizza.

Hubstaff also logged my GPS coordinates, a feature that Mr. Nevogt said was mainly used by companies trying to ensure that their salespeople were visiting clients. Given that London has been on lockdown since late March, I had few movements to track. The software mainly caught me jogging around a nearby park. And going to a wine shop.

Week 2

Adam: Once accustomed to life under surveillance, I made the questionable decision of letting Pui-Wing have access.

Latest Updates: Economy

- [Asia markets rise as investors look for silver linings amid the pandemic.](#)
- [The coronavirus dragged weakened Japanese economy into a recession for first time since 2015.](#)
- [Apple plans to reopen some stores this week, but will require distancing and masks.](#)

[See more updates](#)

Updated 8m ago

More live coverage: [Global](#) [U.S.](#) [New York](#)

“You are agreeing not to fire, judge or blackmail me for whatever this turns up,” I wrote to her in an email beforehand.

Pui-Wing: I was curious, I admit it. But also reluctant because do we really want to see someone’s minute-by-minute location or how often he or she uses Twitter?

With those misgivings, I opened the program and saw a dashboard. It showed various categories, including screenshots of Adam’s computer, his time sheets, apps and URLs he had visited and his whereabouts.

I clicked on screenshots and saw that Adam had been online for 9 hours 42 minutes 17 seconds the previous day. The dozens of screenshots included those of a Google Meet conference call that Adam had participated in, which displayed as extremely close-up photos of the faces of numerous colleagues.

I quickly retreated to the main dashboard. There I saw that Adam’s activity for the week was at a somewhat disappointing 45 percent. He later explained that the number didn’t accurately reflect his time spent working because it logged only when he was typing, not when he was making phone calls or doing other work away from his computer. Right.

Adam: For employers nervous about wasting money in a shaky economy, I could see Hubstaff's appeal. Mr. Nevogt introduced me to Chris Heuwetter, who runs a marketing company in Jupiter, Fla., called 98 Buck Social.

Mr. Heuwetter said he had seen work hours collapse after he let his 20 employees work from home in the virus outbreak. The company was facing a drop in sales, but Mr. Heuwetter said some employees did not start responding to messages until after 10 a.m. Responses to customer questions also slowed.

So he began using Hubstaff on March 31. Once he did, he said, his employees' productivity levels rose "immediately."

I could relate. Hubstaff was starting to affect my behavior. Each day, I logged in early because it was keeping a running clock of my activity. Knowing my online actions could be reviewed, I did not spend (as much) time reading about sports and rarely opened messaging apps on my laptop, nervous about a screenshot catching a private exchange.

But my activity scores stayed stubbornly low, usually from 30 percent to 45 percent. On April 14, Hubstaff showed that I worked for nearly 14 hours but had a productivity score of 22 percent.

Week 3

Adam: The moment when I no longer wanted to be monitored came on April 23 at 11:30 a.m., when Hubstaff caught me doing an internet exercise class. By the time I realized I had not logged out, it had snapped a screenshot of the trainer setting up to teach the class in her living room.

Even though this was just an experiment, it didn't make it any less embarrassing and intrusive. And it goes beyond being caught exercising in the middle of the day. What if other screenshots exposed sensitive health or

financial information?

I trust Pui-Wing, but the monitoring systems have few safeguards to prevent abuse, and they rely on managers exercising judgment and restraint.

Pui-Wing: Fortunately, I did not see Adam's internet exercise class. After poking around the Hubstaff metrics, it was clear it did not capture when he was reporting and talking to sources. It was thus irrelevant — at least to how we work.

Also, did I mention it was yucky to see so much of someone's information? I didn't log back in.

Occasionally, I glanced at the daily emails that Hubstaff sent about Adam. They showed his productivity score at 30 percent, sometimes edging up to 50 percent. I chuckled when I noticed that he began spending more time on news websites as his behavior changed.

Adam: By the end, I found myself trying to cheat the Hubstaff system altogether. As I write this at 11:38 a.m. on April 24, I am about to get some coffee and spend time with my cooped-up kids. But I plan to leave a Google Doc open on my computer that Hubstaff can screenshot to make it look like I was doing work.

Even if my editor says she isn't looking. Just, you know, in case.

Pui-Wing Tam contributed reporting (and surveilling) from San Francisco.

Adam Satariano is a technology reporter based in London. @satariano

A version of this article appears in print on May 7, 2020, Section B, Page 1 of the New York edition with the headline: How My Boss Monitors Me While I Work From Home

7. Tools — Due: project preparation

References:

Williams, Joseph M. (1990), “Ten Principles for Writing Clearly,” in *Style: Toward Clarity and Grace*, University of Chicago Press, Chicago (from the back cover of the 2003 ed.).

Ten Principles for Writing Clearly

1. Distinguish real grammatical rules from folklore.
2. Use subjects to name the character's in your story.
3. Use verbs to name their important actions.
4. Open your sentences with familiar units of information.
5. Begin sentences constituting a passage with consistent topic/subjects.
6. Get to the main verb quickly:
 - Avoid long introductory phrases and clauses.
 - Avoid long abstract subjects.
 - Avoid interrupting the subject-verb connection.
7. Push new, complex units of information to the end of the sentence.
8. Be concise:
 - Cut meaningless and repeated words and obvious implications.
 - Put the meaning of phrases into one or two words.
 - Prefer affirmative sentences to negative ones.
9. Control sprawl.
 - Don't tack more than one subordinate clause onto another;
 - Extend a sentence with resumptive, summative, and free modifiers.
 - Extend a sentence with coordinate structures after verbs.
10. Above all, write to others as you would have others write to you.

(Source: WILLIAMS, Joseph M. (1990), *Style: Toward Clarity and Grace*, Chicago University Press, Chicago, inside cover, ed. 2003).