BABY GALAXIES • ALTERNATIVES TO EUTHANASIA • MAKING AVIATION SAFER

SCIENTIFIC AMERICAN

MAY 1997 $4.95

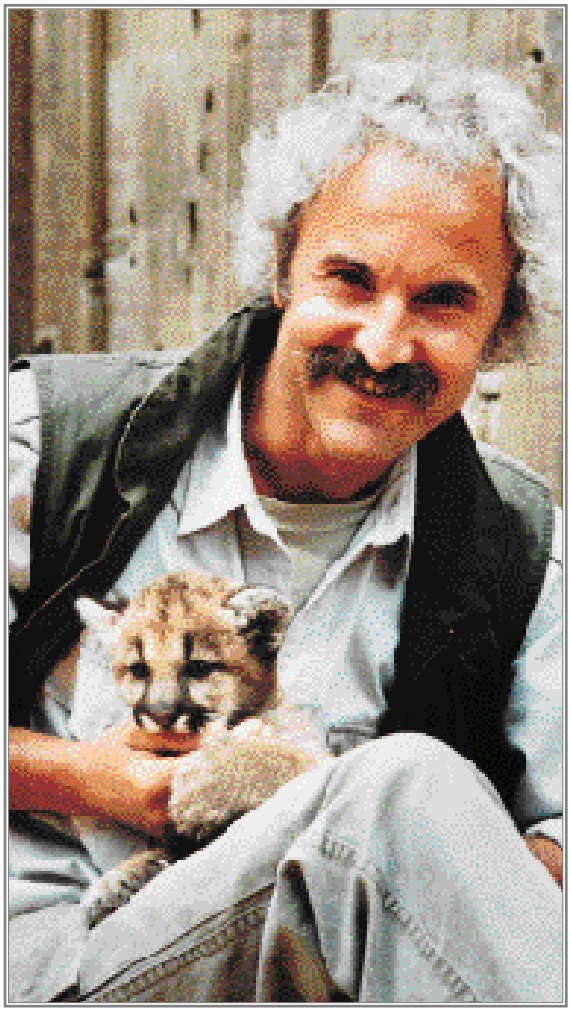
THE 100-YEAR WEATHER FORECAST
PREDICTING WHAT A WARMER CLIMATE WILL REALLY MEAN

The king of beasts masters the politics of survival

Lions seem like the archetypal social animals, working together toward a common goal—such as their next meal. But after many years observing these creatures in the wild, we have a less exalted view.…

—Craig Packer and Anne E. Pusey, *page 52*

The Artist and the Lion's Tale



STEVE MARTIN

Carl Brenders

The artist behind this month's cover of *Scientific American* is Carl Brenders, acclaimed around the world as one of the premier painters of wildlife. The almost photographic realism of his paintings, with its meticulous devotion to anatomical detail, emerges from Brenders's conservationist philosophy that nature is itself perfect. "That is why I paint the way I do," he says. "I want to capture that perfection."

Brenders, who was born and trained in art in Belgium, typically begins his work with extensive field research into the habits and habitats of his wildlife subjects. It was while on a trip to the Kalahari Desert in Botswana that he began tracking and gathering information about lions and their environment. Based on his observations, Brenders created a pencil sketch of a lion (*shown below*) and the *Kalahari* painting (*cover*) in watercolors and gouache, using techniques of his own invention.

Recently Brenders was honored as the Featured Artist at the 1997 Southeastern Wildlife Exposition in Charleston, S.C. A retrospective exhibition of 30 of his works is now in progress at the Carnegie Museum of Natural History in Pittsburgh, Pa. (February 1 through May 18). Other examples of his artwork can be found in the book *Wildlife: The Nature Paintings of Carl Brenders* (published by Harry N. Abrams, 1994) and in the series of limited edition art prints published by Mill Pond Press (Venice, Fla., 1-800-535-0331). *—The Editors*

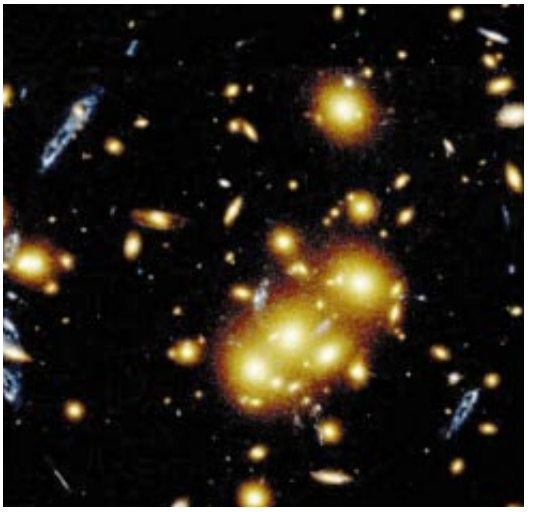
On the Cover:
Detail from *Kalahari,* a mixed media painting by Belgian artist Carl Brenders © 1997. Art courtesy of the artist and Mill Pond Press, Inc.

SCIENTIFIC AMERICAN

May 1997 Volume 276 Number 5

NEWS AND ANALYSIS

Climatologists have concluded that because of the greenhouse effect and other influences, the world will grow a few degrees warmer in the next century. Yet simplistic predictions that scorching summers, more cyclones and heavier rainfall will therefore follow can be far off the mark. These experts offer a more realistic view.



Travel back in time for a few billion years, courtesy of high-powered telescopes, and the universe looks like a very different place. Once it was exceedingly hot, dense and uniform; now it is relatively cool and empty. By peering at the earliest, most distant galaxies, astronomers are learning how this transformation occurred.

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111. Periodicals postage paid at New York, N.Y., and at additional mailing offices. Canada Post International Publications Mail (Canadian Distribution) Sales Agreement No. 242764. Canadian BN No. 127387652RT; QST No. Q1015332537. Subscription rates: one year $34.97 (outside U.S. $47). Institutional price: one year $39.95 (outside U.S. $50.95). Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537. Reprints available: write Reprint Department, Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111; fax: (212) 355-0408 or send e-mail to info@sciam.com Visit our World Wide Web site at http://www.sciam.com/ **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 247-7631.**

About the Cover
This painting is the first of a lion by wildlife artist Carl Brenders. For more information about Brenders and his work, please see the inside of the cover flap.

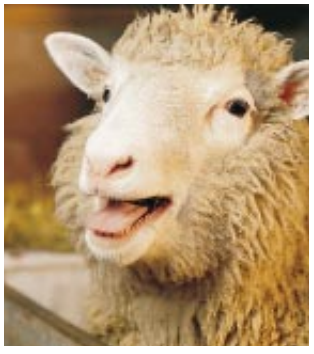
FROM THE EDITORS

The Misunderstood Clone

Oh, give me a clone
Of my own flesh and bone
With its Y chromosome changed to X
And when it is grown
Then my own little clone
Will be of the opposite sex.

The late, great Isaac Asimov co-authored that doggerel with Randall Garrett decades ago, but it fits today with the general giddiness over mammalian cloning. Jokes about cloned sheep and virgin wool abound. Associate editor Tim Beardsley assesses some of the more sobering aspects in his news story beginning on page 15.

It is worth pausing to review everything that cloning *isn't*. First and foremost, it is not a process for making exact copies of grown people. My clones and I would be no more alike and probably less than any identical twins. To strip away cloning's mystique, remember that it was originally a horticultural term ("clone" derives from the Greek word for "twig"). Any gardener who has planted a clipping and seen it take root has cloning credentials. No one expects a cloned rosebush to be a carbon copy of its parent down to the arrangement of the thorns, so it would be equally wrong to expect human clones to match up in the infinite variety of personal characteristics.



NAJLAH FEANNY *SABA*

SHEEPISH GRIN over cloning confusion.

Second, cloning is not yet a technology ready for use on human cells. But because the techniques needed to accomplish cloning are simple as far as biomedical miracles go, it seems all but certain that some clinic or laboratory will quietly start trying at any moment. Yet rushing to human experiments could be tragic.

Finally, even when cloning of humans is safe, it isn't necessarily going to be popular. Cloning won't replace the old style of reproduction: it's not as much fun, and it's a lot more expensive. Cloning commercially valuable animals makes perfect economic sense—it is a potentially surer thing than breeding. Granted, you can't put a price on vanity, so the idea will appeal to people with excesses of cash and ego. Still, most of us will probably eat a cloned mammal before we shake hands with one.

Speaking of mammals, the majestic lion featured on our cover has been greatly admired by people around our office. In response, SCIENTIFIC AMERICAN has decided to make available a limited edition of numbered art prints of Carl Brenders's painting *Kalahari,* signed by the artist. For further information, you are welcome to call 1-800-777-0444.

JOHN RENNIE, *Editor in Chief*
editors@sciam.com

LETTERS TO THE EDITORS

IGNITING CONTROVERSY

I was disappointed with John Horgan's article on the National Ignition Facility (NIF) ["Beyond the Test Ban," News and Analysis, December 1996]. Since 1990, four major committees have reviewed the NIF—the National Academy of Sciences, the Fusion Policy Advisory Committee, Jasons, and the Inertial Confinement Fusion Advisory Committee. All four have strongly supported the NIF. Most technical experts have, thus far, judged the NIF to be an excellent window into the physics of nuclear weapons testing. Even the advocates of other approaches for inertial fusion energy largely support the NIF because it is the only near-term method for providing fundamental information on the basic physics of the fusion process.

DAVID H. CRANDALL
Director, Office of Inertial Fusion and the National Ignition Facility Project
U.S. Department of Energy

FACING THE CRITICS

I feel I must respond to George Styx's [sic] analysis of the Institute for Creation Research that appeared in the series "Science versus Antiscience?" [Trends in Society, January]. Styx's article "Postdiluvian Science" did a disservice to readers by reinforcing previous misrepresentations of creationist thinking and ignoring some major trends in science. Indeed, Styx missed a golden opportunity. This is a momentous time in origins science. The more we learn of life, even microscopic life, the more we see design and order on an elegant level, impelling us to the conclusion that the universe was created. As a result, many evolution professors are forsaking naturalism. Some are becoming creationists. Most are gravitating to illogical New Age thinking—the Gaia hypothesis—that Mother Nature is alive and doing this on purpose. Creationist thinking is not a threat to science. It is a persuasive challenge to a sterile naturalistic religion posing as science.

JOHN D. MORRIS
President, Institute for Creation Research
El Cajon, Calif.

I believe you have severely underestimated the importance of the public's growing acceptance of pseudoscientific claims. By critiquing creationism, feminist science and interest in the paranormal in only a very general way, you have failed to highlight the most significant trends in current New Age culture. Schools and law-enforcement agencies have spent tens of thousands of public dollars to purchase dowsing rods to locate drugs in high school lockers. Public defenders have hired psychics to "read" the auras of prospective jurors. Medical insurance plans are beginning to cover numerous unproved homeopathic and other junk remedies. We all share in these costs. As the introduction to "Science versus Antiscience?" articulated, belief in the supernatural is not new. But these modern examples are different: corporations and public institutions are beginning to entrench such beliefs in their decision-making processes, their policies and their actions.

DOUG FRASER
Haileybury, Ontario

In his article "The Media's Eerie Fascination," Philip Yam concludes with the optimistic view that better science education might create a generation of more skeptical, science-savvy citizens. As a science teacher, I certainly hope so. But where will this education come from? Most teachers at the primary level receive nothing but the most rudimentary introduction to science. We must demand that our children study science and its methods throughout their education; we must also produce teachers who are thoroughly trained in science, who can answer a child's simple (but often profound) questions about nature without feeling intimidated or uncomfortable. If professional scientists disdain to present science to the general public, we will continue to pay the price for this snobbishness. Pseudoscience will prevail by default.

WAYNE R. ANDERSON
Sacramento City College

The Editors reply:

With all respect, Morris seems to have an exaggerated impression of how many mainstream scientists are persuaded that creationism is a convincing or even valid alternative to evolution. (Incidentally, our writer's name is Gary Stix, not George Styx.) The creep of irrationality into public institutions is deplorable and vexing, as Fraser says. Our point was only that it is hard to document clearly that those institutions are more prone toward nuttiness than in the past. And we agree 100 percent with Anderson: much effort and enthusiasm need to go into teaching science more effectively.

THE ONCE AND FUTURE CHAMP

In the article "Understanding Parkinson's Disease," by Moussa B. H. Youdim and Peter Riederer [January], the caption under Muhammad Ali's photograph refers to him as a "once indomitable athlete." I would say that Ali's very presence at the Olympics last summer, his ongoing appearances in public despite his disease and his continuing work to help others are clear proof of his current indomitable spirit and courage—no "once" about it. Ali fights a different battle today, but he remains "The Champ."

GREG GUERIN
Tempe, Ariz.


CORBIS-BETTMANN

Indomitable Ali

Letters selected for publication may be edited for length and clarity.

CLARIFICATIONS

Despite recent maneuvers, Pioneer 10's signal remained sufficiently feeble that instead of collecting data ["In Brief," April], the 25-year-old probe was retired in March. The image on the cover of the January issue, showing turbulent flow around a golf ball, was based on a photograph by F.N.M. Brown.

50, 100 AND 150 YEARS AGO

SCIENTIFIC AMERICAN

MAY 1947

Dr. Felix Bloch at Stanford University is working on new methods of analyzing materials, using the frequency and magnetic reaction of atom nuclei. Based on the principle that the atom nucleus of every element has a characteristic, precise frequency to which it resonates under the influence of radio-frequency current in a magnetic field, the experiments consist of placing test materials in the field of a powerful electro-magnet. The radio-frequency current is induced into the nuclei, and a sensitive receiving set determines the frequency given off by the nuclei. This frequency gives the key to the composition of the material." [*Editors' note: Bloch and Edward M. Purcell of Harvard University won the 1952 Nobel Prize for Physics for their work with nuclear magnetic resonance.*]

Special effects in the service of grand opera, 1897

"Modern commercial radar equipment is being installed in eight Boeing Stratocruiser luxury airliners now under construction for American Overseas Airlines. Storm areas and regions of dangerous icing will be revealed by a radar antenna in the nose, pointing forward, and shorelines will be mapped from many miles out to sea by a 60-inch-diameter antenna in the belly, pointing downward."

MAY 1897

How is the temperature of the sun maintained? Helmholtz suggested in about 1853 that the sun's heat is maintained by its slow shrinkage. Suppose I drop a book on the floor, what happens? Gravity acts upon it, with a little noise; but the main thing is motion has been produced and has been stopped, and a certain amount of heat unquestionably produced. Suppose every portion of the sun's surface drops 150 feet toward the sun's center, diminishing its diameter by about 300 feet; that would account for all the heat the sun sends forth. A yearly shrinkage of 300 feet in diameter would have to go on for 7,000 years before detection by the best telescopes that we or our posterity are likely to possess."

"The visible sign of cobwebs and dust on a bottle of wine used to be taken as convincing evidence of age. Unfortunately, the Division of Entomology of the U.S. Department of Agriculture says that an industry has recently sprung up which consists of farming spiders for the purpose of stocking wine cellars, and thus securing a coating of cobwebs to new wine bottles, giving them the appearance of great age."

"A case in a New York court where an owner, suing for damages from a railroad company for injury done his property by the noise of passing trains, sought to introduce the phonograph, and thus give to the court direct and practical evidence of the sound vibrations caused by the locomotives and cars as they were propagated in the apartments of the plaintiff. The court has held open the admissibility or non-admissibility of such evidence for further consideration."

"Fafner the dragon, in 'Siegfried,' is one of the most interesting properties at the Metropolitan Opera House in New York. It is thirty feet long, made of papier maché and cloth and is painted in shades of green. The jaw, tongue and antennae are all movable. The head is supported by one man and is moved by a second man. A hose runs from offstage through the tail and the body to the mouth, and carries the steam for the sulfurous breath of the terrible monster; the eyes are provided with electric lights. Our illustration shows the dragon standing in the mouth of the cave, belching forth steam, the eyes gleaming fitfully."

MAY 1847

It is ascertained that the planets, like our own, roll in regulated periods around the sun, have nights and days, are provided with atmosphere, supporting clouds, and agitated by winds. Notwithstanding the dense atmosphere and thick clouds with which Venus and Mercury are constantly enveloped, the telescope has exhibited to us great irregularities on their surfaces, and thus proved the existence of mountains and valleys. On Mars, the geographical outlines of land and water have been made apparent, and in its long polar winters snows accumulate in the desolation of the higher latitudes."

"A number of cabs with newly invented wheels have just been put on the road in London. Their novelty consists in the entire absence of springs. A hollow tube of India rubber about a foot in diameter, inflated with air, encircles each wheel in the manner of a tire, and with this simple but novel appendage the vehicle glides noiselessly along, affording the greatest possible amount of cab comfort to the passenger."

NEWS AND ANALYSIS

IN FOCUS

THE START OF SOMETHING BIG?

Dolly has become a new icon for science

NAJLAH FEANNY *SABA*

DOLLY, THE FIRST CLONE OF AN ADULT MAMMAL, poses at the Roslin Institute near Edinburgh, Scotland.

It was supposed to be impossible. When Ian Wilmut, Keith H. S. Campbell and their colleagues at the Roslin Institute near Edinburgh, Scotland, announced in February that they had cloned an adult sheep to create a lamb with no father, they did not merely stun a world unprepared to contemplate human virgin births. They also startled a generation of researchers who had grown to believe, through many failed experiments, that cells from adult animals cannot be reprogrammed to make a whole new body. Dolly, the lamb at the epicenter of the culture-shock waves, developed from a sheep egg whose original nucleus had been replaced by a nucleus from an adult ewe's udder. By starving the donor cells for five days before extracting their nuclei, Wilmut and Campbell made the nuclear DNA susceptible to being reprogrammed once placed in an egg.

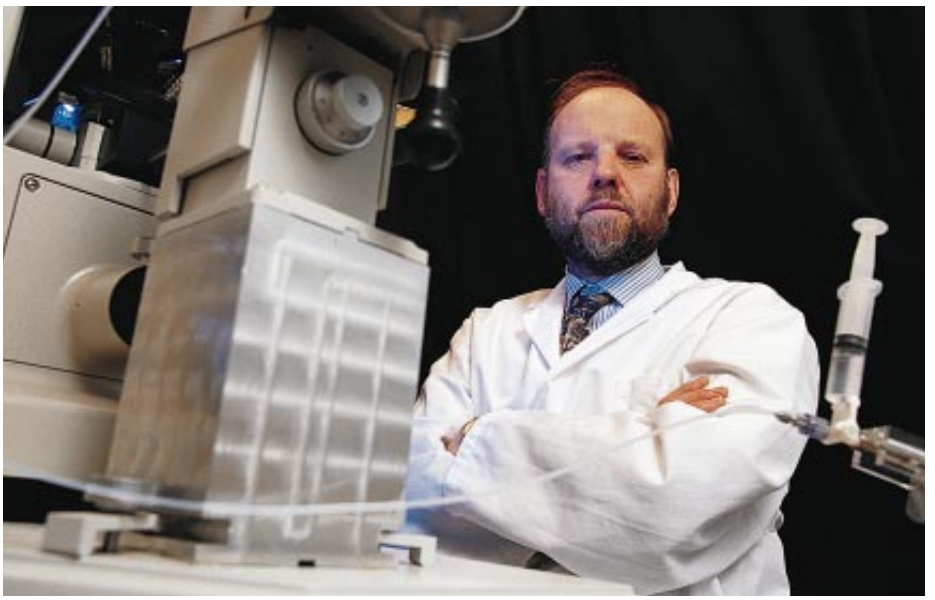
Dolly's birth thus represents an ethical and scientific watershed. Around the world, advisory committees and legislators are frantically trying to decide whether and when it might be ethical to duplicate the feat in humans. Traditional teachings that life begins at conception suddenly seem to be missing the point. "We have to rid our minds of artificial divides," says Patricia King of Georgetown University. President Bill Clinton quickly announced a ban on the use of federal funds for human cloning research and asked the National Bioethics Advisory Commission to recommend some actions.

Many animal development experts now suspect that genetically duplicating humans is possible, especially as Donald Wolf of the Oregon Regional Primate Research Center has already cloned rhesus monkeys from embryonic cells. (Cows, sheep and rabbits have also been cloned from embryonic cells in recent years, but these experiments lacked the emotional impact of a copied mature animal.) Indeed, it took less than two weeks from the date of the Roslin Institute's announcement in *Nature* for Valiant Ventures in the Bahamas to an-

nounce that it will build a laboratory to clone people willing to pay. The company was founded for the purpose by the Raëlian Movement, a self-styled religious organization.

But producing healthy human clones may prove to be extremely difficult. Wilmut, who argues for a moratorium on such attempts, points out that more than half the cloned sheep pregnancies he initiated failed to develop to term. Some had abnormalities. "People have overlooked that three out of eight [cloned] lambs died soon after birth" in an earlier study, he notes. Moreover, it took 277 attempts to produce Dolly from an adult cell.

Should Valiant Ventures's plans ever come to fruition, they would probably produce many unhappy customers and some dead babies before they created a live one. That grim scenario prompts bioethicist Arthur Caplan of the University of Pennsylvania to argue that anyone attempting such a project "ought to be arrested." He predicts that a moratorium will be enforced by government officials. (Such restrictions might spare egotistical millionaires the disappointment of learning that cloned offspring can be just as hard to handle as natural ones.)



NAJLAH FEANNY *SABA*

IAN WILMUT
led the team that cloned sheep, first from embryos and now from a ewe.

Wilmut concurs that there are no ethical grounds to justify duplicating existing humans. He even opposes allowing a couple to copy a child in order to get a source of tissue to save its life (although some years ago a California couple conceived a child in the time-honored manner to supply bone marrow for a sibling). The only human cloning Wilmut would condone is copying an embryo to avoid genetic disease caused by mutations in mitochondria, DNA-bearing structures lying outside cell nuclei. Mutations of mitochondrial DNA can cause devastating afflictions, including blindness. By implanting a nucleus from an embryo with defective mitochondria into an egg donated by a woman with healthy mitochondria, researchers could help a couple have a child free from mitochondrial disease.

Other bioethicists are more receptive to copying people. John C. Fletcher of the University of Virginia believes that society might find it acceptable for a couple to replace a dying child or for a couple with an infertile partner to clone a child from either partner. "I am not scared of cloning," Fletcher declares. The widespread squeamishness toward embryo research suggests, however, that Fletcher may for now be in a minority.

Four years ago the revelation that researchers at George Washington University had divided genetically crippled human embryos provoked a national outcry—even though the investigators never contemplated implanting the multiple embryos into a uterus. Last year the National Institutes of Health terminated an employee who used federal equipment to perform genetic tests on cells from human embryos before implanting them, in violation of a congressional ban.

In the arena of animal husbandry and biomedicine, cloning could bring about big changes—provided the technique works in species other than sheep and can be made more efficient. "I have no doubt this will become the method of choice for producing transgenic animals," says James M. Robl of the University of Massachusetts. Transgenic, or genetically manipulated, animals are typically now made by a laborious hit-or-miss procedure that involves injecting genes into eggs and breeding the few animals that take up the genes. Cloning should expedite the rapid generation of large numbers of creatures with specific alterations, Robl believes.

Robl founded a company, Advanced Cell Technology, that plans to clone transgenic animals that will produce human proteins in their milk or supply tissue for transplants that human immune systems will not reject. (The Roslin Institute has a partnership with PPL Therapeutics, which will also produce animals that secrete human proteins.) And Robl foresees large gains for animal breeding in general. Experiments involving genetically identical clones, he explains, would involve fewer confounding variables and thus should be easier to interpret; moreover, fewer animals may be needed to produce the same results. Breeding programs to rescue endangered species might also become more effective. Cloning could sidestep some of the difficulties of sexual reproduction, although by limiting genetic diversity it might create its own problems.

Looking toward more distant shores, Dolly's existence raises the question of whether cells from patients can be reprogrammed to make genetically compatible therapeutic tissue, such as brain tissue of the type that is destroyed in Parkinson's disease. "The components needed for this kind of manipulation are out there," Robl speculates.

In the meantime, there is much to learn about the potential of genetic reprogramming. Nobody knows whether Dolly will live a healthy life, because her cells may in some respects behave like those of an animal six years old—the age of Dolly's parent when she was copied. It will be scientifically fascinating if Dolly develops strange and fatal afflictions in midlife. It will be even more fascinating if she does not.

—Tim Beardsley in Washington, D.C.

ASTRONOMY

VANISHING WORLD

Could the first planet discovered around a sunlike star be a mirage?

Recent reports of planets circling stars similar to the sun have sent imaginations reeling. Artists have crafted fanciful portraits of the new worlds; theorists have raced to account for the objects' bizarre properties; and everyone has delighted in speculating that maybe, just maybe, one of the planets could support life.

After years of false starts and retracted results, astronomers thought they had finally secured airtight proof that our solar system is not unique. Now it's déjà vu all over again, however, as David F. Gray of the University of Western Ontario has presented evidence that the first of these newfound planets, reportedly circling the star 51 Pegasi, does not really exist.

Gray's work underscores the precarious nature of the planet-hunting business. Ubiquitous science-fiction images notwithstanding, nobody has ever actually set eyes on a planet outside our solar system. All the reported planetary detections—at least eight by the latest count—depend on exceedingly subtle, indirect evidence.

When Michel Mayor and Didier Queloz of the Geneva Observatory examined 51 Pegasi, for instance, they noticed that the star's spectrum shifts slightly back and forth in a regular, 4.23-day period. This result was rapidly confirmed by Geoffrey W. Marcy and R. Paul Butler of San Francisco State University and the University of California at Berkeley, who have since become the leaders in finding new planets.

The two groups interpreted the spectral changes as a Doppler shift—a stretching or compression of the star's light caused by movement of the star. They concluded that a giant planet, at least half the mass of Jupiter, is orbiting 51 Pegasi and pulling it to and fro.

But Gray, who has been observing 51 Pegasi intermittently since 1989, was not convinced. In the February 27 issue of *Nature,* he describes a variation in the absorption lines of the spectrum of 51 Pegasi; the effect also has a 4.23-day period, and it cannot be explained by a planetary influence, Gray asserts. He suspects that the star's surface is oscillating in a manner "analogous to water sloshing in a basin." Those who saw a planet in the data, he says, "got carried away in a tide of enthusiasm."

Before Gray's paper even appeared in print, Mayor, Queloz, Marcy and Butler published a stinging rebuttal—without the delays of peer review—on the Internet. The planet hunters charge that Gray is the one chasing phantoms. "I don't think he has a real spectral signature," Marcy says, citing large errors and a good deal of scatter present in Gray's data points.

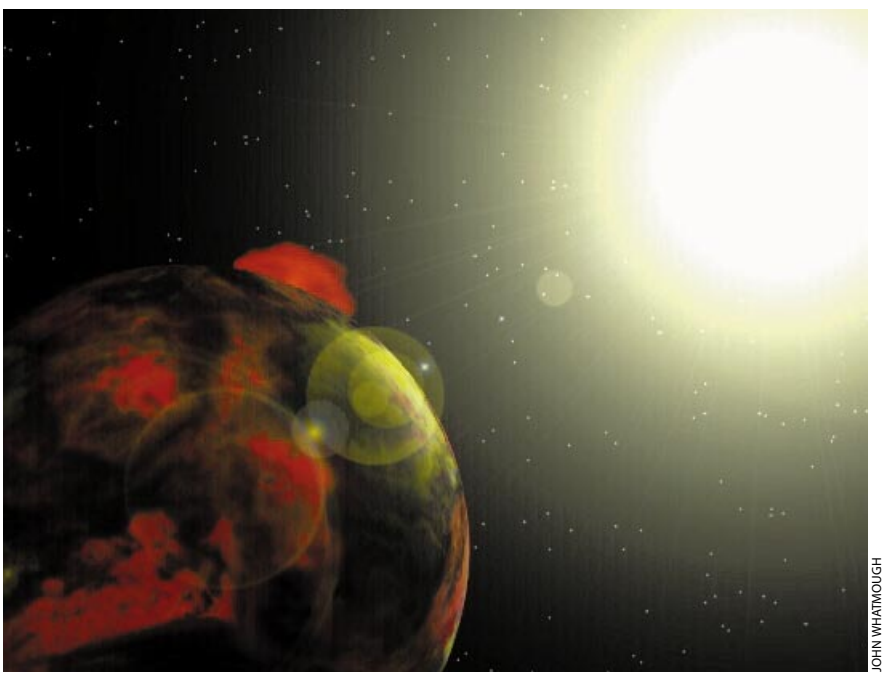
Marcy also assails the logic of Gray's interpretation. Oscillations should affect the star's brightness, Marcy notes, but "51 Peg is not showing brightness variations to one part in 5,000." Moreover, the kind of oscillation Gray proposes is unlike any yet seen or predicted. "That type of oscillation would be far more extraordinary, far more unexplainable, than the planet," Marcy concludes.

Emotion is clearly on the pro-planet side. Although Gray is only questioning the existence of one of the extrasolar planets, his paper has created the perception that he is a scientific Scrooge, snatching away a long-sought discovery. "Frankly, I cannot understand some of the vehement attacks on David Gray's work by some of my colleagues," says Artie Hatzes of the University of Texas, who is now collaborating with Gray on further analysis of 51 Pegasi.

Which side the science favors is not as obvious. Nobody takes Gray's paper lightly. Sallie Baliunas of the Harvard-Smithsonian Center for Astrophysics remarks that Gray's "observations are in general exquisite," although she sees "some problems with his analysis." She also faults his paper for omitting error bars and observation times, which would help other researchers assess his work.

Gordon Walker of the University of British Columbia, who wrote a commentary on Gray's report in *Nature,* agrees that it is preliminary. But Walker says the findings serve as an essential reminder that "stars are not clocks"—they vibrate, rotate and change in ways that can fool the unwary.

For his part, Gray seems slightly bemused by the fuss. "I'm not particularly interested in extrasolar planets," he explains, which is why he did not publish his studies of 51 Pegasi sooner. "I hate to say, 'Who cares?' but to me it was not terribly important." To an astronomer more attuned to the physics of stars



JOHN WHATMOUGH

PUTATIVE PLANET
orbiting the star 51 Pegasi, depicted here by an artist, may not exist.

IN BRIEF

Fowl Play

You can't judge a bird by its feathers. In a show of microsurgical mastery, Evan Balaban of the Neurosciences Institute in San Diego replaced certain brain cells in chicken embryos with like cells from developing quails. When the chimerical chickens hatched 19 days later, they displayed several astonishing, er, quailifications: some sounded like chicks but bobbed like quails, whereas others moved like chickens but sang three-note trills. The find demonstrates that hardwired behavior can be swapped between species and that the neuromechanisms behind many instincts are independent.

ROBERT DOWLING *Animals Animals*

Managing Migraines

Forget Excedrin. A new study presented at the American Academy of Neurology's annual meeting in April revealed that the drug sumatriptan can boost productivity in migraine sufferers by some 50 percent. The researchers gave either sumatriptan or a placebo to 132 volunteers experiencing headaches at work. Two hours later 79 percent of the treated individuals reported relief; only 32 percent of the control subjects felt better. Similarly, treated people lost on average only 86 minutes of work to migraine pain, but those given placebos missed as much as 168 minutes.

A New Take on Telomeres

Aging, it turns out, is not linked to shrinking telomeres—those non-sense stretches of DNA that cap off chromosomes. Because telomeres are not duplicated when a cell divides, scientists had presumed that telomeres continually shortened until the cell died. Immortal cancer cells, they noted, often bore extra long telomeres. The theory was compelling but wrong, several studies now demonstrate. In fact, telomeres appear to change lengths repeatedly. And these phases—from long to short and back again—have more to do with cell division than longevity.

More "In Brief" on page 24

than to the debris that may circle them, "a new oscillatory mode would be more exciting than some planet," he adds.

Regardless of their perspectives, all the participants are eager to settle the dispute. Fortunately, this is one scientific controversy that should not drag on indefinitely. In the coming weeks, astronomers around the world will focus their attention on Tau Bootis, another star with an alleged planet in a short-period (3.3-day) orbit, to see if it shows spectral variations like the ones Gray claims for 51 Pegasi.

Later in the year, when 51 Pegasi is again well placed for observation, both Marcy's and Gray's interpretations will be put to the test. And in a few months, high-precision measurements of stellar positions—a practice known as astrometry—should provide definitive measurements of the wobbly motions of three other stars that Marcy and Butler have reported as having planets. As yet, nobody has questioned those results.

What if even those planets vanish? Marcy's confidence does not waver: "If that happens, I'll take the #28 bus to the Golden Gate Bridge and take a swan dive." —*Corey S. Powell*

ORIGIN OF LIFE

THE SINISTER COSMOS

A meteorite yields clues to life's molecular handedness

This is, superficially, a right-handed world. Roughly nine out of 10 people eat, throw and write with their right hands. But in a deeper sense, we are all lefties. The amino acids of which we and all other known organisms are composed are left-handed.

Life's molecular handedness has long baffled biologists. Amino acids, which cells link together to build peptides and proteins, all come in two versions that, like a pair of gloves, mirror each other's shape. When amino acids are created in a laboratory, the batch is invariably racemic: it contains equal numbers of left- and right-handed molecules. (The original definition of left-handed molecules is that they make polarized light shone through them rotate to the left.) Presumably the same was true inside the earth's primordial ooze. So how did life set out on its sinister course?

Many evolutionary theorists believe chirality of one form or another was inevitable, because racemic chemistry would have been too inefficient for carrying out certain biological functions. Yet natural selection's choice of left-handed amino acids has been deemed simply a matter of chance. Now an analysis of a meteorite that crashed into Murchison, Australia, 28 years ago supports a different scenario. It suggests that "organic matter of extraterrestrial origin could have played an essential role" in nudging life down its left-handed path, as stated by John R. Cronin and Sandra Pizzarello of Arizona State University this past February in *Science*.

The Murchison meteorite has fascinated students of life's origins since 1970, when investigators discovered that the charcoal-hued rock is rich in amino acids and other complex organic compounds. That fact established that such molecules can be generated by nonbiological processes occurring beyond the earth and even beyond the solar system. But did that cosmic chemistry create an excess of left-handed amino acids in the meteorite? Initial studies said no; experiments more than 10 years later said yes. The latter findings were suspect, however, because of the possibility that the samples had been contaminated by terrestrial amino acids.

Cronin began pondering the mystery

CHIP CLARK

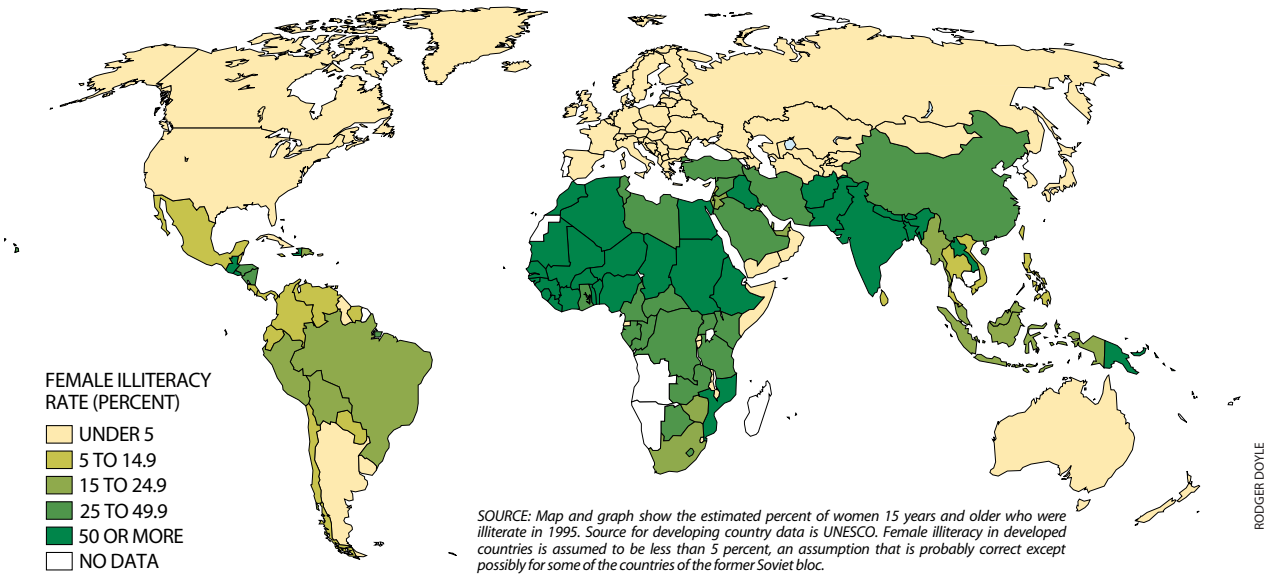
***MURCHISON METEORITE** has a slight excess of "left-handed" amino acids.*

of life's handedness in earnest several years ago. While teaching a class on chemical evolution, he encountered a hypothesis advanced in the 1980s by William A. Bonner of Stanford University and others. Bonner noted that spinning neutron stars are thought to emit electromagnetic radiation that propagates in corkscrew fashion from their poles. This radiation, Bonner speculated, could skew organic molecules toward left-handedness as they form.

"If there was anything to the idea," Cronin says, "the Murchison meteorite was the place to look." To rule out contamination, Cronin and Pizzarello focused on amino acids that occur rarely, if at all, on the earth. Because some

BY THE NUMBERS

Female Illiteracy Worldwide

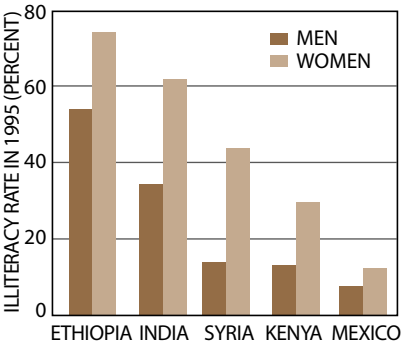


SOURCE: Map and graph show the estimated percent of women 15 years and older who were illiterate in 1995. Source for developing country data is UNESCO. Female illiteracy in developed countries is assumed to be less than 5 percent, an assumption that is probably correct except possibly for some of the countries of the former Soviet bloc.

RODGER DOYLE

In the history of literacy, the Protestant Reformation of the 16th century was a major turning point, for it gave women the first wide-scale opportunities to learn reading and writing. One premise of the radical Protestants, including Lutherans and Calvinists, was that everyone was entitled to read the Bible. Nowhere was this premise more apparent than in Lutheran Sweden, where in the late 17th century, a highly successful literacy program began to promote the Christian faith. The ability of women to read was vital because they were seen as the primary teachers of the young. The Protestant commitment to female literacy was evident in other places, such as Puritan New England, where women were more literate than their sisters in Europe.

The biggest surge in female literacy in Western countries occurred in the 19th century. By 1900 the overwhelming majority of women in several countries, including the U.S., France, England, and the more advanced parts of Germany and the Austrian empire, could read and write. Virtually all Western women are now literate, although a substantial minority have no more than a rudimentary skill, such as the ability to pick out facts in a brief newspaper article. (A 1992 study by the National Center for Education Statistics found that 17 percent of U.S. adults have only this rudimentary ability; 4 percent are unable to read at all. Illiteracy in the U.S. is probably no higher than in western Europe.)



Literacy statistics for most non-Western countries are problematic because there is no uniform worldwide method of measurement. Nevertheless, the map above is useful for highlighting some gross differences. The current major problem areas are in Asia, Africa and Central America. As in Western countries of earlier days, availability of schooling and the traditional notion about the sexual division of labor—the assignment of women to domestic tasks—are probably important factors. Another element, which applies particularly to Asia and Latin America, is the strict supervision by male family members of women's activities outside the home, which tends to inhibit the education of women. In almost all developing countries, women tend to be less literate than men, a circumstance illustrated in the chart below, which shows rates by gender for five typical countries. Literacy among women is associated with low fertility, low infant mortality and better health of children.

One trend that may portend a new literacy pattern is the growing education of women in Western countries, particularly in the U.S., where since 1980 they have increasingly outnumbered men on college campuses. This trend, which apparently reflects the rising aspirations of women, a decline in gender discrimination and the burgeoning need for women to earn a living, suggests that women may eventually surpass men in literacy sophistication. *—Rodger Doyle*

In Brief, *continued from page 18*

Leaping Lizards

Scientists from the National Museum of Natural History in Karlsruhe, Germany, the Royal Ontario Museum and the University of Toronto have dug up a complete skeleton of the oldest flying reptile ever found. It appears that this Late Permian creature, a *Coelurosauravus jaekele,* relied on curved, airfoil-like wings for flight, as do modern-day geckos. Unlike other prehistoric tetrapods, *C. jaekele* had no internal support for its gliding membrane; instead support came from bony rods, placed like battens on the skin.

HANS-DIETER SUES
Royal Ontario Museum

Elusive Leptoquarks?

The latest subatomic assault on the Standard Model comes from the DESY accelerator in Hamburg, Germany. There physicists recently reviewed data collected from millions of collisions involving one kind of lepton, called a positron, and protons, made up of quarks. Most often, the positron bounced off the quarks. In 12 instances, however, the positron made a U-turn and sped off with a surprising amount of energy. This abrupt about-face, the researchers say, may represent random fluctuations. But it may also indicate that a positron and a quark formed a fleeting leptoquark and quickly decayed. The quest for more concrete evidence of leptoquarks continues.

Snakes in Space

A massive, frozen lightning bolt, first seen in 1992, writhes like a snake in Sagittarius. Until recently, scientists knew only that the strange structure was some 150 light-years long, two to three light-years wide and had two giant kinks that shed powerful radio emissions. Now Gregory Benford of the University of California at Irvine has proposed that charged molecular clouds traveling through magnetic fields generate the Snake and similar filaments near the middle of our galaxy. The Snake wiggles, he suggests, because the magnetic force around it is too weak to contain it.

More "In Brief" on page 28

amino acids become racemic over time when exposed to water, the workers also chose molecules that do not react with water. The investigation turned up excesses of left-handed versions of four amino acids ranging from 7 to 9 percent.

Stanley Miller of the University of California at San Diego, a doyen of origin-of-life studies, calls Cronin and Pizzarello "very careful" researchers whose report must be taken seriously. If confirmed, he says, their work establishes that "nonbiological forces can create asymmetries [between left- and right-handed molecules] either on the earth or elsewhere."

But it is still unclear, Miller adds, how the relatively small asymmetries observed in the Murchison meteorite could have triggered the formation of terrestrial organisms composed exclusively of left-handed molecules. Pizzarello agrees that the research raises at least as many questions as it answers about life's murky beginnings. "Everything opens new doors," she says. *—John Horgan*

PSYCHOLOGY

AS TIME GOES BY…

You must remember this. Really

Remember that terrible fight with your best friend when you were seven years old? Be careful. In recent years, psychologists have shown that memories of long-ago events can be altered—intentionally or otherwise—by a psychotherapist or detective. The disturbing implications for criminal justice have stimulated scores of studies of "false-memory syndrome."

Curious to see just how difficult it is to muddle one's memories of reality and fantasy, psychologists Henry L. Roediger III and Kathleen B. McDermott of Washington University have been asking volunteers to remember words in specially constructed lists. They have discovered they can make most people remember—at least for a day—things that never happened. *Scientific American* here offers a bare-bones version of an experiment described by McDermott in the April 1996 issue of the *Journal of Memory and Language,* so that readers can produce robust false memories in their friends and family right in the convenience of the home.

First, recruit your victims by asking them only to participate in a five-minute test of learning—don't tip them off to the real purpose. Next, choose any three of the lists in the table below and read the words to the subject in a neutral voice, pausing for a moment or two between each word but continuing right from one list into the next. Do not read the words in the "unspoken target" column; those are for grading.

Having read all 45 words, ask your subjects to write down, in any order, every word they can clearly remember from those just heard. Allow four minutes, then pencils down. Guessing is *not* allowed. Now scan through your students' answers and see how many words from the unspoken column appear on the answer sheet.

In her study of 40 subjects, McDermott found that on average each volunteer correctly recalled fewer than 40

False-Memory Test

Read any three of these lists consecutively. Then check subject's recall for…	…these unspoken target words.
bed, rest, awake, tired, dream, wake, snooze, blanket, doze, slumber, snore, nap, peace, yawn, drowsy	sleep
nurse, sick, lawyer, medicine, health, hospital, dentist, physician, ill, patient, office, stethoscope, surgeon, clinic, cure	doctor
thread, pin, eye, sewing, sharp, point, prick, thimble, haystack, thorn, hurt, injection, syringe, cloth, knitting	needle
hot, snow, warm, winter, ice, wet, frigid, chilly, heat, weather, freeze, air, shiver, Arctic, frost	cold
apple, vegetable, orange, kiwi, citrus, ripe, pear, banana, berry, cherry, basket, juice, salad, bowl, cocktail	fruit
hill, valley, climb, summit, top, molehill, plain, peak, glacier, goat, bike, climber, range, steep, ski	mountain

In Brief, *continued from page 24*

Physicians' Don'ts Reference

That famed manual for doctors, the *Physicians' Desk Reference,* gives faulty, and possibly fatal, advice on treating overdoses, say physicians and pharmacists who answer emergency calls at the San Francisco Poison Control Center. The group surveyed 80 doctors who had called in for help and found that in the past year, half had turned to the seven-pound, 3,000-page tome, listing information from manufacturers. It further reviewed entries in the 1994 edition on six drugs often used in deadly overdoses. In each case, the *PDR* recommended treatments that were dangerous, ineffective or simply outdated. The *PDR*'s publisher, Medical Economics, states that several flaws have been fixed in the 1997 edition.

FOLLOW-UP

Bhopal: A Decade Later

A recent paper in the *National Medical Journal of India* looks at long-term consequences of the 1984 gas leak in Bhopal, India. By administering questionnaires and clinical tests in 1994, the authors found that Bhopal residents who had the highest gas exposure 10 years earlier suffered the largest number of general health problems, fevers, birth complications and respiratory symptoms. Neurological, psychiatric and ophthalmic diseases were also most prevalent among the most heavily exposed. In an accompanying paper, the International Medical Commission, Bhopal, argues for the creation of a worldwide bill of rights for health and safety to prevent such tragedies in the future. They specifically condemn Union Carbide for being less than straightforward about the quantity and composition of leaking gases at the time, failing to have provided any emergency preparation and, among other things, failing to deliver adequate compensation to the afflicted population. (See June 1995, page 16.) *—Kristin Leutwyler*

WIDE WORLD PHOTOS

percent of the words read to them. But here is the interesting part: the average participant also claimed to remember hearing 57 percent of the unspoken target words associated with his or her lists.

Varying the test to try to pin down the source of the effect, McDermott and Roediger put aside the first list of words their human guinea pigs remembered and made them start over. Given a second chance, the typical subject proceeded to include even more false memories than before.

Other researchers had male and female assistants take turns reading each successive word in the lists. Then the psychologists handed each test taker a page of multiple-choice questions. The page listed, in random order, half the words just read aloud plus the unspoken target words and a bunch of completely unrelated terms. The questions were the same for each word in the list: Did you hear this spoken? Who uttered it, a man or a woman—or don't you remember? The result was alarming: not only did these intelligent people often say they recalled hearing a target word that was never voiced, but many also recollected which experimenter supposedly pronounced it.

It is not too hard to see why. Each list collects words that all have to do with a target word. The longer a list, McDermott and Roediger discovered, the more likely people are to falsely remember hearing its target. The researchers hypothesize that as we hear the words "rest," "slumber" and "doze," the web of neurons in our brain naturally fetches the word "sleep" and adds it to our memories of those words actually heard. This simple theory does not explain, however, why some lists—words associated with "butterfly," for example—do not seem to produce false memories. Other factors must be at work.

Although humdrum words in a five-minute test lack the emotional weight and temporal distance of the traumatic, decade-old recollections at issue in false-memory syndrome, McDermott says her findings should apply "to all sorts of episodes ranging from minutes to the whole of one's life." Psychologists consider all memories that last for more than about 30 seconds to be "long-term" and thus susceptible to similar influences, McDermott maintains. She notes that her subjects were motivated to be accurate and knew that errors would be detected.

So, are you still confident about remembering that childhood argument? Certain it isn't just a story your grandmother once told you? If so, *Scientific American* wishes to remind you that you were planning to send in the check for your subscription renewal today.

—Tim Beardsley in Washington, D.C.

FIELD NOTES

SCENT PROSPECTORS

Looking for new "notes"

When I call Ken Purzycki and ask whether I can watch him field-test his portable scent-collection device, he demurs. There may be no fragrant flowers in the woods of northern New Jersey in early March, he says. Inexpert in the olfactory sciences, I blunder by asking whether I can pick up something that he can stick into his odor gatherer, maybe a Big Mac.

Purzycki says he doesn't do hamburgers, just the kind of scents that go into

ORCHID UNDER GLASS emanates odor molecules for the delectation of Scent Trek, a device invented by Ken Purzycki of Givaudan-Roure in Teaneck, N.J.

fine perfumes or dishwashing detergents. He does confide that once he captured and then faithfully reproduced the bouquet of the New York City subways when challenged to do so by a reporter, who went away suitably impressed.

Despite the pesterings of frivolous journalists, the director of fragrance science for Givaudan-Roure, one of the world's largest flavor and fragrance companies, proves himself a gentleman. He agrees to accommodate my request to inspect his scent collector. A week later I arrive at Givaudan's "creative center" in Teaneck.

First, I receive an introduction into the state of the art in olfactory research from Purzycki and his boss, Thomas McGee, the senior vice president for corporate development and innovation. The conversation ranges from the prospects for electronic noses (moderate) to virtual reality. Yes, that technology is finally gaining its last sensory input, a kind of postmodern version of the 1950s Smell-O-Vision. (Purzycki may have some use for that subway scent after all.)

After McGee gives me a whiff of a chemical that really does replicate the smell of a tropical beach, we move to the laboratory to observe Scent Trek, the reason for my visit. There, beside a gas chromatograph and a mass spectrometer, sits a potted orchid (genus *Cattelya*) with a glass bubble around its sumptuous pink petals. An outlet at the side of one of the two semicircular glass hemispheres allows the molecules emitted by the store-purchased flower to be sucked down a plastic tube and trapped onto one of 12 polymer filters that sits in a metal carrying case.

PETER MURPHY

A filter can be removed from the case and analyzed by chromatography and spectrometry to ascertain the identity and quantity of each odor molecule. Then the scent can be reconstituted, mixed with other fragrances and incorporated into a perfume or a shampoo.

Purzycki developed Scent Trek because of too many long nights spent in botanical gardens waiting for a plant to reach its "peak olfactive moment." Scent emission occurs only at the time of day when the plant is most likely to be pollinated. In the past, Purzycki would sit bleary-eyed beside a flower with "headspace" technology—a handheld filter and a gas flowmeter. Then he would return to the laboratory to analyze the sample.

Scent Trek is intended to automate

headspace (a brewing-industry term that refers to the foam produced by beer). The steel case has a specialized computer that activates a pump for intervals of one to two hours, drawing in the emanations from the bubble-enclosed flower. The filters in the kit allow for separate samples to be taken at different times throughout the day. For example, the peak scent for the orchid purchased from the northern New Jersey florist was between 5 A.M. and 7 A.M.

When he invented his scent collector, Purzycki was thinking about the Costa Rican rain forest rather than the wilds of New Jersey. Like many drug companies, scent manufacturers are seeking new chemicals from nature. Givaudan-Roure, owned by pharmaceutical maker Roche, already has a database of more than 15,000 natural chemicals amassed over the past 20 years. The biodiversity of the rain forest offers an opportunity for new "notes": the complex of chemicals from an individual flower or a material or place. A note may suggest a mood, an environment or even someone's interests. The fragrance named for Michael Jordan mixed notes from a beach, a golf course and a baseball glove.

Scent Trek was designed so that the entire apparatus could fit into a knapsack and be easily assembled in the field by a nontechnician. Givaudan-Roure has worked with Costa Rica's National Institute for Biodiversity, a nonprofit group that has supplied samples to Merck for new drug leads. Costa Rica has already yielded a few high notes. Take *Leueha candida*, a white flower that Givaudan-Roure describes as "reminiscent of a gardenia but without the harsh green note and with tones of tuberose but without the animalic note." Is there anything more to say? —*Gary Stix*

ANTI GRAVITY

Coffee Talk

Clearly, things have now officially gone too far. Incontrovertible evidence that coffee mania is out of control could be found in February at the annual meeting of the American Association for the Advancement of Science, where that august body devoted an entire session to the liquid the Food and Drug Administration should really consider designating as a "caffeine delivery system." Such a session was in keeping with the setting, for this year's meeting was held in a town where French Roast is easier to score than french fries—the Medellín of caffeine, Seattle.

Actually, the time was vine-ripe for a scientific look at coffee, what with it trailing only oil as the world's most widely traded commodity and what with caffeine being the world's most widely used psychoactive substance. Its insidious effects can be seen at virtually any of the legion of Seattle coffee bars, where burly, bearded, plaid-shirted timbermen wait patiently in long lines only to ask contritely for concoctions such as a "tall, 2 percent mocha latte."

Kate LaPoint, chief editor of *Coffee Talk,* a Seattle-based trade publication serving the coffee industry, told the AAAS session's audience of her own experience with what we can only hope is the limit of the mania. "I was driving down the highway," she said, "and I saw an ambulance driving really slowly. It was an `Espresso Ambulance.' They carry emergency espresso." Then again, perhaps even more fanatical is the coffee brewer she spoke about who checks the barometric pressure before brewing, so he can fine-tune his alchemy.

MICHAEL CRAWFORD

Jeffrey Parrish of the U.S. Agency for International Development noted coffee's influence on the switch from an industrial to an information-based culture. "I would contend," he remarked, "that the higher education and computer revolution that have become the very fabric of our society would not exist if a cup of java were not beside the keyboard." An ornithologist by training, Parrish went on to give a talk as eye-opening as the four varieties of coffee the session attendees were free to sample. Because coffee consumes 44 percent of the permanent arable cropland in northern Latin America, real environmental concerns surround its production. In particular, growers are moving toward environmentally hostile "sun coffee," grown in fields open to sunlight, and away from "shade coffee," where the fields still include a canopy of trees. Sun-coffee fields give higher yields but harbor as little as 3 percent of the number of bird species that shade-coffee areas do. The change thus eats away at wintering grounds for many songbirds familiar in the U.S. (Note to baseball fans at Camden Yards: as sun-coffee plots have become more common, oriole populations have dropped, so drink enough joe and the last oriole you see could be Cal Ripken, Jr.)

John Potter of Seattle's Fred Hutchinson Cancer Research Center talked about coffee's health effects (which for the average drinker, having one or two cups a day, are few) and gave a brief history. "The world's first coffee shop opened in Constantinople in 1475," he stated, "and shortly after that a law was passed making it legal for a woman to divorce her husband for an insufficient daily quota of coffee." (The headline in the Constantinople paper had to have been "Coffee Grounds for Divorce.")

The event that must get credit for giving rise to the current coffee frenzy, however, is Pope Clement VIII's decision 400 years ago, when he was urged to ban the substance because it came from the Islamic world. "He tasted it," Potter explained, "decided it was delicious and actually baptized it." One can only wonder what Clement, known for his piety, blurted out when he realized that he had watered down one terrific cup of cappuccino. —*Steve Mirsky*

PROFILE: ALAN COCCONI

Electric Cars and Pterosaurs Are My Business

On a warm and clear February afternoon, I am cruising southern California's Foothill Freeway in a one-of-a-kind electric roadster. I've got the San Gabriel Mountains on my left and Alan Cocconi on my right, in the passenger seat. Cocconi, who created this charged-up chariot, is egging me on. We are already pushing 90 miles an hour.

An electrical engineer, Cocconi's specialty is power electronics. Instead of fiddling with the usual milliwatts and microtransistors, he designs circuits in which tens of kilowatts course through transistors the size of jacket buttons. And in the U.S., at least, no one does it better than Cocconi, colleagues insist. "I am just good enough as an engineer to know how good he is," says Wally E. Rippel, a senior engineer at the prestigious high-tech consulting firm AeroVironment and a former staff physicist at the Jet Propulsion Laboratory in Pasadena, Calif.

Like most successful engineers, Cocconi is less well known than the creations that have flown (sometimes literally) because of his circuitry. One of them was a giant, flapping pterosaur, the star of the IMAX motion picture *On the Wing*. Another was SunRaycer, General Motors's winning entry in the landmark 1987 race across Australia of solar-powered electric vehicles.

These projects were just warm-up exercises for Cocconi's work on the Impact, the sleek, prototype electric vehicle that GM unveiled in 1990 to a blitz of media attention. Cocconi's circuitry converted direct current from the vehicle's batteries to the alternating current that ran its motor; it also converted AC to DC to charge the batteries. Given the late-1980s technology he had to work with, this circuitry, called an inverter, was a stupendous piece of engineering—and a major reason why the Impact was such a breakthrough.

With a few—but significant—electrical modifications, the Impact became the EV1, which GM released into southern California and Arizona last December. Cocconi, who had disapproved of most of the modifications, had long since left the project, for which he had been a handsomely paid subcontractor. His abrupt departure, in 1991, was characteristic. Colleagues describe him as a loner who has never been able to work with people, organizations or even ideas he does not hold in high regard.

Cocconi now runs his own company, AC Propulsion. Working out of a small, cluttered warehouse in a nondescript industrial park in San Dimas, Calif., Cocconi and his seven employees derive much of their income from converting gasoline-powered cars to battery power.

According to Cocconi, one of his converted 1993 Honda Civics, without any special streamlining, outperforms GM's arduously designed, highly aerodynamic EV1 in range and in the length of time needed to charge the batteries. AC Propulsion charges $75,000 to $120,000 to do a conversion; it has done 11 of them so far, while also selling 45 electric drivetrains to do-it-yourself converters. The company is profitable, Cocconi says, "if we don't pay ourselves too much."

Lately, in his spare time and with $200,000 of his own money, Cocconi built a flashy electric sports car, dubbed Tzero, that he hopes to market soon. "If you pay $75,000 for a car, you just don't want to come back and show the neighbors a Honda," he notes. The Tzero is, in fact, the red road monster in which I find myself tearing up the California pavement on this fine afternoon.

Cocconi's aptitude came from his parents, both Italian-born physicists. His father, Giuseppe, studied under En-

DAN WINTERS

***POWER ELECTRONICS WHIZ ALAN COCCONI** gets a charge out of taking on the titans.*

rico Fermi in Rome and in 1959 wrote a famous paper with Philip Morrison, then at Cornell University, proposing the use of the hydrogen emission spectra in the search for extraterrestrial intelligence.

"I'm lucky I did something different, so I didn't have to compete with him," Cocconi says of his father. "I don't understand his physics, and he doesn't understand my electronics."

In 1962, when Cocconi was four years old, his parents left Cornell for CERN, the European center for particle physics. Raised in nearby Geneva, Switzerland, Cocconi immersed himself in building radio-controlled model airplanes—not from kits, like most hobbyists, but from scratch. Accepted to the California Institute of Technology, he arrived in Los Angeles in the late summer of 1976 with a coffin-size Styrofoam box containing his precious planes.

At the airport he made a rude reacquaintance with the country of his birth. The bus driver who was to take him to the campus refused to take the box; a stalemate ensued until Cocconi realized that the man simply wanted a bribe. "He said he wanted a six-pack. I didn't know what that was, so I told him I'd just pay him the money for it. It turned out all he wanted was two dollars."

At Caltech, Cocconi concentrated on electronics. "My motivation was simple and not that noble," he says. "I wanted to build better model airplanes." He later realized that in electronics, as opposed to aeronautics, it would be somewhat easier for him to steer clear of military work.

This theme is a recurring one with Cocconi; asked if he is antimilitary, he thinks for a moment and replies that he is "reasonably antimilitary. I just don't want to actively contribute to the effort." He shrugs. "I guess that growing up in Switzerland gave me a slightly different outlook."

After college he worked for a couple of years designing power electronics circuitry for a small company called Teslaco; it was the only time in his life that he has been an employee. He saved up $7,000 and promptly quit, because he had decided that what he really wanted to do was design and build remotely piloted airplanes.

"My parents were upset," he recalls. "Two years out of school, I quit my job with no prospects for a new one. My biggest fear was that I'd be a model-airplane bum for the rest of my life."

Working alone in his tiny Pasadena apartment, Cocconi designed flight surfaces, airframes, control electronics and even antennas and crafted little aeronautical gems out of fiberglass, foam and carbon fiber. He installed video cameras and flew the planes high above dry lake beds in the Mojave Desert, taking close-up shots of snowcapped peaks. After a year, a friend in the drone business introduced him to an engineer from the National Aeronautics and Space Administration. So impressed was the engineer by one of Cocconi's planes that he awarded him a contract to build a drone for aerodynamics research.

Cocconi subsequently contacted an acquaintance, Alec Brooks of AeroVironment, and the company soon provided him with some contract work. In 1984 the Smithsonian Institution and Johnson Wax Company agreed to fund *On the Wing,* the IMAX motion picture, whose script demanded a flying mechanical pterosaur. AeroVironment, which was hired to build the beast, in turn paid Cocconi to create the circuits that would flap the wings and guide and stabilize the contraption in the air.

DAN WINTERS

ELECTRIC ROADSTER: ***the "Tzero" will provide guilt-free thrills for the wealthy and environmentally conscious.***

The task was tricky. Most flying machines are not unstable in any axes of motion, whereas the pterosaur robot was unstable in both pitch and yaw. The pterosaur managed to soar as required for the movie's rousing climax, but later it crashed at a military air show. The recollection still amuses Cocconi to no end. "There were articles on how taxpayer dollars were being wasted on pterodactyls that crashed," he says between convulsive giggles.

Through AeroVironment, Cocconi got the contracts for his work on the SunRaycer and Impact vehicles. These jobs gave him the expertise—and funds—to launch AC Propulsion. Besides converting gas-powered cars to electric and building the red roadster, Cocconi has built a little trailer that houses a small gasoline engine and converts any of his electric cars into a hybrid vehicle capable of cruising easily at highway speeds and with essentially unlimited range.

Cocconi has driven the car-trailer hybrid across the U.S. twice, once in September 1995 to a meeting of the Partnership for a New Generation of Vehicles (PNGV) group in Washington, D.C. Under the advocacy of Vice President Al Gore, the PNGV consortium was formed to develop advanced vehicle technologies, including hybrids.

At the Washington meeting, Cocconi was "pretty disheartened by what I was hearing." While conceding that his car-trailer combination may not be precisely the configuration the PNGV envisioned in its long-term plans, he claims that his vehicle already meets all other PNGV specifications for a hybrid. Nevertheless, Cocconi says, Gore's representatives at the PNGV meeting ignored him: "If they were serious about getting something on the road, you'd think they would have at least wanted to ask me a few questions."

"The approach is to hand out money to the automakers and to justify handing it out" rather than to try to get a practical car on the road, he declares. This mentality, in his opinion, was also responsible for what he sees as the major flaw in General Motors's EV1: the car has an inductive charger, rather than the conductive one Cocconi favored and used in the Impact. The use of an inductive system requires those who drive an EV1 to have a $2,000 charger that must be installed by a local utility.

GM claims the system is safer, but Cocconi disagrees, asserting that the company's choice of the inductive charger "worked as a very effective sponge to soak up all the federal and state dollars that could have gone into creating a much cheaper and pervasive infrastructure" for electric vehicles.

Later on, we take one of the converted Hondas out for an evening spin. In the belief that more information is better, Cocconi has arrayed on the dash a generous assortment of gauges and dials, in whose amber glow the inventor appears beatific. Under dark, unkempt hair, his toothy smile reveals the creator's contented bliss. —*Glenn Zorpette*

MICROELECTRONICS

UNDER THE WIRE

Chipmakers face a looming performance barrier

Gordon Moore, meet Georg Simon Ohm and Michael Faraday. Moore's law—the oft-cited dictum of Intel's chairman emeritus that projects huge leaps in chip power in ridiculously short periods—is coming under siege from a surfeit of ohms and farads, the units of resistance and capacitance named after the two renowned 19th-century scientists.

Moore's law postulates that by continually making transistors smaller and squeezing them closer together, the number of tiny switches on a chip doubles every 18 months. But the principle of getting more for less does not apply to the wires connecting the millions of transistors that populate the fastest microprocessors. "For the first decades of this industry, transistors, not their interconnections, were what mattered," says Mark T. Bohr, Intel's director of process architecture and integration. "But there is no good way to shrink down the interconnections the way you can the transistors."

As the width of the wires decreases and the distance electrons must travel between the multitude of transistors lengthens, the resistance of the wires to the flow of current increases. And capacitance—an unwanted transfer of electrical energy among closely spaced wires—can cause current to slow or a digital bit to shift erroneously from a 0 to a 1.

The hundreds of meters of wiring that crisscross some of today's most advanced microprocessors already account for about half of the delay in signals traversing the chip. Moreover, connecting the seven million or so transistors on these chips—a feat accomplished by the skyscraperlike stacking of five or six layers of wiring on top of the chip's surface [*see illustration below*]—now consumes about half of the costs of manufacturing. Five years ago these expenses represented about a third of the fabrication bill.

For years, chip manufacturers paid wiring little heed. "The interconnections were just an afterthought," says Kenneth Monnig, a program manager at the industry consortium SEMATECH. The wiring issue, however, has moved to the forefront of the industry's attention, requiring major changes in materials and manufacturing processes.

One seemingly simple solution is to switch from aluminum wiring, which is now the standard, to copper. Copper exhibits lower resistance, and it is also less subject to a phenomenon known as the electron wind, the tendency of a dense flow of current coursing through a narrow wire to erode the metal. "You get a gap in the metal because the electrons blow the metal molecules down the wires," remarks G. Dan Hutcheson, president of VLSI Research.

The switch to copper, however, is not being warmly anticipated in the billion-dollar-plus fabrication plants. Although chip companies already alloy aluminum with small amounts of copper, a wholesale adoption of this technology may require changes in a basic manufacturing step.

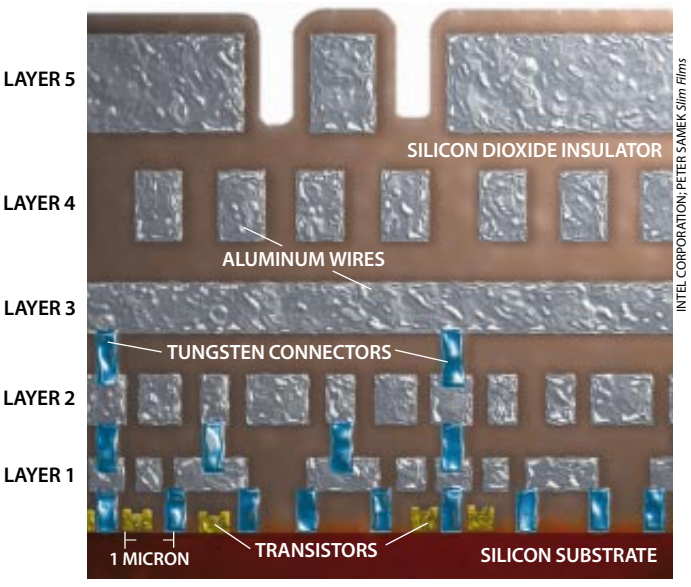
Aluminum can be laid down on a chip as a thin metal film and etched with a plasma of ions into metal wiring lines. Copper, in contrast, may have to be deposited into narrow grooves that have been carved out of the surrounding insulating material. The metal must uniformly fill trenches that may be a quarter of a micron or less in width. Copper also has a tendency to diffuse into the surrounding silicon, which can destroy a transistor's switching ability.

Larger gains in performance could be forthcoming from the use of better insulators, or dielectrics, that reduce the capacitance buildup that causes signal delays. As two wires are placed ever closer, they begin to function more like a capacitor—a device that stores electrical charge—and less like a highway through which electrons speed from one point to another.

Chipmakers now seek a substitute for silicon dioxide, the reigning dielectric. Silicon dioxide's only sin is too high a dielectric constant—a measure of its ability to keep a signal in one wire from disrupting neighboring signals. The lower the dielectric constant, the easier it is to avoid signal interference.

Substitutes for silicon dioxide, however, all come with a set of unwanted baggage. "The Styrofoam cup on your desk has a low dielectric constant, but getting it to withstand 400 degrees Centigrade during the manufacturing process is a whole other challenge," says Intel's Bohr. Polymers such as Teflon are being tested, but they tend to soften at high processing temperatures. "When you have five levels of that stuff, you don't want that to happen," says Fabio Pintchovski, director of materials research for Motorola Semiconductor.

Adding fluorine or carbon atoms to silicon dioxide can make it a better dielectric. A Chatsworth, Calif., company,



INTEL CORPORATION; PETER SAMEK *Slim Films*

SKYSCRAPER STACKING OF WIRES *above the surface of a microchip—in this artist's rendition—is needed to connect the roughly seven million transistors on some advanced microprocessors. Tungsten connectors allow wires to be linked from layer to layer. The wires become thicker in the upper layers to reduce resistance so that widely separated groups of transistors can be linked by high-speed connections.*

Trikon Technologies, recently claimed that carbon halved silicon dioxide's dielectric constant. But the industry has taken a wait-and-see attitude, because silicon dioxide additives may cause undesirable chemical reactions during manufacturing.

Recognizing that air has the lowest dielectric constant of any substance, one company has taken a Swiss-cheese approach to fashioning dielectrics. Nanoglass, a joint venture between Allied-Signal and a New Mexico start-up called NanoPore, has crafted a silicon dioxide material with air-filled pores that can be as small as 10 nanometers in diameter.

The material, which is similar to a class of substances known as aerogels, can achieve a dielectric constant only slightly above that of air itself. But manufacturers must still determine whether the porous material will withstand the stresses of the fabrication line.

Time is running out to solve the host of remaining technical problems. "The solutions are not obvious," laments Robert Havemann, a Texas Instruments fellow. Even if chipmakers can introduce copper wiring and a dielectric as good as air, they will buy themselves only another decade until some radically new technology is needed: circuit connections using light waves or radio signals or some wholly new chip designs that forgo the smaller-is-better approach. Unless new approaches emerge, the history books may look back on Moore's law as an artifact of the electronic industry's adolescence. —*Gary Stix*

CHEMISTRY

STRANGE ATTRACTORS

Chemists make magnets without metal

Wave a magnet over a cup filled with iron filings, and it is hardly surprising to see them stand on end. It would be quite something else if passing a magnet over a cup of coffee could suddenly pull all the caffeine to the surface. Or if an old blueprint could stick to the refrigerator all by itself. Admittedly, these wonders are pretty farfetched, but chemists have recently rearranged the same organic constituents that make up caffeine and blueprint dye to produce two new kinds of magnets that are lighter, more flexible and easier to make than the common metal variety.

Nonmetallic magnets work because magnetism is not a property of metals per se but of the electrons in them. Electrons have a property called spin that makes them behave like tiny magnets, each with north and south poles. When the spins on many adjacent electrons all point in the same direction, the overall effect produces the familiar poles of any magnet. Certain metals are easy to magnetize because they have an abundance of electrons just waiting to line up in magnetic order. But a number of nonmetallic substances have electrons to play with as well.

Joel S. Miller, now at the University of Utah, and Arthur J. Epstein of Ohio State University discovered the first such organic magnet in 1985. Although the compound did contain iron atoms, it was various organic additives that really made it work. Other researchers have since assembled completely organic magnets. In 1991 scientists in Japan created a magnetic compound, called 4-nitrophenyl nitronyl nitroxide, that contains just carbon, hydrogen, nitrogen and oxygen—the same ingredients in caffeine and a host of biological chemicals.

These early materials were impractical because they became magnetic only when cooled nearly to absolute zero. That is no longer a problem. Miller and Epstein have now developed an organic-based magnet that retains its properties up to 75 degrees Celsius (167 degrees Fahrenheit). The compound consists of the nonmagnetic metal vanadium surrounded by the organic molecule tetracyanoethylene, or TCNE.

A French team led by Michel Verdaguer of the University of Pierre and Marie Curie in Paris has also produced room-temperature magnets related to the pigment Prussian blue, once used to color blueprints and fabric. These deep-blue compounds, made with vanadium and chromium atoms surrounded by organic groups, will stick to other magnets up to approximately 42 degrees C (108 degrees F).

Gregory S. Girolami of the University of Illinois, who has worked with the Prussian blue magnets, explains that the new nonmetallic materials magnetize because their atoms are arranged in rigid lattices that tighten interactions between electrons, encouraging them to align their spins. Chemists are fiddling with these lattices to produce organic magnets that work at even higher temperatures and can compete with the strength of their iron counterparts.

Now that organic magnets work at room temperature, engineers are starting to speculate about ways to exploit their advantages over metals. For one, they should bend and spread more easily. They might also be cheaper than metal magnets, which are typically produced at vulcanian temperatures. Flexible magnetic coatings or high-density magnetic data storage systems are two obvious application possibilities.

Soon after his paper on room-temperature organic magnets appeared, Miller received calls from a cosmetic company ("I'm not sure what they wanted," he says) and from a doctor hoping to improve the magnetic valves in artificial hearts. But because the vanadium-TCNE compound reacts explosively with oxygen, and the Prussian blue magnets weaken with time, widespread applications will have to wait.

Nevertheless, the promise of lightweight, plasticlike magnets has excited many scientists. Some are now studying the materials' unusual ability to change magnetic properties when exposed to light—an attractive feature for high-density optical data storage systems. —*Sasha Nemecek*

JOEL S. MILLER *University of Utah*

NONMETALLIC MAGNET **(inside tube, at right)** ***is made from the two liquids shown at the left.***

RENEWABLE ENERGY

POWER TO THE PEOPLE

"Net metering" makes producing energy at home more economical

After an environmentally conscious home owner installs solar panels on the roof or a wind-driven generator in the backyard, clean electricity flows for free—but only while the sun shines and the wind blows. One way to cope with this intermittence is to store energy in batteries. But a growing number of utility companies now allow home owners a less expensive option. They can deposit the excess electricity they produce into the power grid and withdraw it at later times using just their standard household electrical meter, which can run equally well backward or forward. Permitting such "net metering" gives a single home owner a privilege normally exercised only among giant utility companies: trading electricity generated at one time for the power required at another.

Net metering should help spur small-scale production of renewable energy. Otherwise, a home owner receives only the so-called avoided cost for any electricity exported to the power grid, and this rate is just a fraction of what the utilities typically charge residential customers. But with net metering, individuals can get, in essence, the full retail price for the electricity they generate, so long as they buy it back during the same billing period. (Any surplus production at the end of the month still earns only the wholesale price.)

Although net metering alone does not normally make home generation of renewable energy economical, it does bring such efforts somewhat closer to the break-even point. In sunny Hawaii, for instance—where the state government offers a solar-energy tax credit and the cost of electricity is especially high—net metering could make home solar-power systems cost-effective.

Net metering may also encourage people in more marginal situations to invest in solar or wind generators. According to Michael L. S. Bergey, president of Bergey Windpower, when net metering is not offered to them, some home owners will balk at the idea of selling their excess electricity at a discount to the local utility, which then resells the power for several times the price. "The biggest thing [net metering] does is change the mind-set," concurs Christopher Freitas, director of engineering for Trace Engineering, a company that makes power conditioning equipment used in home installations. "The idea of being able to spin a utility meter backward really appeals to people."

RICHARD CHOY *Peter Arnold, Inc.*

***SOLAR PANELS** can run the electricity meter backward.*

Yet some utilities and government officials have resisted net metering, which is now available in Japan and Germany but only in 16 U.S. states. When advocates of renewable energy proposed a net-metering law in California in 1995, Pacific Gas and Electric, a major utility, fought against it—but lost. New York State governor George Pataki vetoed a net-metering law passed last year, citing concerns that energy from homes might continue to flow through lines during general outages, endangering power-company workers.

"That argument was absolutely bogus," says Thomas J. Starrs, a lawyer at Kelso, Starrs and Associates, who helped to write the legislation for net metering in California. Power from home generators, he notes, runs through a device called an inverter that converts direct current to alternating current; should power in an area fail, the inverter automatically cuts off the flow. But green advocates recognize that they will still have to expend some energy of their own to persuade all concerned parties—from state governors and utility officials to local building inspectors and insurers—that individual home owners can safely generate power from sources that do not create radioactive waste or greenhouse gases. As Starrs quips, "We're still working out the kinks." —*David Schneider*

MEDICINE

TAMING TREMOR

A pacemaker for the brain nears approval

Bearing down on the drill, a surgeon bores a two-centimeter-wide hole into the top of Russell D. Sherman's skull. Sherman hardly notices. Conscious and smiling, the 69-year-old plumber from Carrington, N.D., raises his voice over the noise to explain why he has made the 800-mile trip—his fifth in nine months here to the University of Kansas Medical Center in Kansas City, Kan.—to have a thin electrical wire threaded into the center of his brain. "If you couldn't write or drink from a cup, you'd understand," he says.

Since he was a young man, Sherman has suffered the effects of essential tremor, a hereditary degenerative disease that causes limbs to tremble, heads to nod, voices to quaver. It affects both of Sherman's hands, and there is no cure. Yet he greeted me before the operation with a steady handshake. An electrode in the left side of his brain, attached to a pacemakerlike power cell tucked inside his chest, was overriding the tremor in his right hand. Now he is having his right brain wired to control his left hand.

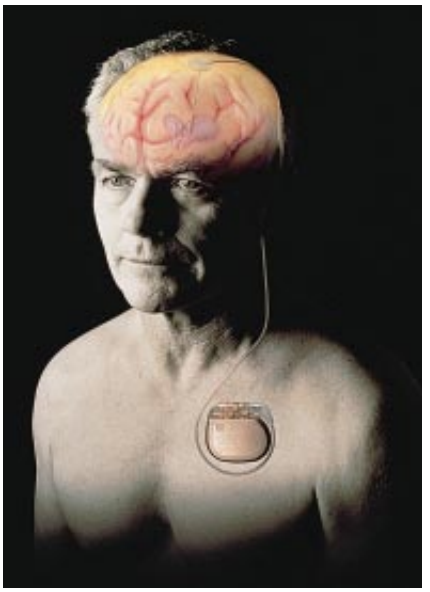
The next day Sherman touches a magnet to his chest to start the power cell pulsing. "Oh, honey, look!" his wife exclaims as the uncontrollable wave of his left hand dies to a mere jitter. "He's had to wear snap shirts and shoes with Velcro straps," she says. "Now he can go back to buttons and laces."

Sherman owes the steadiness of his hand to the slip of a surgeon's some 45 years ago. In an attempt to remove a different part of a patient's brain, the doctor accidentally cut off blood to the thalamus. When she awoke, her tremor was gone. Neurosurgeons eventually identified a pea-size region of cells in the thalamus that, when killed, often stops the shakes. But the therapy, called a thalamotomy, sometimes causes sensory and speech problems. When three

drugs were approved in the late 1960s to treat essential tremor and the quakes of Parkinson's disease, they almost completely replaced surgical treatment. The drugs help most patients for a while, but for about 5 percent of those diagnosed with essential tremor and about 10 percent of Parkinson's patients, drugs no longer offer sufficient relief.

In 1987 Alim L. Benabid of the Joseph Fourier University in Grenoble, France, tried a different operation. Instead of turning up the power on a probe to burn the thalamic sweet spot, he simply left it in the brain, emitting low-voltage pulses 130 times a second. The implant seemed to calm tremors as well as a thalamotomy did, but at lower risk.

Benabid took his results to Minneapolis-based Medtronic. Zapping a recalcitrant body into submission is one business Medtronic, as the leading producer of defibrillators and heart pacemakers, knows well. In 1993 the company set up clinical trials around the world and in 1995 began selling the systems in western Europe, Australia and Canada. About 80 percent of the more than 2,000 patients who have received the implant report complete or partial calming of their tremor, claims Donald H. Harkness, Medtronic's clinical manager for the device. Benabid's studies confirm that success rate.



COURTESY OF MEDTRONIC

***ELECTRONIC IMPLANT** in the brain quiets quaking limbs.*

In March a U.S. Food and Drug Administration review panel unanimously recommended approving the device for the American market as well. With perhaps 200,000 potential patients, market analysts predict Medtronic could sell more than $100-million worth of the devices in the year 2000.

To do so, the firm will have to persuade insurers to cover the cost. Harkness reports that the Netherlands has already requested a study comparing the long-term cost and benefits of the implant with those of destructive surgery. U.S. insurers may follow suit. "Scientifically, the question is very interesting," Harkness admits. But for Medtronic, he adds, it is risky business: "I don't try to answer questions that I don't really want to know the answers to. We're able to position [the device] now without data. If we did the studies, and thalamotomy, say, proved superior—well, what would we do?"

Ultimately, the question may be moot. Any side effects caused by a thalamotomy are irreversible. Faced with that unpleasant prospect, patients may scrape together the additional $11,000 or so themselves for a device that can be controlled and, if necessary, removed. Sherman says he doesn't know whether his insurance will cover his second implant. "I'll pay for it myself if I have to," he declares. "Whatever it costs, it's worth it."

—W. Wayt Gibbs in Kansas City, Kan.

CYBER VIEW

World Wide Widgets

You would think that a message scrambled with RSA Data Security's RC5 encryption software would be safe from hackers' prying eyes. After all, in order to break the code one must try 281 trillion possible keys to find the one that fits. The National Security Agency may possess the monstrous processing power to do this before the coded message grows moot, but few others do—the fastest desktop machine now available would take about two years to do the job.

Yet on February 10 a German deciphered an RC5-encrypted message just 13 days after it was released. Had he not found the key, a Swede soon would have, or a South African or another of the more than 5,000 volunteers, linked via the Internet, who had divvied up the problem and attacked its parts simultaneously. Each participant downloaded a widget, or small program, that checks a chain of keys when the computer is otherwise unoccupied. On finishing those calculations, each worker posted the results to a central computer and downloaded more untested keys to try. For two weeks, this simple scheme created an ad hoc supercomputer as powerful as any yet built. Just imagine if the process were automated, repeatable—and profitable.

You'll have to imagine, because such a thing does not yet exist. It's not for lack of trying. For years, computer scientists have dreamed of software that, given a really large task, could borrow any kind and number of idle machines to which it was connected and speed up the job by spreading it around. As long ago as 1992 the concept was considered feasible and cool enough to warrant a name: metacomputing. But attempts to link supercomputers into a metacomputer foundered. Interest waned.

Now metacomputing buffs are buzzing again, for two reasons. The Internet has dramatically expanded the pool of connected machines. And Java, a set of software standards produced by Sun Microsystems, now allows programmers to write code that can run on many previously incompatible platforms.

A spate of recent experiments, both real and thought, has some optimistic researchers bubbling with ideas about how to use metacomputers. Ian Foster of Argonne National Laboratory, among others, imagines the World Wide Web evolving into a "computational grid" one can plug into almost as easily as an electrical outlet. Want to render *Toy Story II* but can't afford 100 high-end workstations? Rent spare capacity on the computers of Pixar, the creator of the original graphics tour de force. The possibilities for ubiquitous scientific supercomputing are enough to make researchers giggle with glee.

Or chuckle with cynicism. Metacomputing still faces formidable obstacles. Some are technical. Putting the data that a program requires into a form all machines can understand, for example. Or splitting tasks into independent chunks so that separate machines can work in parallel. That is easy for only a small set of problems, such as rendering three-dimensional images or breaking codes.

DAVID SUTER

Economic realities impose other constraints. It is much cheaper and faster to process bits than to move them around. If a helper machine takes one hour to solve part of your problem but 30 minutes to download the widget and return its answer, the cost is probably higher than the benefit. And many parallel programs grind to a halt if the cooperating machines cannot share their results very quickly.

The potential showstoppers, however, are mostly social hang-ups. How could VISA or IBM trust strange machines to work on their data without peeking at them? If a scientist is paying others to run a simulation for her, how can she be certain they aren't taking the money and returning fake or flawed results?

Of the dozen or so metacomputing projects under way, several have nearly cleared some of these obstacles. At the University of California at Berkeley, computer scientists are working on a project named WebOS that incorporates URLs (the file names used by Web browsers) into a workstation's normal file system so that files can be written to, as well as read from, the Web. To reduce network congestion, the researchers propose scattering copies of shared files onto widely separated servers. They have built a system called Smart Clients that can automatically keep the copies in sync. When users need to open a file, Smart Clients sends them a little Java program that determines which server is closest and least busy.

The Javelin prototype at the University of California at Santa Barbara creates software brokers to balance demand with supply. Jobs are submitted as Java programs to the broker, which then forwards copies to idle helper machines. Because Web browsers force Java programs to run within a "playpen," they can (in principle) do no harm to the helper computers. But because the playpen denies programs access to the disk and other important functions, it can be difficult to get real work out of such programs.

Perhaps the most promising experiment to date is Charlotte, built at New York University. Charlotte distributes its Java programs to volunteer computers that then race one another to finish the work. The strategy is inefficient but effective: when Charlotte parceled out a model of magnet interaction to 10 machines, the problem was solved nine times as quickly. The system arrived at correct answers even when several of the machines were deliberately crashed.

So far not even the most ambitious metacomputing prototypes have tackled accounting: determining a fair price for idle processor cycles. It all depends on the risk, on the speed of the machine, on the cost of communication, on the importance of the problem—on a million variables, none of them well understood. If only for that reason, metacomputing will probably arrive with a whimper, not a bang. It will squeeze more power from supercomputer centers, campus networks and corporate intranets. But don't expect your home PC to start earning its keep anytime soon.

—W. Wayt Gibbs in San Francisco

Divided We Fall: Cooperation among Lions

Although they are the most social of all cats, lions cooperate only when it is in their own best interest

by Craig Packer and Anne E. Pusey

In the popular imagination, lions hunting for food present a marvel of group choreography: in the dying light of sunset, a band of stealthy cats springs forth from the shadows like trained assassins and surrounds its unsuspecting prey. The lions seem to be archetypal social animals, rising above petty dissension to work together toward a common goal—in this case, their next meal. But after spending many years observing these creatures in the wild, we have acquired a less exalted view.

Our investigations began in 1978, when we inherited the study of the lion population in Serengeti National Park in Tanzania, which George B. Schaller of Wildlife Conservation International of the New York Zoological Society

YOUNG FEMALE LIONS, shown here, band together in groups of six to 10, called prides. Such togetherness does not always make them more successful hunters, as scientists once presumed; loners frequently eat more than individuals in a pride do. Instead communal living makes lions better mothers: pridemates share the responsibilities of nursing and protecting the group's young. As a result, more cubs survive into adulthood.

PHOTOGRAPHS BY CRAIG PACKER

SISTERHOOD makes it possible for pridemates to protect their cubs against invading males (*top*). Angry groups can ward off lone males, which are on average nearly 50 percent larger than females (*middle*). And they will frequently attack and kill less powerful trespassing females (*bottom*).

began in 1966. We hoped to discover why lions teamed up to hunt, rear cubs and, among other things, scare off rivals with chorused roars. All this togetherness did not make much evolutionary sense. If the ultimate success of an animal's behavior is measured by its lifetime production of surviving offspring, then cooperation does not necessarily pay: if an animal is too generous, its companions benefit at its expense. Why, then, did not the evolutionary rules of genetic self-interest seem to apply to lions?

We confidently assumed that we would be able to resolve that issue in two to three years. But lions are supremely adept at doing nothing. To the list of inert noble gases, including krypton, argon and neon, we would add lion. Thus, it has taken a variety of research measures to uncover clues about the cats' behavior. Indeed, we have analyzed their milk, blood and DNA; we have entertained them with tape recorders and stuffed decoys; and we have tagged individuals with radio-tracking collars. Because wild lions can live up to 18 years, the answers to our questions are only now becoming clear. But, as we are finding out, the evolutionary basis of sociality among lions is far more complex than we ever could have guessed.

Claiming Territory

Male lions form lifelong alliances with anywhere from one to eight others—not out of any fraternal goodwill but rather to maximize their own

MALES are quick to challenge lions they do not know—real or not. When the authors played tape recordings of strange males roaring within a coalition's turf, representatives from that coalition immediately homed in on the sound. Moreover, they often took the offensive, pouncing on decoys placed nearby.

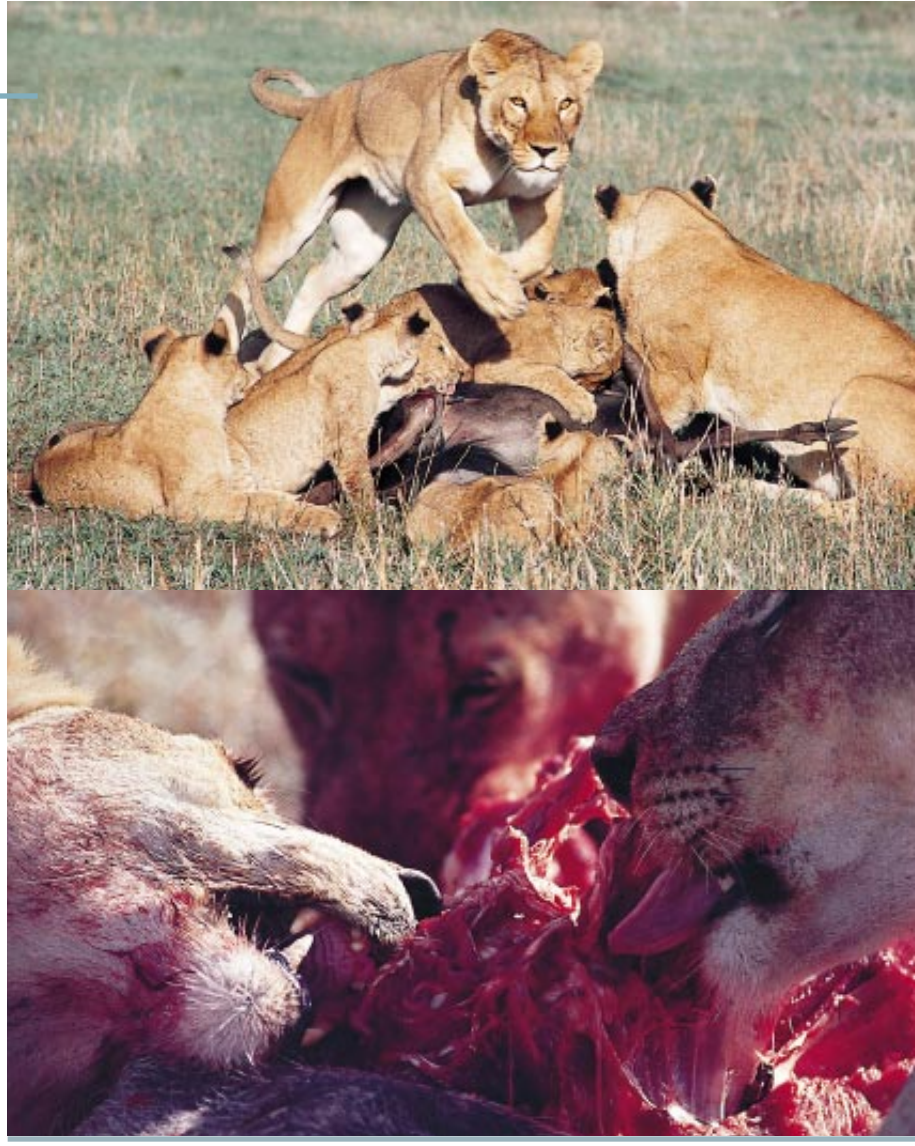
chances for reproducing. Most companions are brothers and cousins that have been reared in the same nursery group, or crèche. Others consist of nonrelatives that teamed up after a solitary nomadic phase. Once matured, these coalitions take charge of female lion groups, called prides, and father all offspring born in the pride during the next two to three years. After that, a rival coalition typically moves in and evicts them. Thus, a male lion's reproductive success depends directly on how well his coalition can withstand challenges from outside groups of other males.

Male lions display their greatest capacity for teamwork while ousting invaders—the situation that presents the greatest threat to their common self-interest. At night the males patrol their territory, claiming their turf with a series of loud roars. Whenever we broadcast tape recordings of a strange male roaring within a coalition's territory, the response was immediate. They searched out the speaker and would even attack a stuffed lion that we occasionally set beside it. By conducting dozens of these experiments, our graduate student Jon Grinnell found that unrelated companions were as cooperative as brothers and that partners would approach the speaker even when their companions could not monitor their actions. Indeed, the males' responses sometimes bordered on suicidal, approaching the speaker even when they were outnumbered by three recorded lions to one.

In general, large groups dominate smaller ones. In larger coalitions, the males are typically younger when they first gain entry into the pride, their subsequent tenure lasts longer and they have more females in their domain. Indeed, the reproductive advantages of cooperation are so great that most solitary males will join forces with other loners. These partnerships of nonrelatives, how-

SERENGETI NATIONAL PARK in Tanzania houses a population of lions that has been studied by scientists since 1966.

PREY CAPTURE is usually done by a single lion, when the group is hunting warthog and wildebeest (*photographs*). Because she will very likely succeed in capturing such easy prey, her sisters will probably eat even if they refrain from the chase. Thus, the pride will often stand back at a safe distance, awaiting a free meal. But when a single lion is less likely to make a kill—say, if she is stalking zebra or buffalo—her pridemates will join in to pursue the prey together (*charts*).



KILLS are shared by the entire pride. If kills are made close to home, mothers bring their cubs to the feast. But they deliver nourishment from more distant kills in the form of milk.

ever, never grow larger than three. Coalitions of four to nine males are always composed of close relatives. Why do not solitary males recruit more partners until their groups also reach an insuperable size? The reasons again come down to genetic self-preservation and, in particular, weighing the odds of gaining access to a pride against those of actually fathering offspring.

Although large coalitions produce the most offspring on a per capita basis, this averaging assumes fair division among companions—a form of cooperation that does not happen in the Serengeti. In fact, the first male to find a female in estrus will jealously guard her, mating repeatedly over the next four days and attacking any other male that might venture too close. Dennis A. Gilbert, in Stephen J. O'Brien's laboratory at the National Cancer Institute, performed DNA fingerprinting on hundreds of our lion samples and found that one male usually fathered an entire litter. Moreover, reproduction was shared equally only in coalitions of two males. In the larger coalitions, a few males fathered most of the offspring. Being left childless is not too bad from a genetic standpoint if your more successful partner is your brother or cousin. You can still reproduce by proxy, littering the world with nephews and nieces that carry your genes. But if you are a lone lion, joining forces with more than one or two nonrelatives does not pay off.

Hunting

Traditionally, female lions were thought to live in groups because they benefited from cooperative hunting. (The females hunt more often than the resident males.) But on closer examination, we have found that groups of hunting lions do not feed any better than solitary females. In fact, large groups end up at a disadvantage because the companions often refuse to cooperate in capturing prey.

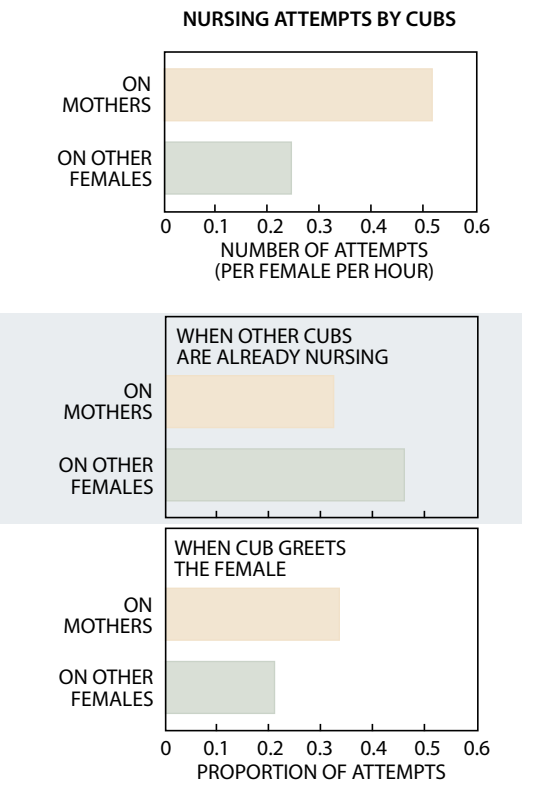
Once one female has started to hunt, her companions may or may not join her. If the prey is large enough to feed the entire pride, as is the usual case, the companions face a dilemma: although a joint hunt may be more likely to succeed, the additional hunters must exert themselves and risk injury. But if a lone hunter can succeed on her own, her pridemates might gain a free meal. Thus, the advantages of cooperative hunting depend on the extent to which a second hunter can improve her companion's chances for success, and this in turn depends on the companion's hunting ability. If a lone animal is certain to succeed, the benefits of helping could never exceed the costs. But if she is incompetent, the advantages of a latecomer's assistance may well exceed the costs.

Evidence from a wide variety of bird, insect and mammalian species suggests that, as expected, cooperation is most wholehearted when lone hunters do need help. The flip side of this trend is that species are least cooperative when hunters can most easily succeed on their own. Consistent with this observation, our graduate student David Scheel found that the Serengeti lions most often work together when tackling such difficult prey as buffalo or zebra. But in taking down easy prey—say, a wildebeest or warthog—a lioness often hunts alone; her companions watch from the sidelines.

Conditions are not the same throughout the world. In the Etosha Pan of Namibia, lions specialize in catching one of the fastest of all antelopes, the springbok, in flat, open terrain. A single lion could never capture a springbok, and so the Etosha lions are persistently cooperative. Philip Stander of the Ministry of Environment and Tourism in Namibia has drawn an analogy between their hunting tactics and a rugby team's strategy, in which wings and centers move in at once to circle the ball, or prey. This highly developed teamwork stands in sharp contrast to the disorganized hunting style of the Serengeti lions.

All female lions, whether living in the Serengeti or elsewhere, are highly cooperative when it comes to rearing young. The females give birth in secrecy and

NURSING is a job shared by all mothers in a pride, not out of generosity but, rather, fatigue. Cubs feed when their mothers return from hunting (*top*). If the mothers stay awake, they will not let cubs other than their own, such as the large adolescent shown, take milk from them (*bottom*). Although cubs try to nurse most often from their own mothers, they can be quite cunning in their attempts to nurse from other females (*charts*).



keep their litters hidden in a dry riverbed or rocky outcrop for at least a month, during which time the cubs are immobile and most vulnerable to predators. Once the cubs can move, though, the mothers bring them out into the open to join the rest of the pride. If any of the other females have cubs, they form a crèche and remain in near-constant association for the next year and a half before breeding again. The mothers lead their cubs to kills nearby but deliver nourishment from more distant meals in the form of milk. When they return from faraway sites, the mothers collapse, leaving their youngsters to nurse while they sleep. We have studied over a dozen crèches, and in virtually every case, each cub is allowed to nurse from each mother in the group. Communal nursing is a major component of the lion's cooperative mystique.

And yet, as with most other forms of cooperation among lions, this behavior is not as noble as it seems. The members of a crèche feed from the same kills and return to their cubs in a group. Some are sisters; others are mother and daughter; still others are only cousins. Some have only a single cub, whereas a few have litters of four. Most mothers have two or three cubs. We milked nearly a dozen females and were surprised to discover that the amount of milk from each teat depended on the female's food intake and not on the actual size of her brood.

Because some females in a pride have more mouths to feed, yet all produce roughly the same amount of milk, mothers of small litters can afford to be more generous. And in fact, mothers of single cubs do allow a greater proportion of their milk to go to offspring that are not their own. These females are most generous when their crèchemates are their closest relatives. Thus, milk distribution depends in large part on a pattern of surplus production and on kinship. These factors also influence female behavior across species: communal nursing is most common in those mammals—including rodents, pigs and carnivores—that typically give birth to a wide range of litter sizes and live in small kin groups.

Although female lions do nurse the offspring of other females, they try to give milk primarily to their own cubs and reject the advances of other hungry cubs. But they also need sleep. When they doze for hours at a time, they present the cubs with an enormous temptation. A cub attempting to nurse from a lioness who is not its mother will generally wait until the female is asleep or otherwise distracted. The females must therefore balance the effort needed to resist the attentions of these pests against their own exhaustion.

Generosity among female lions, then, is largely a matter of indifference. Females that have the least to lose sleep best—owing either to the small size of

AFFECTION is common among pridemates, which rely on one another to help protect their young. Male lions present one of the greatest threats: if one coalition takes over a new pride, the newcomers—eager to produce their own offspring—will murder all the pride's small cubs and drive the older cubs away.

their own litter or to the company of close relatives. Female spotted hyenas have resolved this conflict by keeping their cubs in a well-protected den. Mothers return to their cubs for short periods, feed their brood and then sleep somewhere else in peace. By watching hyenas at the den, we found that mother hyenas received as many nursing attempts from the cubs of other females as did mother lions, but the hyenas were more alert and so prevented any other than their own offspring from nursing.

Surviving in the Serengeti

As we have seen, female lions are most gregarious when they have dependent young; the crèche is the social core of the pride. Childless females occasionally visit their maternal companions but generally keep to themselves, feeding well and avoiding the social complexities of the dining room or nursery. Mothers do not form a crèche to improve their cubs' nutrition. And gregarious mothers may actually eat less than solitary mothers; they have no system of baby-sitting to ensure a more continuous food supply. Instead mother lions form a crèche only to defend themselves and their cubs.

A female needs two years to rear her cubs to independence, but should her cubs die at any point, she starts mating within a few days, and her interval between births is shortened by as much as a year. Male lions are rarely affectionate to their offspring, but their territorial excursions provide effective protection. Should the father's coalition be ousted, however, the successors will be in a hurry to raise a new set of offspring. Any cubs left over from the previous regime are an impediment to the new coalition's immediate desire to mate and so must be eliminated. More than a quarter of all cubs are killed by invading males. The mothers are the ultimate victims of this never-ending conflict, and they vigorously defend their cubs against incoming males. But the males are almost 50 percent larger than the females, and so mothers usually lose in one-on-one combat. Sisterhood, on the other hand, affords them a fighting chance; in many instances, crèchemates succeed in protecting their offspring.

Male lions are not their only problem. Females, too, are territorial. They defend their favorite hunting grounds, denning sites and water holes against other females. Large prides dominate smaller ones, and females will attack and kill their neighbors. Whereas most males compress their breeding into a few short years, females may enjoy a reproductive life span as long as 11 years. For this reason, boundary disputes between prides last longer than do challenges between male coalitions, and so the females follow a more cautious strategy when confronted by strangers. Karen E. McComb, now at the University of Sussex, found that females would attempt to repel groups of tape-recorded females only when the real group outnumbered the taped invaders by at least two. Females can count, and they prefer a margin of safety. Numbers are a matter of life and death, and a pride of only one or two females is doomed to a futile existence, avoiding other prides and never rearing any cubs.

The lions' pride is a refuge in which individuals united by common reproductive interests can prepare for the enemy's next move. The enemy is other lions—other males, other females—and they will never be defeated. Over the years, we have seen hundreds of males come and go, each coalition tracing the same broad pattern of invasion, murder and fatherhood, followed by an inevitable decline and fall. Dozens of prides have set out to rule their own patch of the Serengeti, but for every new pride that has successfully established itself, another has disappeared. Lions can seem grand in their common cause, battling their neighbors for land and deflecting the unwanted advances of males. But the king of beasts above all exemplifies the evolutionary crucible in which a cooperative society is forged. SA

The Authors

CRAIG PACKER and ANNE E. PUSEY are professors in the department of ecology, evolution and behavior at the University of Minnesota. They conducted their studies at the Serengeti Wildlife Research Institute, the University of Chicago and the University of Sussex. Packer completed his Ph.D. in 1977 at Sussex. That same year, Pusey received her Ph.D. from Stanford University.

Further Reading

A Molecular Genetic Analysis of Kinship and Cooperation in African Lions. C. Packer, D. A. Gilbert, A. E. Pusey and S. J. O'Brien in *Nature,* Vol. 351, No. 6327, pages 562–565; June 13, 1991.

Into Africa. Craig Packer. University of Chicago Press, 1994.

Non-offspring Nursing in Social Carnivores: Minimizing the Costs. A. E. Pusey and C. Packer in *Behavioral Ecology,* Vol. 5, No. 4, pages 362–374; Winter 1994.

Complex Cooperative Strategies in Group-Territorial African Lions. R. Heinsohn and C. Packer in *Science,* Vol. 269, No. 5228, pages 1260–1262; September 1, 1995.

Managing Human Error in Aviation

Mistakes by flight crews contribute to more than two thirds of aviation accidents. Training to enhance team performance may reduce potentially fatal errors

by Robert L. Helmreich

In 1978 a United Airlines DC-8 carrying 189 people crashed while attempting to land in Portland, Ore. Immediately after lowering the landing gear on the approach to the airport, the pilots noticed that an indicator light had failed to go on. The failure implied that one set of wheels and its support structure might collapse on landing, potentially causing a fire or otherwise leading to injuries. Instead of continuing the approach, the crew decided to circle in a holding pattern while they determined if the landing gear was indeed compromised. As the delay increased, the fuel became dangerously low. The captain, preoccupied with the light, failed to monitor the overall situation and ignored repeated warnings from the flight engineer about the dwindling fuel. By the time the captain reacted and tried to land, it was too late. All four engines quit, and the airplane crashed in a wooded area short of the runway, killing 10 of the people on board.

The accident investigation revealed that the only problem with the airplane was that the warning light had malfunctioned. The captain's error was not his attempt to deal with a potentially life-threatening mechanical problem but rather his failure to maintain awareness of other critical aspects of flying an aircraft under highly stressful conditions.

This accident coincided with investigations conducted by the National Aeronautics and Space Administration into the causes of airline accidents since the introduction of highly reliable turbojet aircraft in the late 1950s. The research showed clearly that more than 70 percent of airline accidents involved some degree of human error. More surprising was that most of these errors stemmed from failures in communication, teamwork and decision making rather than from technical shortcomings.

The airline industry was shocked to realize that well-trained and technically proficient crews could crash airworthy craft because of failures of human interaction and communication—areas in which neither training nor formal evaluation was required by the Federal Aviation Administration (FAA) or any other country's aviation regulatory agency. This realization led to the development of programs—collectively known as crew resource management, or CRM for short—that targeted the team and leadership aspects of piloting an aircraft. The programs, though focusing on the cockpit crew (pilot, copilot and flight engineer), also include flight attendants, air traffic controllers and other support staff. CRM extends beyond the cockpit because other aviation professionals have a role in determining the safety of flight.

Learning from Mistakes

A major goal of CRM is to get pilots to work as a team to reduce errors. Having two or three crew members in the cockpit provides a measure of redundancy; one person may notice something that escapes the attention of another. But multiple perspectives on the conditions of a flight are useless unless the information is shared. Through training focusing on the inherent limitations of human performance, including the impact of stress on the ability to absorb information and make decisions, pilots and other flight personnel become more aware of the importance of collaboration as a countermeasure against error.

CRM is grounded in social, cognitive and organizational psychology as well as in human factors research, which focuses on how people interact with machines. Individual programs are tailored separately to each airline. The airline—usually with the help of outside experts—first conducts an organizational analysis to identify procedures that might impede training. For instance, the airline might survey pilots' assessments of the airline's culture, information that can sometimes identify unsafe practices embedded within an organization. These surveys are supplemented by observations of the behavior of an airline's crews during routine flights.

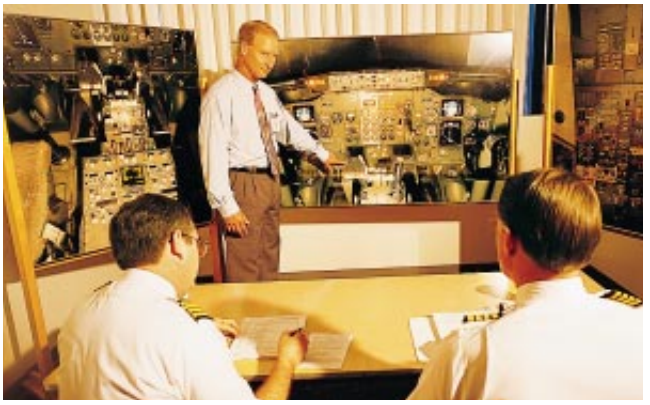
A still broader understanding of the problems experienced within an airline may be derived from analyses of pilot errors. To gather such data, an airline might adopt a nonpunitive policy toward the reporting of mistakes, to encourage pilots to share their experiences. One airline that instituted such a policy received more than 5,000 reports from its pilots in 21 months. The volume of reports does not indicate that this airline is unsafe; rather it highlights the number of errors that occur during normal flights but that are usually caught and corrected without consequence.

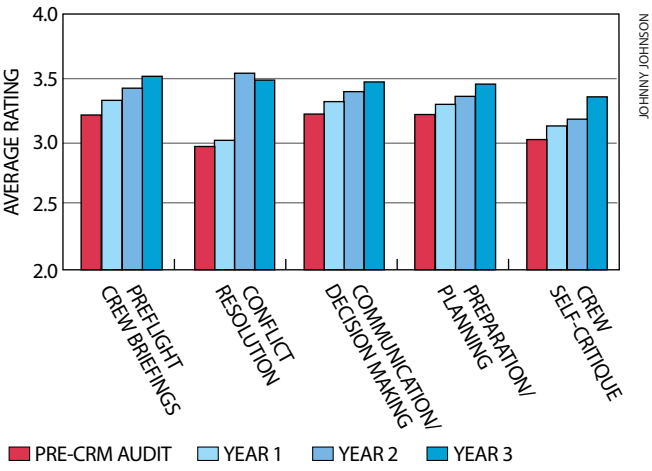
A typical CRM program begins with a seminar that provides background in

J. L. ATLAN *Sygma*

CRASH of an Air Florida Boeing 737 into the Potomac River near Washington, D.C., in 1982 occurred after a pilot failed to heed the copilot's warnings that the airplane was moving too slowly during the acceleration before takeoff.

PHOTOGRAPHS BY ERIC O'CONNELL





JOHNNY JOHNSON

HIGH-FIDELITY FLIGHT SIMULATOR (*top left*) enables pilots undergoing crew resource management (CRM) training (*top right*) to hone collaboration skills. A trainer (*foreground, top right*) programs into a computer stressful flight situations—an engine failure, for instance. Later the pilots attend debriefing sessions in which they review the group's performance (*bottom left*). Gradual improvement in performance emerged in a study of the impact of CRM at one airline (*bottom right*).

group dynamics, the nature of human error and the issues that arise when people work with machines. Members of a cockpit crew are asked to review accident case studies that highlight the importance of the interactions among crew members. An often cited example is the 1982 crash of an Air Florida Boeing 737 near Washington National Airport. The crew took off with ice on the wings and ice in a sensor, which caused the speed indicators to read too high. Because of the erroneous speed reading, too little power was applied. As the following dialogue indicates, the first officer sensed a problem with the instrument readings and power setting, but he did not communicate his concerns clearly.

First officer: Ah, that's not right.
Captain: Yes, it is, there's 80 [referring to ground speed].
First officer: Nah, I don't think it's right. Ah, maybe it is.
Captain: Hundred and twenty.
First officer: I don't know.

Shortly after takeoff, the plane stalled and crashed into a bridge over the Potomac River. Accidents are also the basis for many of the scenarios that pilots confront in flight simulators. Sessions in a simulator are part of the annual training to reinforce basic CRM concepts.

Beyond Stick and Rudder

High-fidelity simulators consist of a cockpit with working instruments and controls, the sensation of motion and a visual representation of the environment outside the cockpit windows. CRM has expanded the use of the simulator as a training tool. Initially simulators were employed only to teach and evaluate pilots' flying skills ("stick and rudder" techniques). Today they enable crews to test themselves in tackling complex problems—ranging from bad weather to mechanical failures—that cannot be resolved by simply following a procedure outlined in the flight manual. (CRM in the simulator is known as line-oriented flight training, or LOFT.)

During a LOFT session, a full crew conducts a complete flight, beginning with the necessary paperwork and crew briefings. An instructor who has received special training in analyzing group behavior directs the session. The instructor also plays the roles of air traffic controller and flight attendant in communication with the cockpit.

The simulator might be programmed for an engine failure to see if the crew remains fixated on this problem and ignores other factors such as fuel or weather, as did the Portland DC-8 captain. Each member of a flight crew is encouraged to develop an overall understanding of how a flight is unfolding—and to communicate any concerns that may emerge. The training session is not a test, so the crew is free to experiment with new behaviors: a captain who is normally rigidly controlling might choose to solicit advice from a copilot, for example. An important component of LOFT is the postsimulation debrief-

ing, in which video recordings of the simulator session are reviewed for both flawed and exemplary behaviors by the flight crew.

CRM assists pilots in interacting not only with one another but with the flight management computer in the cockpit, which is sometimes described as an "electronic crew member." The courses teach that automation of aircraft flight systems, from navigation to landing, has not proved to be the safeguard against error that many aircraft designers envisioned. In fact, automation has heightened the need for crews to communicate more effectively. Some of the most sophisticated commercial airliners have been involved in accidents that resulted from the misuse of automated systems.

Computerized flight systems pose a number of challenges for crews. Although the technology can guide an aircraft with precision, a long journey under the control of the flight computer can be an exercise in tedium, complicating the task of maintaining a high level of vigilance. Further, programmed flight computers can produce confusing results. In our studies, we have often heard crew members ask, "What is it doing? What will it do next?" Although computerized flight systems are frequently described as being "smart," in reality they will faithfully execute erroneous commands, sometimes with tragic results.

In older, less automated aircraft, jokingly called "Jurassic jets," changes in flight controls by one crew member to alter altitude or speed settings are immediately apparent to others. They can see one another move the instrument dials. A change is immediately reflected on a gauge on the control panel. In automated aircraft, one crew member enters a change—say, in altitude—into a keyboard, and so the alteration may not be observed by another crew member, thereby eliminating a valuable cross-checking procedure. As a result, airplanes equipped with advanced automated equipment actually require a greater level of communication among crew members to ensure that everyone understands what is taking place. Ideally, the pilots should verbalize every keyboard entry so that others can catch errors before disaster occurs.

When Teamwork Saved Lives

The 1989 flight of a United Airlines DC-10 over the Midwest is an often studied case for its lessons on how a crew handled many tasks at once while facing imminent disaster. During Flight 232 from Denver to Chicago, the center engine disintegrated, severing the hydraulic lines. Pressure from fluid in the lines is needed to move the rudder, ailerons and other control surfaces that maneuver the aircraft. Because of the failure, the pilots were unable to control the direction of the airplane.

In this dire situation, the three-member crew became the model of a team, even recruiting a pilot seated in the first-class section for assistance. Together they devised a technique to steer the aircraft by increasing and decreasing power from the two remaining engines. Although the airplane hit the ground just short of the runway, many passengers survived. In its accident report, the National Transportation Safety Board singled out the crew's performance and cited the value of its training in crew resource management (CRM).

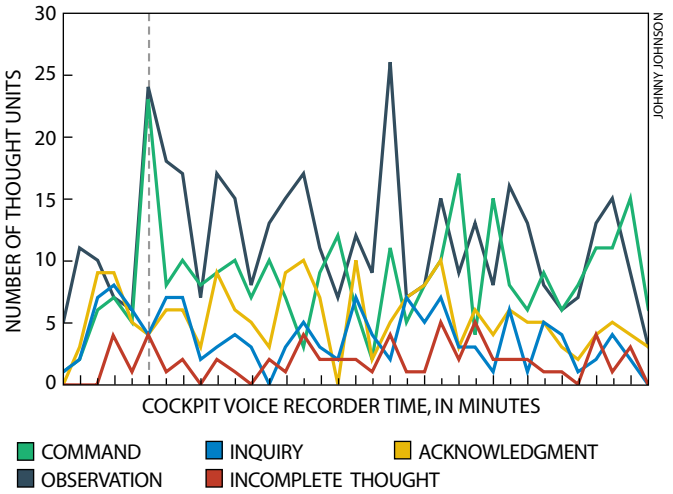
To further understand the dynamics of this case history, Steven C. Predmore, now at Delta Airlines, analyzed the cockpit voice recordings of the crew as part of his doctoral research in my University of Texas group. Predmore classified what the crew said to one another into "thought units" that quantified a single thought, intent or action.

The crew had to deal with controlling the aircraft, assessing damage, choosing a landing site and preparing the cabin crew and passengers for an emergency landing. Unlike the crew that crashed in Portland, this one never fixated on any single task. It coped with multiple issues simultaneously during the 34 minutes of the recording.

The crew members also effectively prioritized their work, abandoning tasks that could distract them. For example, about 12 minutes before the crash, the crew shifted its focus from corrective action and assessing damage to concentrating on executing the descent. (Of the 296 people on board, 111 died; all the pilots survived.)

Predmore's investigation points to the sheer volume of communication that took place among crew members. At its peak, 59 thought units were transmitted during a single minute, some as brief as a hastily uttered "okay." The overall average was 31 a minute, about twice the level encountered during demanding periods of routine flight. The nature of the interactions—a continuing effort to describe the situation at hand—demonstrates how the crew managed to keep one another aware of the events unfolding and how they went about making decisions from this information.

Not only were appropriate commands issued, but junior crew members were free to suggest alternative courses of action. Bursts of social conversation—providing emotional support or inquiring about the level of anxiety being experienced by others—were interspersed throughout the transcript; these asides proved to be an effective way of coping with the overwhelming stress confronted by the crew. —*R.L.H.*



MINUTE-BY-MINUTE analysis of the cockpit voice recorder showed intensive communication among crew members before the DC-10 crashed in Sioux City, Iowa, in 1989. The types of interactions were known as "thought units." The total number of units peaked at 59, during the fifth minute shown (*gray line*). The number is the sum of all communications in that minute.

Automation can also have the unintended effect of increasing pilots' workload at the times they should be looking for traffic and managing navigation in the congested sky around an airport. For example, a change in the runway assigned for landing may require extensive keyboard entries into the flight management computer, which controls speed, direction and altitude. Riding as an observer in the cockpit of automated aircraft, I have seen both pilots, looking down, reprogramming the approach into the computer while the airport was clearly visible out the window. Disengaging the automation would have allowed them to reach the runway without becoming preoccupied and losing awareness of traffic and other threats in busy airspace.

CRM attempts to get pilots to think of the automated cockpit system as another crew member—one with specific strengths and weaknesses. In the simulator, pilots may confront scenarios in which their ability to cope with the tasks at hand may be taxed unless they switch the automation off.

PHOTOGRAPHS BY ERIC O'CONNELL

COCKPIT INSTRUMENTATION in older aircraft (*top*) allows the pilot to see changes readily in speed or direction of the aircraft. In a newer airplane (*bottom*), the pilot types into a computer the course to follow after takeoff, as the crew member in the left seat is doing. If distractions mount, however, the entry may not be noticed by the other pilot.

CRM goes beyond just providing crews with an understanding of the role of automation. When pilots enter the cockpit, they carry the baggage of three cultures—the professional culture of the pilot, the business culture of an airline and national culture. All these influences can affect their performance. Ashleigh C. Merritt, a postdoctoral member of our research group at the University of Texas at Austin, which has been studying team performance and CRM for 20 years, investigated the three cultures as part of her doctoral dissertation at the university.

In her work, she found that the culture of pilots is a strong one—exemplified by the rugged individualism vividly portrayed by Tom Wolfe in *The Right Stuff*. (The book details the lives of the test pilots who first broke the sound barrier and the astronauts in the early manned space program.) In addition to having great professional pride, many pilots strongly deny susceptibility to stress—they are unwilling to acknowledge that fatigue and sudden danger can dull their thinking and slow response times. This sense of invulnerability can manifest itself in a desire to play the role of the white-scarfed, lone aviator battling the elements. By emphasizing the limits of human performance and the inevitability of error, CRM attempts to alter these ingrained beliefs and make pilots see that working as a team can help avoid mistakes that stem from high stress and a large workload.

Culture of the Airline

The corporate culture of an airline also promotes or detracts from safety. Training alone is unlikely to produce lasting changes in behavior if an airline has little commitment to CRM—the reason much of the initial work by outside consultants focuses on analyzing airline procedures and management practices. The investigation of an Air Ontario Fokker F-28 that crashed in Canada in 1989 after taking off with ice on the wings in a severe snowstorm demonstrates an example of lapses within an organization. Pilots were accustomed to looking out the window to see if snow had blown off when an aircraft had reached 80 knots during the acceleration for takeoff. This nonstandard check usually worked well with slower, turboprop planes, whose straight wings were readily observable from the cockpit. But it did not readily detect the presence of ice on a jet's swept-back wings.

Merritt's data from surveys of more than 13,000 pilots in 16 countries also found that behavior in the cockpit was influenced by nationality. The popular view of the cockpit as a "culture free" zone, in which a pilot from any country performs the same tasks identically in the same kinds of aircraft, is a myth. Perceptions of the appropriate roles for captains and junior crew members and attitudes about the importance of rules and written procedures differed significantly from one culture to another. In a survey in an Asian country, only 36 percent of pilots agreed that crew members should voice concerns about the safety of a flight, whereas that figure climbed to 98 percent in one Western country.

Our research suggests that no nation's culture produces the "ideal" crew. The optimal crew would be strongly oriented toward teamwork and a consultative style of leadership in which junior officers felt encouraged to speak up to share information and advocate alternative courses of action. The most effective crew would adhere to standard procedures but could still use their judgment to deviate from rules in the interest of safety. The challenge for CRM developers is to harmonize the training with local conditions and culture.

Any airline that undertakes a full program of CRM must be prepared to commit substantial time and resources. It is fair to ask, then, whether the training

Can CRM Reduce Human Error in the Operating Room?

A distinguished neurosurgeon persists in operating on the wrong side of a woman's brain, in spite of vague protests by a resident who is aware of the error. In another hospital operating room, a surgeon and anesthesiologist resolve their differences by fisticuffs while an elderly patient lies anesthetized on the table.

Several physicians, including David M. Gaba of Stanford University and the late Hans-Gerhard Schaefer of the University of Basel/Kantonsspital in Switzerland, have recognized parallels between interpersonal communications problems in the cockpit and those encountered in the operating room. Their work has resulted in the development of human factors training programs for surgical and anesthetic teams. This type of training is now being practiced around the world.

R. F. SMITH *Sygma*

DUMMY PATIENT is used in mock surgical operations by a medical team in Britain being trained with techniques derived from crew resource management (CRM) for pilots.

The effort in Basel most closely resembles the aviation model. The program there is built around an operating room simulator in which an instrumented, computer-controlled mannequin (nicknamed Wilhelm Tell) can be anesthetized and can undergo laparoscopic, abdominal surgery. The "patient" breathes, coughs, responds to the introduction of drugs and bears the liver of a recently slaughtered pig with circulating (artificial) blood.

During the operation itself, various crises may be introduced by the computer—a hemorrhage, collapsed lung or cardiac arrest—that increase the need for joint decision making. The session is videotaped and is followed by a debriefing. Similar training may one day extend to emergency rooms, helicopter and ambulance emergency teams, intensive care units and delivery rooms. *—R.L.H.*

achieves its stated goals of reducing error and improving team effectiveness. The question cannot be answered simply. Because commercial aviation is, in fact, extremely safe, it will take many years to collect adequate statistics to determine whether CRM programs have reduced the overall frequency of air crashes. There are, however, a number of other measures that can be used to judge its usefulness. For instance, we now have data from more than 8,000 flights tracked at major airlines. At one airline, for which we have amassed data annually over a four-year period, we noted steady improvements in such factors as the ability to distribute workload.

Based on accumulating evidence of the value of CRM, the FAA has moved to make the training mandatory for flight crews at all major and regional airlines. Many airlines have initiated joint CRM courses for pilots and flight attendants that focus on the coordination of cockpit and cabin activity during emergencies. The FAA has also developed CRM for air traffic controllers. And the International Civil Aviation Organization, the United Nations agency that regulates worldwide aviation, requires CRM for all airlines that operate internationally.

Although the potential benefits of CRM seem clear, a small subset of pilots, roughly 5 percent, rejects its lessons. The few pilots whose attitudes toward CRM concepts worsen after training have become known as "boomerangers." Pilots who actively reject CRM practices pose serious threats to safety. Airlines have responded by placing increased emphasis on selecting pilots who are not only technically competent but also show themselves able to function as members of a team.

The lessons drawn from data on the performance of flight crews under stress in aviation are being generalized to other high-risk professions—for example, medical operating rooms [*see box above*], ships and the control rooms of nuclear and petrochemical plants. The use of the technique outside aviation suggests that it is proving to be an effective training strategy in any environment where a team of professionals interacts with a complex technological system. SA

The Author

ROBERT L. HELMREICH is professor of psychology at the University of Texas at Austin, where he has taught since 1966. He is principal investigator of a research project, initially funded by the National Aeronautics and Space Administration and currently supported by the Federal Aviation Administration, that examines the individual and team performance of flight crews as well as the influence of organizational and national culture in the cockpit. He also serves as visiting professor at the University of Basel/Kantonsspital in Switzerland, where he is studying interpersonal issues in medicine. In addition to those cited in the text, the research group at the University of Texas includes John A. Wilhelm (co-investigator), John Bell, Roy Butler, Peter Connelly, William E. Hines, James Klinect, Lou E. Montgomery, Sharon Jones Peeples, Bryan Sexton and Paul J. Sherman.

Further Reading

WHY CREW RESOURCE MANAGEMENT? EMPIRICAL AND THEORETICAL BASES OF HUMAN FACTORS TRAINING IN AVIATION. R. L. Helmreich and H. C. Foushee in *Cockpit Resource Management*. Edited by E. L. Wiener, B. Kanki and R. Helmreich. Academic Press, 1993.

TEAM PERFORMANCE IN THE OPERATING ROOM. R. L. Helmreich and H.-G. Schaefer in *Human Error in Medicine*. Edited by M. S. Bogner. Lawrence Erlbaum Associates, 1994.

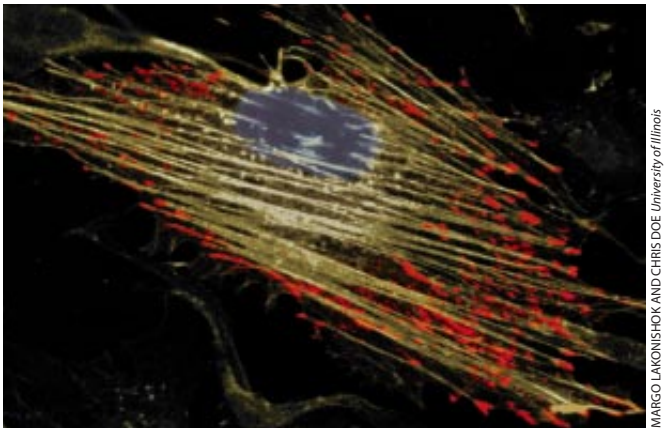
HUMAN FACTORS ON THE FLIGHTDECK: THE INFLUENCE OF NATIONAL CULTURE. Ashleigh C. Merritt and Robert L. Helmreich in *Journal of Cross-Cultural Psychology,* Vol. 27, No. 1, pages 5–24; January 1, 1996.

Home page on the World Wide Web for the Aerospace Crew Research Project at the University of Texas at Austin: http://www.psy.utexas.edu/psy/helmreich/nasaut.htm

Integrins and Health

Discovered only recently, these adhesive cell-surface molecules have quickly revealed themselves to be critical to proper functioning of the body and to life itself

by Alan F. Horwitz



MARGO LAKONISHOK AND CHRIS DOE *University of Illinois*

CELL has been colored to reveal the actin fibers (*tan*) in its internal scaffolding and the integrins (*orange*) that link the scaffolding to the protein fibronectin in the external matrix.

The cells of the body stick to one another and to the packing material, or extracellular matrix, around them. As might be expected, this adhesion holds tissues together and is therefore essential to survival. Less obviously, it helps to direct both embryonic development and an array of processes in the fully formed organism, including blood clotting, wound healing and eradication of infection. Unfortunately, the stickiness of cells can also contribute to a number of disorders, among them rheumatoid arthritis, heart attack, stroke and cancer.

Although scientists have long recognized the importance of adhesive interactions in the body, until recently they knew little about how such interactions exert their diverse effects on physiology. The fog began to lift about 20 years ago, when investigators isolated some of the matrix molecules that stick to cells. During the past 15 years, they have learned that cell-surface molecules called integrins are central players in many adhesion-related phenomena. Not surprisingly, drugmakers are already capitalizing on the findings to develop novel treatments for a number of diseases.

I feel fortunate to have been among the investigators who identified the first integrins and uncovered their activities. But the integrin story is not the tale of a single laboratory. More than in many areas of biology, understanding of integrins has unfolded through the cooperation of teams exploring widely divergent processes. Some of us started with an interest in embryonic development. Others were more concerned with the functioning of the mature body or the progression of specific diseases. The extent of the cross-fertilization and the pace of progress have been nothing short of exhilarating.

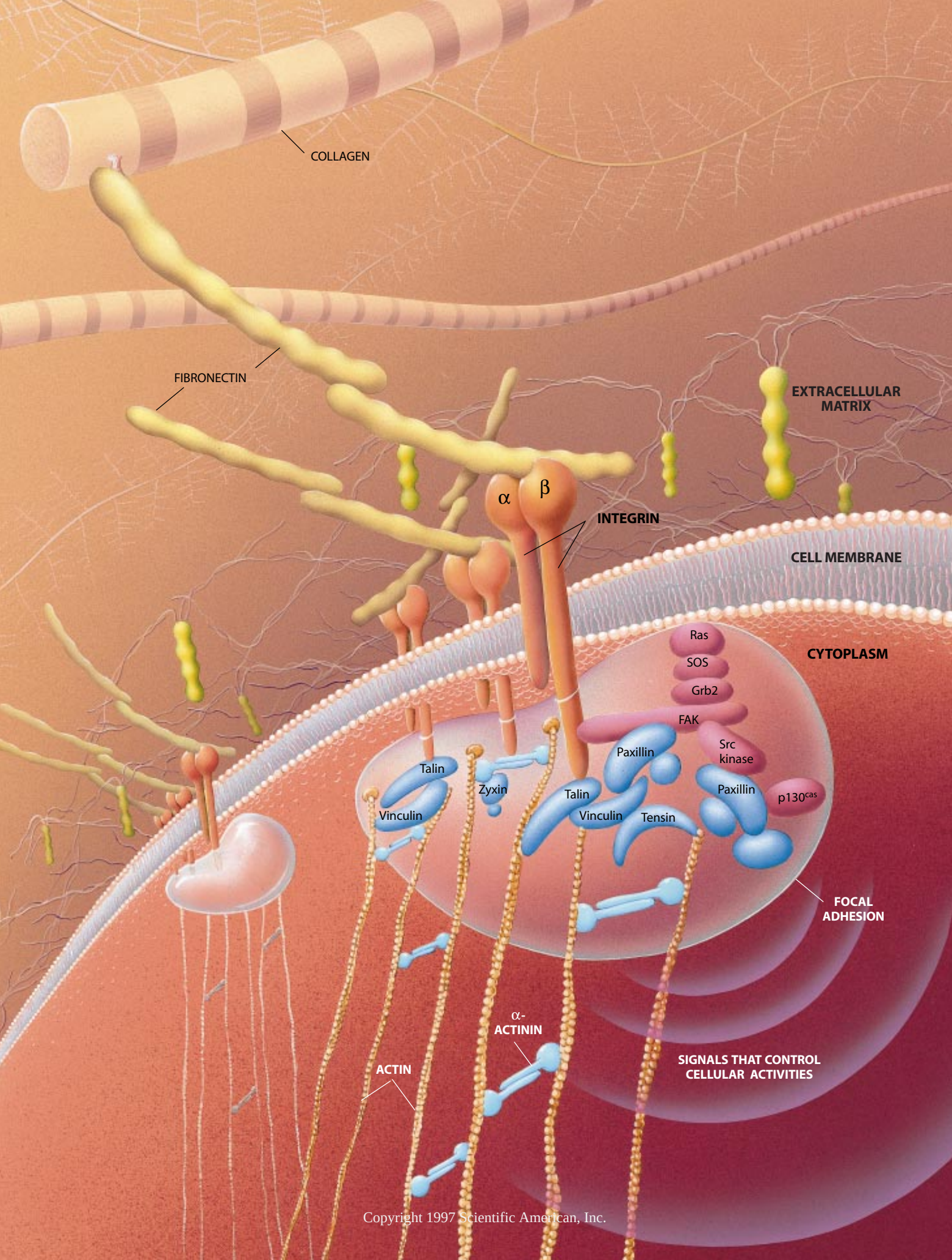
A dramatic example of the importance of adhesion to proper cell function comes from studies of the interaction between matrix components and mammary epithelial cells. Epithelial cells in general form the skin and the lining of most body cavities; they are usually arranged in a single layer on a specialized matrix called the basal lamina. The particular epithelial cells that line mammary glands produce milk in response to hormonal stimulation. If mammary epithelial cells are removed from mice and cultured in laboratory dishes, they quickly lose their regular, cuboidal shape and the ability to make milk proteins. If, however, they are grown in the presence of laminin (the major adhesive protein in the basal lamina), they regain their usual form, organize a basal lamina and assemble into glandlike structures capable once again of producing milk components.

By the early 1980s scientists interested in how the extracellular matrix can control the activity of adherent cells had made some headway in studies focused on the matrix itself. They knew that the matrix consists primarily of gel-like chains of sugars and interconnected fibrous proteins, although the amount of matrix and the details of its structure can vary from one tissue to the next. The proteins include laminin and fibronectin (another adhesion molecule) as well as collagen, which is sometimes adhesive but is the primary structural component of most matrices. And microscopy had indicated that adhesive matrix molecules were linked—presumably through one or more intermediary molecules—to the system of intracellular fibers (the cytoskeleton) that gives cells their three-dimensional shape.

Investigators were also well aware that formation of attachments between cells and a matrix can affect the cells in any number of ways; the response of

TOMO NARASHIMA

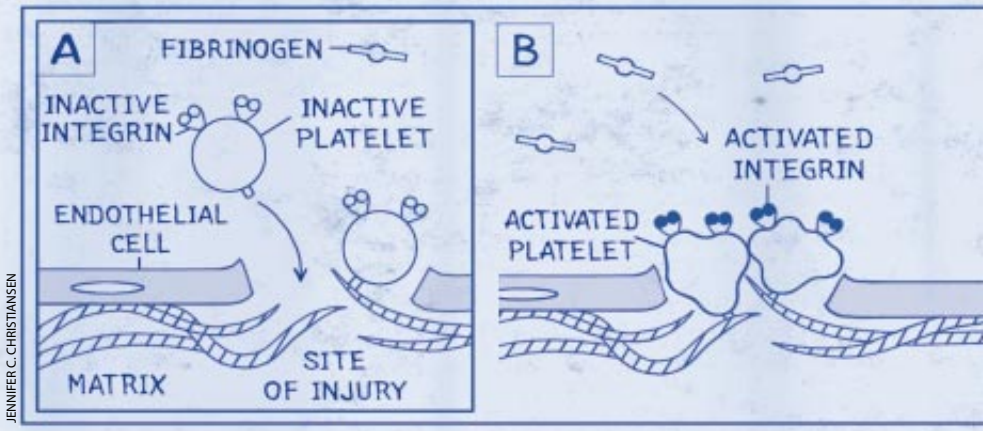
INTEGRINS (*orange*) span cell membranes. They hold a cell in place by attaching at one end to molecules of the extracellular matrix (or to molecules on other cells) and at the other end to the cell's own scaffolding, or cytoskeleton. They connect to this scaffolding through a highly organized aggregate of molecules—a focal adhesion—that includes such cytoskeletal components as actin, talin, vinculin and α-actinin. Integrins have also recently been found to relay messages from the matrix into the cell. The process seems to involve stimulation of dedicated signaling components (*magenta*) in focal adhesions. (The configuration of focal adhesions can vary.)



the cells depends on the type of cells involved, their state at the moment and the specific makeup of the matrix. Sometimes cells respond by changing shape. Other times they migrate, proliferate, differentiate (become more specialized) or revise their activities more subtly. Often the various changes issue from alterations in the activity of genes. Genes specify the sequences of amino acids in proteins, which carry out most cellular functions. When genes are active, or expressed, they give rise to the specified proteins. As genes are switched on and off, the mix of proteins in a cell changes and so does the operation of the cell.

To understand how the extracellular matrix elicits such diverse changes in cells, however, researchers needed to identify the receptors on cells that serve as the docking sites for matrix proteins. Scientists often isolate a receptor for a selected molecule by noting which components in cell extracts stick to copies of that molecule. But the receptors in this case were frustratingly difficult to isolate.

Part of the reason, we now see, was that adhesion molecules of the matrix bind to a variety of substances, such as other matrix components. This problem diminished when the cell-binding site on fibronectin was finally pinpointed. Remarkably, this region consists essentially of a sequence of just three amino acids—arginine, glycine and aspartate. (It is commonly called RGD, for the single letter codes assigned to those amino acids.) Another reason for the difficulty was that individual receptors for matrix molecules often do not bind tightly to matrix proteins. Strong adhesion is achieved through a kind of Velcro effect: many relatively weak attachments collectively form a stronger one. Weak binding by the individual receptors meant they did not remain attached long enough to allow for ready detection.

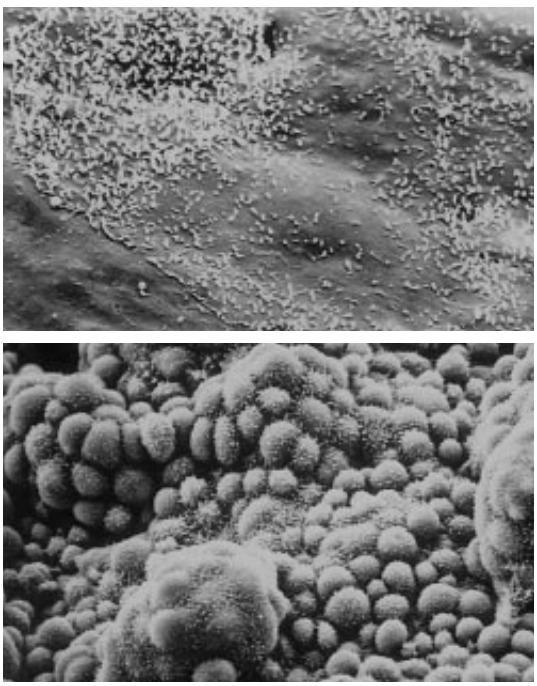


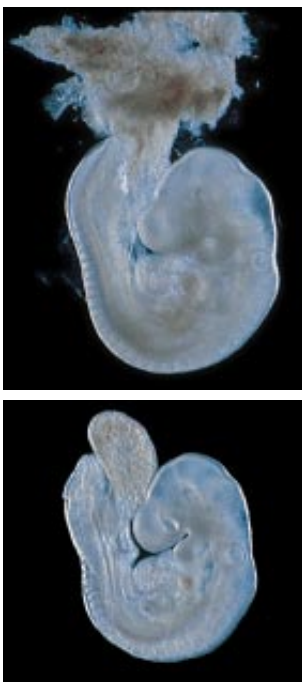
All in the (Integrin) Family

In spite of the difficulties, by the mid-1980s scientists had managed to isolate several cell-surface adhesion receptors. They had also determined the sequences of some of the amino acids in the molecules. In 1987 the combined results revealed that the receptors belonged to a large family of structurally related molecules, one or more of which appeared on virtually every cell type in the animal kingdom. In recognition of the family's importance to the structural integrity of cells and tissues, the molecules were named "integrins." Later, as I shall explain, the name proved apt in yet another way.

Meanwhile research into the structure and function of integrins took off. The finding that integrins come in many flavors and that they occur on most cell types suddenly meant that questions difficult to address with one kind of cell could potentially be tractable with another. This flexibility and the sharing of discoveries from many disciplines led to quick progress and kept most of us brimming with new ideas to explore.

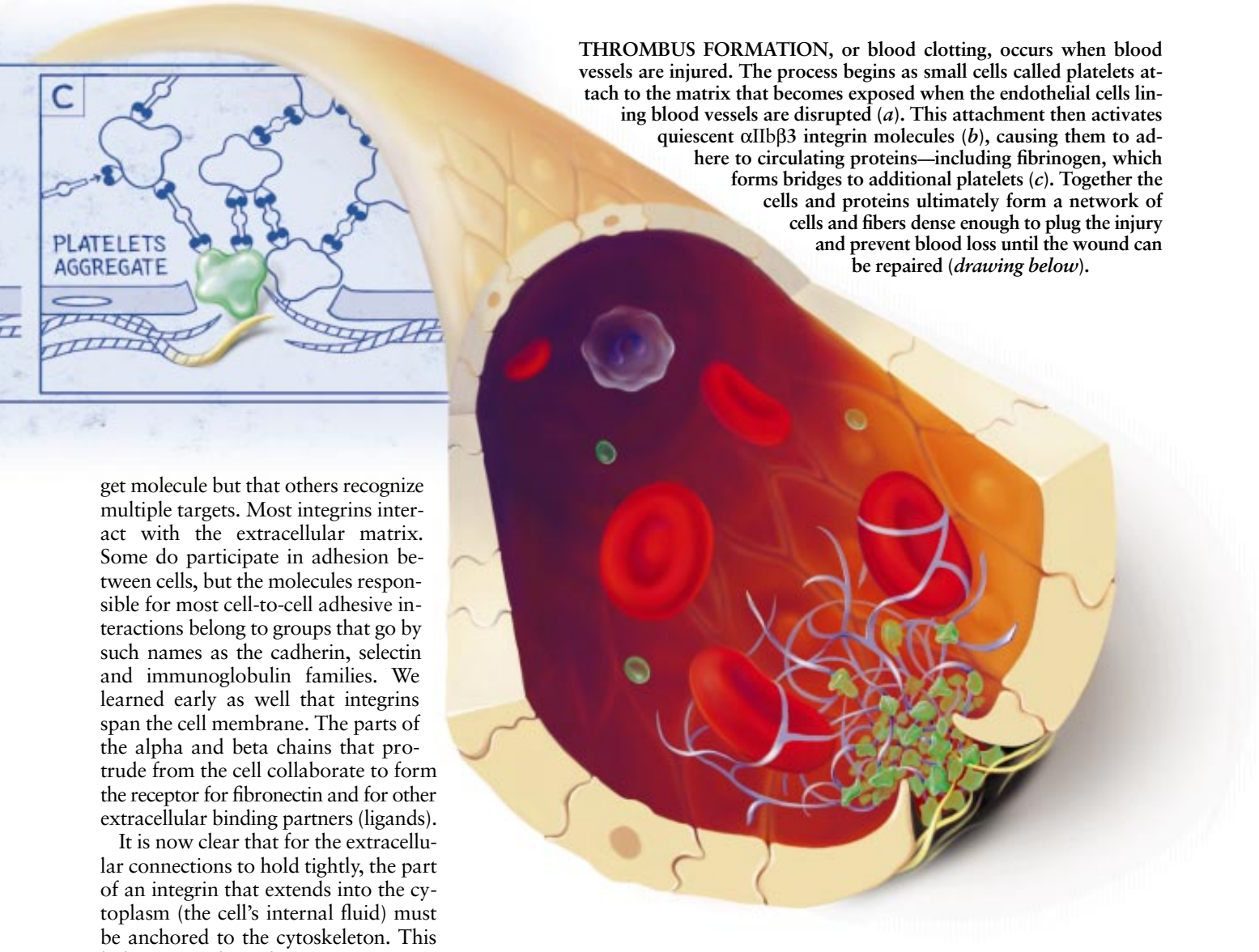
Early on we understood that integrins consist of two protein chains, or subunits. The subunit designated "alpha" is today known to have about 15 variants, and the "beta" subunit has about eight variants. The chains, which are generally named by a number or letter, combine into at least 20 different integrins. We also realized that certain integrins adhere to only a single type of tar-





left micrographs: COURTESY OF JOANNE EMERMAN *University of British Columbia* AND MINA J. BISSELL *University of California, Berkeley*
right micrographs: RICHARD O. HYNES AND JOY T. YANG *Howard Hughes Medical Institute, Massachusetts Institute of Technology*

MATURATION OF CELLS and embryos requires attachment of cells to the extracellular matrix, often via integrins. Mammary epithelial cells removed from mice and grown in a culture dish lose their normal shape and the ability to produce milk proteins (*top left*); they regain their structure (*bottom left*) and function only when specific matrix molecules are provided. A normal mouse embryo began to form a placenta (*top right*) by its ninth day of gestation. An embryo lacking the α4 integrin, however, did not form a placenta (*bottom right*) and soon died.



THROMBUS FORMATION, or blood clotting, occurs when blood vessels are injured. The process begins as small cells called platelets attach to the matrix that becomes exposed when the endothelial cells lining blood vessels are disrupted (*a*). This attachment then activates quiescent αIIbβ3 integrin molecules (*b*), causing them to adhere to circulating proteins—including fibrinogen, which forms bridges to additional platelets (*c*). Together the cells and proteins ultimately form a network of cells and fibers dense enough to plug the injury and prevent blood loss until the wound can be repaired (*drawing below*).

get molecule but that others recognize multiple targets. Most integrins interact with the extracellular matrix. Some do participate in adhesion between cells, but the molecules responsible for most cell-to-cell adhesive interactions belong to groups that go by such names as the cadherin, selectin and immunoglobulin families. We learned early as well that integrins span the cell membrane. The parts of the alpha and beta chains that protrude from the cell collaborate to form the receptor for fibronectin and for other extracellular binding partners (ligands).

It is now clear that for the extracellular connections to hold tightly, the part of an integrin that extends into the cytoplasm (the cell's internal fluid) must be anchored to the cytoskeleton. This linkage is achieved in an interesting way. When the receptor becomes bound by a molecule outside the cell and aggregates with other bound integrins, highly organized complexes known as focal adhesions form. These adhesions, which can vary in size and structure, incorporate a variety of molecules, including the cytoplasmic tails of the clustered integrins and components of the cytoskeleton. Receptor binding and aggregation also cause the cytoskeleton to become organized: cells that were rounded take on a definite structure.

Integrin-mediated attachment of cells to a matrix can also facilitate travel by peripatetic cells, notably those of developing organisms and the white blood cells of the immune system. Migration begins with the extension of footlike protrusions at the front of a cell. Next, integrin molecules near the leading edge form traction-providing links to the underlying matrix. At about the same time, similar linkages at the rear of the cell release, causing the back part of the cell to move forward, as if one end of a spring were let go. We do not yet know what forces induce the cells to release the matrix connections at the rear. Nor do we fully understand the biochemical and mechanical forces that produce forward motion of the front or the rear. Nevertheless, the cycle, once accomplished, is repeated.

From Structure to Function

Along with acting as cellular "glue" and at times facilitating migration, integrins have a third—quite powerful—effect on cells. By about 1990, attachment of integrins to molecules outside the cell was found to activate so-called signal transduction pathways on the inside. These pathways, consisting of molecules that relay messages through the cytoplasm, evoke such responses as gene expression, cell division and induction of processes that prevent cells from self-destructing.

Shortly after, scientists discovered that signaling by integrins can modulate the messages relayed into cells by growth factors. Growth factors, which can travel from one cell to another, had earlier been shown to activate specific signal transduction pathways that control which genes are expressed and whether a target cell reproduces or even continues to survive. Now studies of integrins demonstrated that to live and proliferate, normal (nonmalignant) cells must be attached to a specific matrix when they are stimulated by growth factors. If either growth factors or adhesive contacts are lost, a dividing cell will stop proliferating and will eventually die. In short, roughly three years after the integrins were named, they were found to live up to their name in a new way. They help to integrate many of the diverse signals that impinge on cells; in

so doing, they determine a cell's fate.

It seems that much of this internal signaling results from activation of molecules that reside with integrins in focal adhesions. This function has been studied most thoroughly in fibroblasts (connective tissue cells), where the complexes are quite large, consisting of more than 20 different molecules.

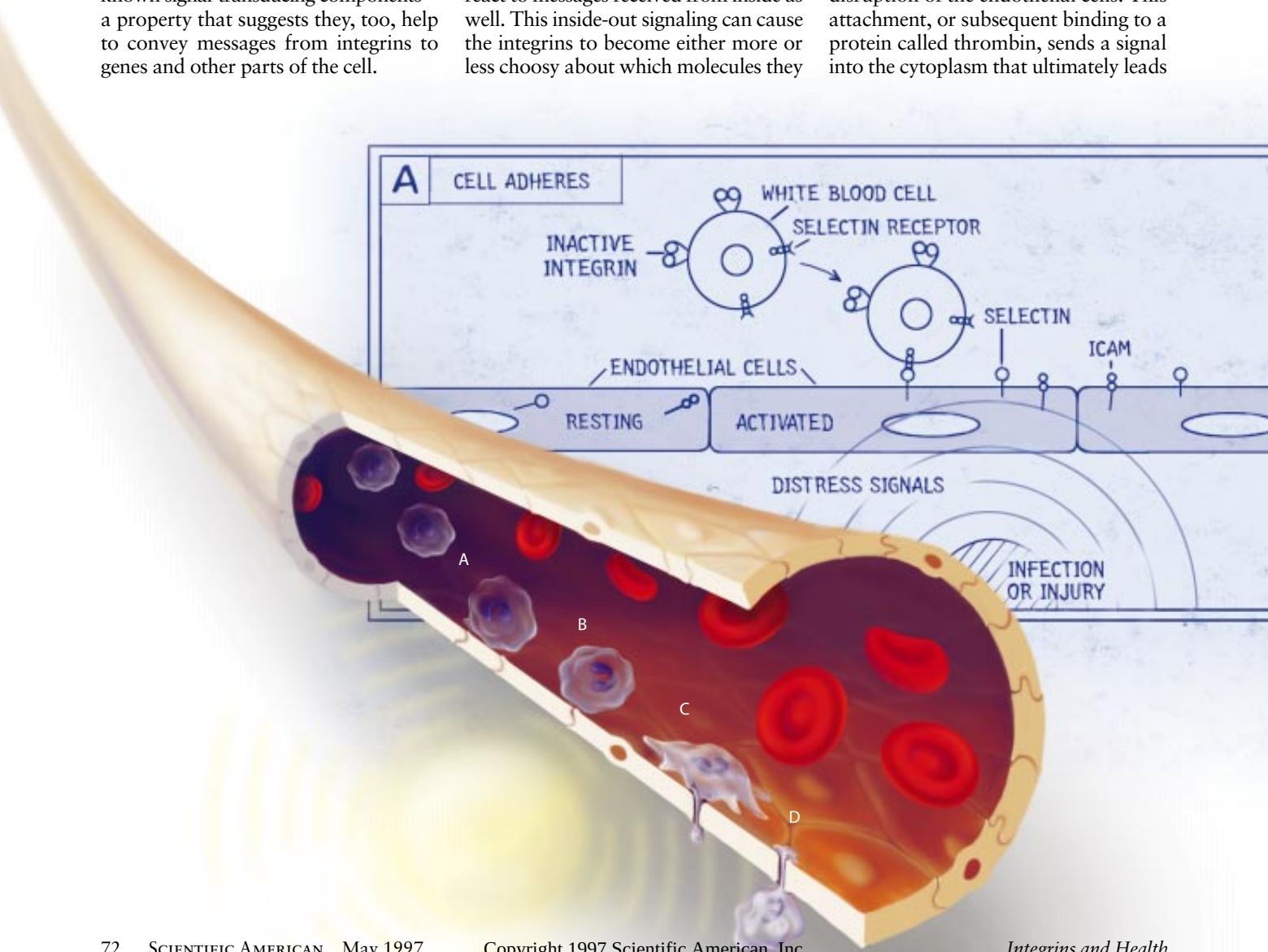
Certain of the molecules in those focal adhesions (such as enzymes of the Src kinase family) have long been known to participate in signaling pathways switched on by growth factors. Their presence in the complexes suggests that matrix molecules and growth factors may sometimes modulate one another's messages by sending signals down convergent or intersecting pathways. Other focal adhesion molecules that are activated by integrins (focal adhesion kinase, paxillin and tensin) typically are not much affected by growth factor activity. Yet they display binding sites for known signal-transducing components—a property that suggests they, too, help to convey messages from integrins to genes and other parts of the cell.

What is less clear is exactly how integrins stimulate signaling molecules in focal adhesions. Certain kinds of cell-surface molecules, notably many growth factor receptors, are tyrosine kinases: they add phosphate groups to tyrosine amino acids on other proteins and, in so doing, regulate the activity of target molecules. But integrins exhibit no kinase activity. Nor are they phosphatases—enzymes that control other molecules by removing phosphates from them. And integrins lack the usual docking sites by which typical signal-transducing molecules attach to one another—which means they probably do not alter those molecules directly. How, then, do the integrins elicit so much internal signaling? They might help signaling molecules to come into contact with one another, but no one yet has a definitive answer.

Integrins do not respond solely to signals received from outside the cell; they react to messages received from inside as well. This inside-out signaling can cause the integrins to become either more or less choosy about which molecules they will accept as binding partners, or such signaling can change the strength with which the integrins bind. The $\alpha 2\beta 1$ integrin, for example, can be inactive, a receptor for collagen or a receptor for both collagen and laminin, depending on the cell that produces it and the signals impinging on it from within the cell.

Inside-out signaling has been studied most thoroughly in platelets—small blood cells that lack a nucleus and help to form blood clots known as thrombi [*see illustration on preceding two pages*]. Thrombi form at injured areas of blood vessels and, like a finger in a dike, temporarily prevent blood from escaping. Platelets circulating in the blood travel singly and are nonadherent. They switch gears, however, when blood vessels, which are lined by a single layer of endothelial cells, become damaged.

First, the platelets stick (without the help of integrins) to areas of extracellular matrix that have become exposed by disruption of the endothelial cells. This attachment, or subsequent binding to a protein called thrombin, sends a signal into the cytoplasm that ultimately leads



to inside-out activation of the αIIbβ3 integrin on the platelet surface; in this case, the signaling causes the integrin to become more adhesive. Now the αIIbβ3 integrin grasps circulating molecules of fibrinogen or von Willebrand factor, which in turn form molecular bridges to more platelets and to the matrix. The resulting aggregate of platelets and proteins culminates in a dense meshwork of cells and fibers.

Integrins in Health and Illness

Beyond examining the effects of integrins on individual cells, scientists have been exploring the molecules' roles in the body as a whole. Ironically, understanding of how integrins direct the maturation of an embryo is sketchier than knowledge of how they participate in some other processes, even though curiosity about the molecular underpinnings of development fueled much of the research leading to the discovery of integrins and other adhesion molecules.

We do have proof, though, that integrins must be operational for development to proceed normally. As the cells of an embryo proliferate and differentiate to form the tissues and organs of the body, they add and subtract integrins from their surface, a sign that the added integrins are involved in the passage through each new stage. Another indication of the need for integrins is the discovery that the cells of embryos rely on particular integrins to help them travel to and find their final destinations.

Moreover, genetic engineering has let scientists produce animals, typically fruit flies and mice, that lack one integrin or another. The animals often become deformed or die during development, demonstrating that the absence of the "knocked out" integrins is at fault. In the malformed embryos, tissues become disorganized apparently because cells fail to travel to the proper places or to form the adhesive contacts required for cohesion. Indeed, in certain mutant fruit flies, newly completed muscles fall apart during the first contractions, as muscle cells detach from connective tissue.

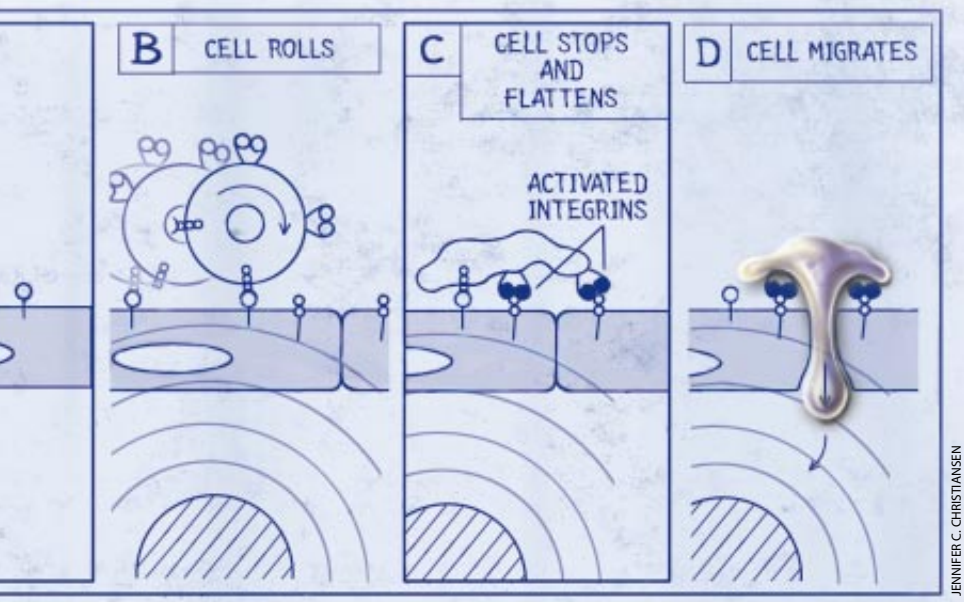
Ongoing research has also established the importance of integrins to normal physiology in mature organisms and has identified their contribution to various diseases. One critical process that requires integrins is inflammation—the complex set of responses that are set in motion by injury or infection. When tissue is damaged or colonized by a disease-causing microbe, certain white blood cells (leukocytes), such as neutrophils and monocytes, leave the bloodstream and race to the troubled region. There they mop up debris and foreign substances and attack any pathogens. The leukocytes additionally secrete substances that retard the spread of infection; if necessary, the cells also recruit other white cells known as lymphocytes to destroy invaders.

For white cells to make their way into compromised tissue, they must first be drawn from the main current of the blood. They are plucked out by endothelial cells that have detected a problem in the vicinity. Adhesive reactions that typically involve selectins, not integrins, cause the leukocytes to slow down and roll along the endothelial lining of the vessel walls. Then, inside-out signaling causes certain integrins on the leukocytes (mainly those containing β1 or β2 subunits) to gain affinity for molecules of the immunoglobulin family—particularly those called ICAMs (intercellular adhesion molecules)—on endothelial cells. These attachments help the leukocytes to stop, squeeze between endothelial cells and cross the blood vessel wall into the damaged or infected tissue.

The importance of integrins in inflammation is highlighted by a disease, called leukocyte adhesion deficiency, in individuals who lack the β2 integrin subunit or who produce a defective version. Because their leukocytes are unable to migrate to sites of injury and trauma, these people suffer from repeated life-threatening infections.

The inflammatory response protects the body from serious infections, but it can contribute to disease if it persists too long or occurs inappropriately. Hoping to improve treatment of disorders that involve chronic inflammation, investigators are developing a variety of compounds (mainly targeted to the α4β1 and β2 integrins) that interfere with integrin-mediated adhesion of white cells to endothelial cells. Indeed, such drugs are now being tested in patients with asthma, inflammatory bowel disease and rheumatoid arthritis.

Inflammation can also be destructive acutely, during reperfusion—the restor-



DEFENSE AGAINST INFECTION requires white blood cells (leukocytes) to leave the bloodstream and migrate into damaged or diseased tissues (*drawing at far left and details above*). First the leukocytes adhere weakly to endothelial cells through interaction with molecules called selectins (*a*) and then roll along the blood vessel wall (*b*). Next, integrins (usually containing β1 or β2 chains) on the white cells become activated (*c*) and adhere to ICAMs (intercellular adhesion molecules). These attachments cause the leukocytes to stop moving and to flatten out. The white cells then squeeze between endothelial cells (*d*), leave blood vessels and, using similar attachments, follow a trail of chemical attractants to the site of infection.

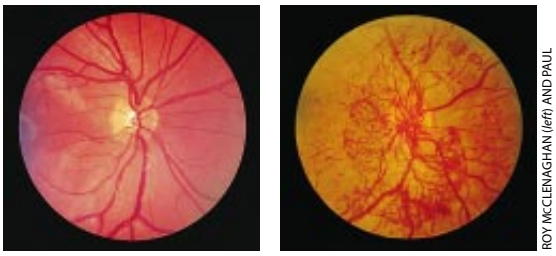
ation of blood flow to tissue that has been denied blood during, say, frostbite, heart attack or stroke. (Many heart attacks and strokes result from occlusion of a major blood vessel feeding the heart or brain.) The temporary loss of blood can kill and injure tissue. When blood flow is restored, neutrophils in the bloodstream sense the damage, migrate to the disrupted areas and release substances known as oxidants.

Oxidants can destroy pathogens, but they can also harm fragile cells and wreak further damage. Investigators are attempting to halt this so-called reperfusion injury with drugs that target either β2 integrins on neutrophils or their ICAM partners on endothelial cells; by keeping the integrins and ICAMs apart, such drugs should prevent neutrophils from leaving the bloodstream to invade reperfused tissue.

As is true of inflammation, too much or too little thrombus formation can be dangerous. People who lack the αIIbβ3 integrin (the one that is so important to platelet aggregation) suffer from Glanzmann's thrombasthenia; their platelets do not aggregate properly, and the victims bleed excessively. At the other end of the spectrum, excessive thrombus development, which often occurs where fatty deposits (atherosclerotic plaque) build up in blood vessels, may lead to heart attack or stroke if a thrombus obstructs a blood vessel.

Doctors have ways to clean out atherosclerotic arteries, such as by balloon angioplasty. Unfortunately, this process can be unkind to endothelial cells, and so worrisome thrombi occasionally form in the arteries within about a day after treatment. Further, in a process called restenosis, vessels in many patients can become occluded again over a period of many months, this time mainly by smooth muscle cells that migrate and proliferate, perhaps initially in response to substances present in thrombi that form near the sites of damage. Medicines that temporarily stop the αIIbβ3 integrin from interacting with fibrinogen are already in use to keep treated vessels clear of thrombi in the days after treatment. By retarding thrombus formation, the agents might additionally prove helpful in controlling restenosis.

Restenosis could potentially be avoided by targeting a different integrin: αvβ3, which appears rather prominently on smooth muscle cells after blood vessels are injured. Because this integrin seems to promote the survival and migration of these cells, blockade of its activity might limit the cells' contribution to restenosis.



ROY MCCLENAGHAN (*left*) AND PAUL PARKER (*right*) *SPL Photo Researchers, Inc.*

RETINA of a normal eye (*left*) is nourished by intact blood vessels, but the retina of a diabetic with proliferative retinopathy (*right*) contains a profusion of abnormal vessels that can damage the retina and cause blindness. Growth of new vessels requires abundant display of the αvβ3 integrin by endothelial cells. As the table at the right shows, this integrin is one of several now attracting the attention of drugmakers.

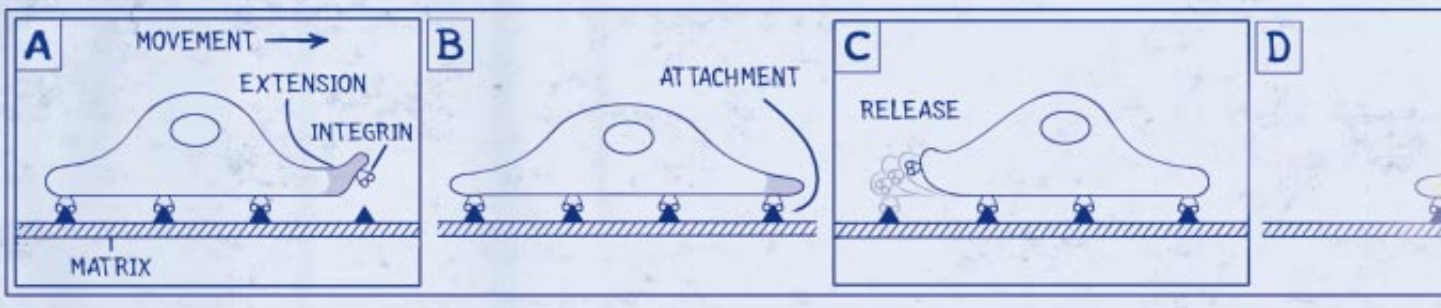
A Role in More Diseases

Other disorders involving undesirable activity by integrins include osteoporosis, a growing list of infectious disorders and cancer. Osteoporosis is familiar as the loss of bone, and attendant increased risk of fracture, that occurs with age, particularly in women. The condition can result from the relative overactivity of cells that bind to bone and degrade it. The binding in question occurs via the αvβ3 integrin, and so researchers are testing whether decoys that will essentially shield the integrins might prevent the destructive cells from adhering to bone.

In the case of infections, certain microbes apparently enter cells at least in part by latching onto integrins. These organisms include ones that produce flulike ailments, meningitis, diarrhea and paralysis. Such findings suggest the integrins that are involved could make good targets for new drug therapies.

Integrins clearly participate in cancer as well, but the details of what they do are incomplete. Cancer arises when cells evade the usual controls on cell division and migration; they then reproduce uncontrollably and gain the ability to invade local tissue and to metastasize—that is, to migrate to distant sites and to grow in unfamiliar territory. Researchers have found that various tumor types stop producing specific integrins, display integrins that their tissues do not normally make or alter their distribution of the usual integrins. The consequences of these changes are not always obvious, but the display of certain integrins is thought to sometimes facilitate migration by cells that normally would not travel.

Integrins can also promote the formation of blood vessels (angiogenesis) in tumors. These vessels nourish the tumors and provide access routes into the bloodstream for metastatic cells [see "How Cancer Spreads," by Erkki Ruoslahti, and "Fighting Cancer by Attacking Its Blood Supply," by Judah Folkman; Scientific American, September 1996]. To construct new blood



JENNIFER C. CHRISTIANSEN

CELLS "WALK" from one site to another by forming and breaking integrin-mediated attachments to a matrix. After attaching to a matrix, they produce footlike extensions (*a*) and form new attachments at their front end (*b*). Next, they release the connections at the rear (*c*), which causes the back end to move forward. Then the cycle begins anew (*d*).

Some Integrins under Study as Potential Drug Targets

Integrins contribute to a number of diseases. The pharmaceutical industry is now attempting to develop therapies that will prevent selected integrins, such as those listed below, from adhering to their usual binding partners in the body.

Integrin	Cell Types Displaying Integrin	Binding Partners	Examples of Disorders Involving Integrin Adhesion
αIIbβ3	Platelets	Fibrinogen, von Willebrand factor, others	Acute return of blood clots (abrupt closure) and later vessel occlusion (restenosis) after angioplasty has cleaned out a clogged artery
αvβ3	Endothelial cells	Fibronectin, vitronectin, others	Angiogenesis contributing to tumor progression and diabetic retinopathy
	Smooth muscle cells	Same as above	Restenosis
α4β1	Various white blood cells	Fibronectin, VCAM-1	Chronic inflammatory diseases, such as asthma and arthritis
β2 integrins	Neutrophils	ICAMs	Reperfusion injury
	Various white blood cells	Same as above	Chronic inflammatory diseases

vessels, proliferating endothelial cells must form adhesive attachments to one another and to the matrix around them. Increasing evidence implies that the αvβ3 integrin appears in quantity on endothelial cells that are forming new blood vessels and that this display keeps the proliferating cells from dying. For instance, compounds that impair endothelial cell adhesion via the αvβ3 integrin induce improperly proliferating endothelial cells to kill themselves. The agents also inhibit the growth of new blood vessels and produce tumor regression in animals, presumably by cutting off the supply of blood to the tumor.

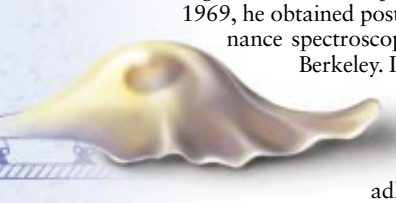
On the basis of such results, researchers are investigating the value of αvβ3 inhibitors as anticancer agents. The same compounds or related ones may also help patients suffering from proliferative retinopathy, a complication of diabetes in which the retina sprouts weak and leaky blood vessels that can destroy the retina and cause blindness. Such drugs should leave healthy blood vessels intact, because the αvβ3 integrin is not displayed abundantly on the endothelial cells of vessels that are no longer growing.

Most proposed therapies for adhesion-related disorders aim to interdict attachment of integrins to their extracellular partners. But agents that increased such attachments could facilitate the migration of healthy skin cells into severely injured skin and so could potentially hasten wound healing. Similar products might be incorporated into artificial matrices, to guide the growth of new, organic tissues that will replace ones damaged by disease, invasive treatments or injury.

Those of us who have studied integrins over the past decade are delighted that therapies based on our research are becoming available. Most of us, however, are just as enthralled by the basic science aspects of the work: the effort to better understand how integrins and other adhesion molecules control so many different aspects of development and physiology. Whether we entered the field by accident or design, we have been treated to an electrifying scientific experience. We each started with a relatively narrow focus and have been informed and dazzled by the breadth of the effects produced by integrins and other adhesion molecules. The rapid progress that has been made testifies to the advances that can be achieved when scientists interested in different research areas put their minds together. SA

The Author

ALAN F. HORWITZ is professor of cell and structural biology at the University of Illinois at Urbana-Champaign. After earning his Ph.D. in biophysics from Stanford University in 1969, he obtained postdoctoral training in magnetic resonance spectroscopy at the University of California, Berkeley. In 1974 he joined the department of biochemistry and biophysics at the University of Pennsylvania School of Medicine, where he became interested in adhesive phenomena. In 1987 Horwitz moved to the University of Illinois to build his current department. In 1995 the siren of adhesion research called him back to the laboratory full time. He now focuses on studying adhesion in cell migration and differentiation.



Further Reading

INTEGRINS: VERSATILITY, MODULATION, AND SIGNALING. R. O. Hynes in *Cell,* Vol. 69, No. 1, pages 11–25; April 3, 1992.

TRAFFIC SIGNALS FOR LYMPHOCYTE RECIRCULATION AND LEUKOCYTE EMIGRATION: THE MULTISTEP PARADIGM. T. A. Springer in *Cell,* Vol. 76, No. 2, pages 301–314; January 28, 1994.

INTEGRINS: EMERGING PARADIGMS OF SIGNAL TRANSDUCTION. M. A. Schwartz, M. D. Schaller and M. H. Ginsberg in *Annual Review of Cell and Developmental Biology,* Vol. 11, pages 549–599; 1995.

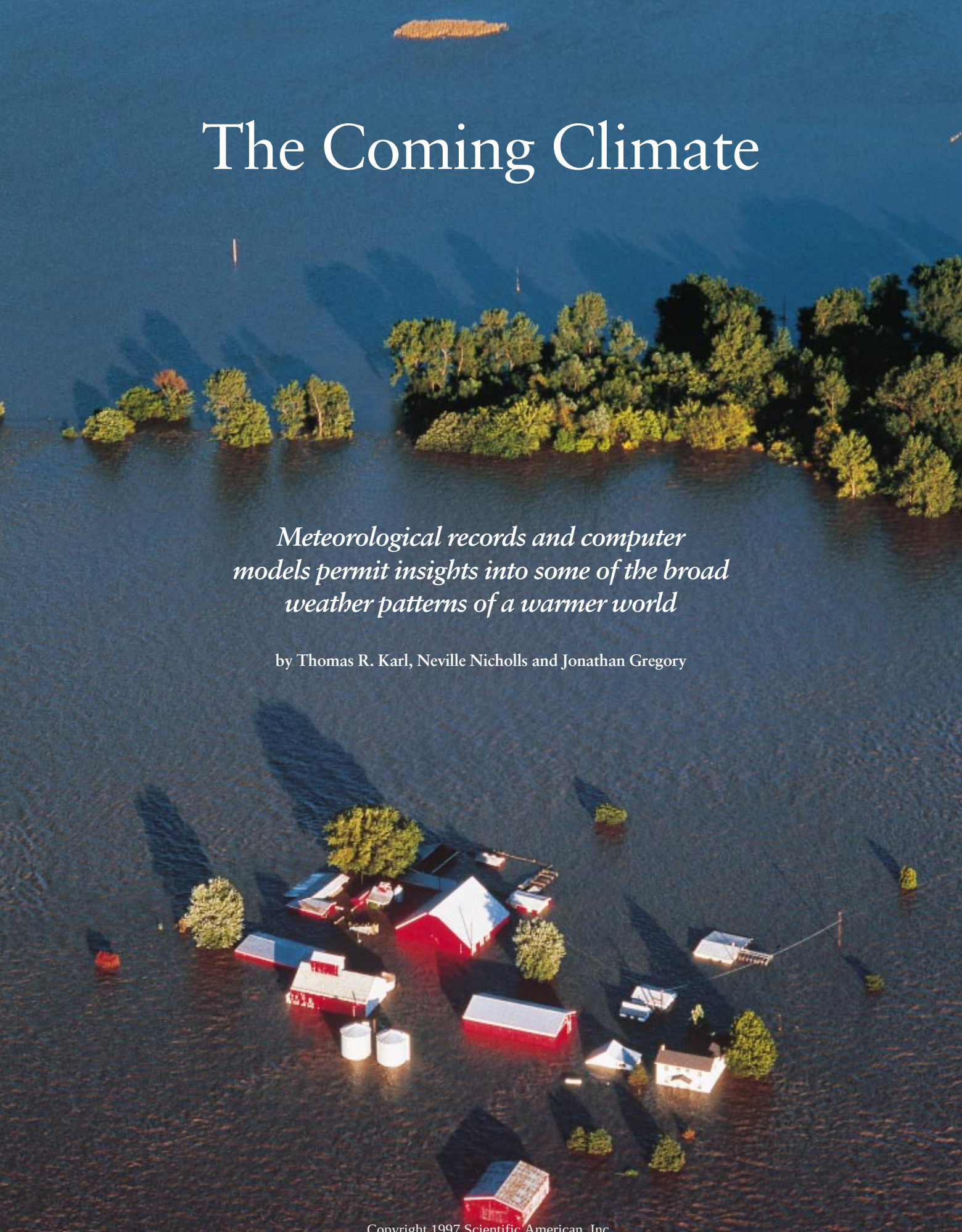
CELL MIGRATION: A PHYSICALLY INTEGRATED MOLECULAR PROCESS. D. A. Lauffenburger and A. F. Horwitz in *Cell,* Vol. 84, No. 3, pages 359–369; February 9, 1996.

THE EXTRACELLULAR MATRIX IN EPITHELIAL BIOLOGY: SHARED MOLECULES AND COMMON THEMES IN DISTANT PHYLA. J. Ashkenas, J. Muschler and M. J. Bissell in *Developmental Biology,* Vol. 180, No. 2, pages 433–444; December 15, 1996.

The Coming Climate

Meteorological records and computer models permit insights into some of the broad weather patterns of a warmer world

by Thomas R. Karl, Neville Nicholls and Jonathan Gregory



Human beings have in recent years discovered that they may have succeeded in achieving a momentous but rather unwanted accomplishment. Because of our numbers and our technology, it now seems likely that we have begun altering the climate of our planet.

Climatologists are confident that over the past century, the global average temperature has increased by about half a degree Celsius. This warming is thought to be at least partly the result of human activity, such as the burning of fossil fuels in electric power plants and automobiles. Moreover, because populations, national economies and the use of technology are all growing, the global average temperature is expected to continue increasing, by an additional 1.0 to 3.5 degrees C by the year 2100.

Such warming is just one of many consequences that climate change can have. Nevertheless, the ways that warming might affect the planet's environment—and, therefore, its life—are among the most compelling issues in earth science. Unfortunately, they are also among the most difficult to predict. The effects will be complex and vary considerably from place to place. Of particular interest are the changes in regional climate and local weather and especially extreme events—record temperatures, heat waves, very heavy rainfall, or drought, for example—which could very well have staggering effects on societies, agriculture and ecosystems.

Based on studies of how the earth's weather has changed over the past century as global temperatures edged upward, as well as on sophisticated computer models of climate, it now seems probable that warming will accompany changes in regional weather. For example, longer and more intense heat waves—a likely consequence of an increase in either the mean temperature or in the variability of daily temperatures—would result in public health threats and even unprecedented levels of mortality, as well as in such costly inconveniences as road buckling and high cooling loads, the latter possibly leading to electrical brownouts or blackouts.

CAMERON DAVIDSON *Tony Stone Images*

FLOODED FARM near the Mississippi River in 1996 illustrates one likely consequence of warming trends. Rainfall will not only increase overall, but individual events will become more intense.

Climate change would also affect the patterns of rainfall and other precipitation, with some areas getting more and others less, changing global patterns and occurrences of droughts and floods. Similarly, increased variability and extremes in precipitation can exacerbate existing problems in water quality and sewage treatment and in erosion and urban storm-water routing, among others. Such possibilities underscore the need to understand the consequences of humankind's effect on global climate.

Two Prongs

Researchers have two main—and complementary—methods of investigating these climate changes. Detailed meteorological records go back about a century, which coincides with the period during which the global average temperature increased by half a degree. By examining these measurements and records, climatologists are beginning to get a picture of how and where extremes of weather and climate have occurred.

It is the relation between these extremes and the overall temperature increase that really interests scientists. This is where another critical research tool—global ocean-atmosphere climate models—comes in. These high-performance computer programs simulate the important processes of the atmosphere and oceans, giving researchers insights into the links between human activities and major weather and climate events.

The combustion of fossil fuels, for example, increases the concentration in the atmosphere of certain greenhouse gases, the fundamental agents of the global warming that may be attributable to humans. These gases, which include carbon dioxide, methane, ozone, halocarbons and nitrous oxide, let in sunlight but tend to insulate the planet against the loss of heat, not unlike the glass of a greenhouse. Thus, a higher concentration means a warmer climate.

Of all the human-caused (anthropogenic) greenhouse gases, carbon dioxide has by far the greatest impact on the global heat budget (calculated as the amount of heat absorbed by the planet less the amount radiated back into space). Contributing to carbon dioxide's greenhouse potency is its persistence: as much as 40 percent of it tends to remain in the atmosphere for centuries. Accumulation of atmospheric carbon dioxide is promoted not only by combustion but also by tropical deforestation.

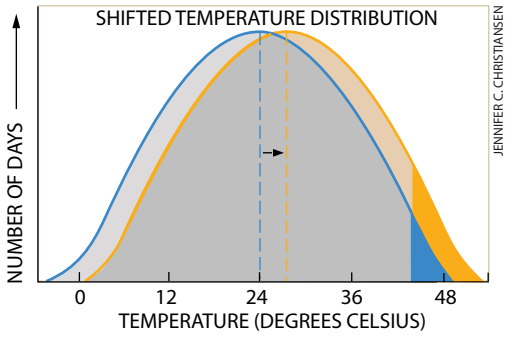
The second most influential human-caused effect on the earth's radiation budget is probably that of aerosols, which are minute solid particles, sometimes covered by a liquid film, finely dispersed in the atmosphere. They, too, are produced by combustion, but they also come from natural sources, primarily volcanoes. By blocking or reflecting light, aerosols tend to mitigate global warming on regional and global scales. In contrast to carbon dioxide, aerosols have short atmospheric residence times (less than a week) and consequently are concentrated near their sources. At present, scientists are less certain about the radiative effects of aerosols than those of greenhouse gases.

By taking increases in greenhouse gases into account, global ocean-atmosphere climate models can provide some general indications of what we might anticipate regarding changes in weather events and extremes. Unfortunately, however, the capabilities of even the fastest computers and our limited understanding of the linkages among various atmospheric, climatic, terrestrial and oceanic phenomena limit our ability to model important details on the scales at which they occur. For example, clouds are of great significance in the atmospheric heat budget, but the physical processes that form clouds and determine their characteristics operate on scales too small to be accounted for directly in global-scale simulations.

How Hot, and How Often?

The deficiencies in computer models become rather apparent in efforts to reproduce or predict the frequency of climate and weather extremes of all kinds. Of these extremes, temperature is one of the most closely studied, because of its effect on humanity, through health and mortality, as well as cooling loads and other factors. Fortunately, researchers have been able to garner some insights about these extremes by analyzing decades of weather data. For statistical reasons, even slight increases in the average temperature can result in big jumps in the number of very warm days [*see top illustration on next page*].

One of the reasons temperature extremes are so difficult to model is that they are particularly sensitive to unusual circulation patterns and air masses, which can occasionally cause them to



SMALL SHIFTS in the most common daily temperature cause disproportionate increases in the number of extremely hot days. The reason is that temperature distributions are roughly Gaussian. So when the highest point in the Gaussian "bell" curve moves to the right (*above*), the result is a relatively large increase (*yellow area*) in the probability of exceeding extremely high temperature thresholds. A greater probability of high temperature increases the likelihood of heat waves (*right*).

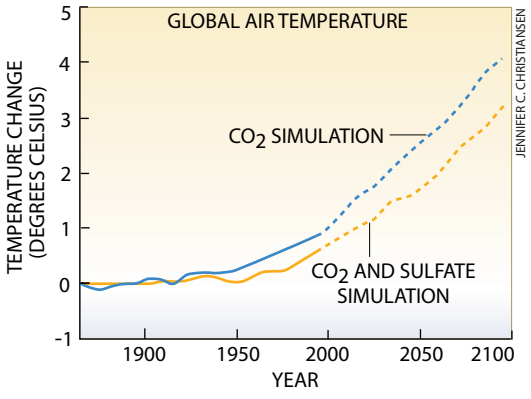
follow a trend in the direction opposite that of the mean temperature. For example, in the former Soviet Union, the annual extreme minimum temperature has increased by a degree and a half, whereas the annual extreme maximum showed no change.

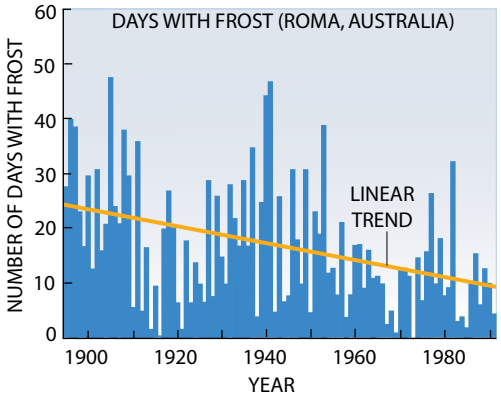
The National Climatic Data Center, which is part of the U.S. National Oceanic and Atmospheric Administration (NOAA), has developed a statistical model that simulates the daily maximum and minimum temperatures from three properties of a plot of temperature against time. These three properties are the mean, its daily variance and its day-to-day correlation (the correlation is an indication of how temperatures persist—for example, how often a hot day is followed by another hot day). Given new values of mean, variance and persistence, the model will project the duration and severity of extremes of temperature.

Some of its predictions are surprising. For example, Chicago exhibits considerable variability of temperature from week to week. Even if the mean January temperature went up by four degrees C (an occurrence that may actually take place late in the next century) while the other two properties remained constant, days with minimum temperatures less than –17.8 degrees C (zero degrees Fahrenheit) would still occur. They might even persist for several days in a row. There should also be a significant reduction in the number of early- and late-season freezes. And, not surprisingly, during the summer, uncomfortably hot spells, including so-called killer heat waves, would become more frequent. With just a three degree C increase in the average July temperature, the probability that the heat index (a measure that includes humidity and measures overall discomfort) will exceed 49 degrees C (120 degrees F) sometime during the month increases from one in 20 to one in four.

Because of their effects on agriculture, increases in the minimum are quite significant. Observations over land areas during the latter half of this century indicate that the minimum temperature has increased at a rate more than 50 percent greater than that of the maximum. This increase has lengthened the frost-free season in many parts of the U.S.; in the Northeast, for example, the frost-free season now begins an average of 11 days earlier than it did during the 1950s. A longer frost-free season can be beneficial for many crops grown in places where frost is not very common, but it also affects the growth and development of perennial plants and pests.

The reasons minimum temperatures are going up so much more rapidly than maximums remain somewhat elusive. One possible explanation revolves around cloud cover and evaporative cooling, which have increased in many areas. Clouds tend to keep the days





GLOBAL AIR TEMPERATURE rise was simulated (*above, left*) by a climate model at the U.K. Meteorological Office's Hadley Center. The blue line is from a simulation based on carbon dioxide only; the yellow line also takes into account sulfate. As the global temperature has increased, the number of days with minimums below zero degrees Celsius has gone down. This example (*left*) shows the annual number of days with frost in Roma, Queensland, in Australia.

WARREN MARR *Panoramic Images*

cooler by reflecting sunlight and the nights warmer by inhibiting loss of heat from the surface. Greater amounts of moisture in the soil from additional precipitation and cloudiness inhibit daytime temperature increases because part of the solar energy goes into evaporating this moisture. More conclusive answers, as well as a prediction about whether the asymmetry in daytime and nighttime warming will continue, await better computer models.

Projections of the day-to-day changes in temperature are less certain than those of the mean, but observations have suggested that this variability in much of the Northern Hemisphere's midlatitudes has decreased as the climate has become warmer. Some computer models also project decreases in variability. The variability depends on season and location and is also tied to surface characteristics, such as snow on the ground or moisture in the soil. In midlatitudes, changes in the daily variability of temperature have also been linked to changes in the frequency and intensity of storms and in the location of the paths commonly taken by storms. These storm tracks are, in effect, a succession of eastward-moving midlatitude depressions whose passage dominates the weather.

The relation between these storms and temperature is complex. In a warmer world, the difference of temperature between the tropics and the poles would most likely cover a smaller range, because greater warming is expected near the poles. This factor would tend to weaken storms. On the other hand, high in the atmosphere this difference would be reversed, having the opposite influence. Changes in storms could also happen if anthropogenic aerosols continue to cool the surface regionally, altering the horizontal temperature contrasts that control the location of the storm tracks.

More Precipitation

The relation between storms and temperature patterns is one of the reasons it is so difficult to simulate climate changes. The major aspects of climate—temperature, precipitation and storms—are so interrelated that it is impossible to understand one independently of the others. In the global climate system, for example, the familiar cycle of evaporation and precipitation transfers not only water from one place to another but also heat. The heat used at the surface by evaporation of the water is released high in the atmosphere when the water condenses again into clouds and precipitation, warming the surrounding air. The atmosphere then loses this heat by radiating it out into space.

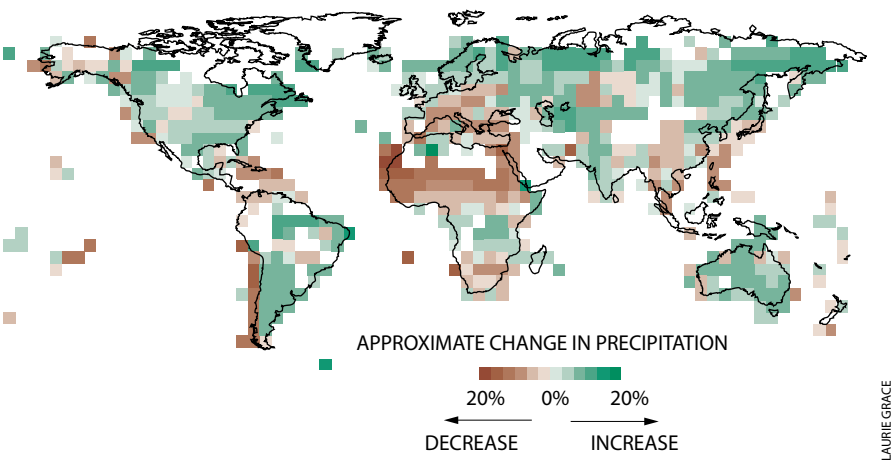
With or without additional greenhouse gases, the earth takes in the same amount of solar energy and radiates the same amount back out into space. With a greater concentration of greenhouse gases, however, the surface is better insulated and can radiate less heat *directly* from the ground to space. The efficiency with which the planet radiates heat to space goes down, which means that the temperature must go up in order for the same amount of heat to be radiated. And as the temperature increases, more evaporation takes place, leading to more precipitation, averaged across the globe.

Precipitation will not increase everywhere and throughout the year, however. (In contrast, all areas of the globe should have warmer temperatures by the end of the next century.) The distribution of precipitation is determined not only by local processes but also by the rates of evaporation and the atmospheric circulations that transport moisture.

For instance, most models predict reduced precipitation in southern Europe in summer as a result of increased greenhouse gases. A significant part of the rainfall in this region comes from local evaporation, with the water not precipitated locally being exported to other areas. Thus, in a warmer climate, increased evaporation in the spring would dry out the soil and lead to less water being available for evaporation and rainfall in the summer.

On a larger scale, most models predict an increase in average precipitation in winter at high latitudes because of greater poleward transport of moisture derived from increased evaporation at low latitudes. Since the turn of the century, precipitation has indeed increased in the high latitudes of the Northern Hemisphere, primarily during the cold season, as temperatures have increased. But for tropical and subtropical land areas, precipitation has actually decreased over the past few decades. This is especially apparent over the Sahel and eastward to Indonesia.

In northernmost North America (north of 55 degrees) and Eurasia, where conditions are normally far below freezing for much of the year, the amount of



LAURIE GRACE

PRECIPITATION TRENDS between 1900 and 1994 reveal a general tendency toward more precipitation at higher latitudes and less precipitation at lower ones. Green indicates more rain; brown less.

snowfall has increased over the past several decades. Further increases in snowfall are likely in these areas. Farther south, in southern Canada and the northern U.S., the ratio of snow to rain has decreased, but because of the increase in total precipitation there has been little overall change in the amount of snowfall. In the snow transition belts, where snow is intermittent throughout the cold season, the average snowfall will tend to diminish as the climate warms, before vanishing altogether in some places. Interestingly, areal snow cover during spring and summer abruptly diminished by nearly 10 percent after 1986. This decrease in snow cover has contributed to the rise of spring temperatures in the middle and high latitudes.

Besides the overall amounts of precipitation, scientists are particularly interested in the frequency of heavy downpours or rapid accumulations because of the major practical implications. Intense precipitation can result in flooding, soil erosion and even loss of life. What change do we expect in this frequency?

Whether precipitation occurs is largely determined by the relative humidity, which is the ratio of the concentration of water vapor to its maximum saturation value. When the relative humidity reaches 100 percent, water condenses into clouds, making precipitation possible. Computer models suggest that the distribution of relative humidity will not change much as the climate changes.

The concentration of water vapor needed to reach saturation in the air rises rapidly with temperature, however, at about 6 percent per degree Celsius. So in a warmer climate, the frequency of precipitation (which is related to how often the relative humidity reaches 100 percent) will change less than the amount of precipitation (related to how much water vapor there is in the air). In addition, not only will a warmer world be likely to have more precipitation, but the average precipitation event is likely to be heavier.

Various analyses already support the notion of increased intensity. In the U.S., for example, an average of about 10 percent of the total annual precipitation that falls does so during very heavy downpours in which at least 50 millimeters falls in a single day. This proportion was less than 8 percent at the beginning of this century.

As incredible as it may seem with all this precipitation, the soil in North America, southern Europe and in several other places is actually expected to become drier in the coming decades. Dry soil is of particular concern because of its far-reaching effects, for instance, on crop yields, groundwater resources, lake and river ecosystems and even on down to the foundations of buildings. Higher temperatures dry the soil by boosting the rates of evaporation and transpiration through plants. Several models now project significant increases in the severity of drought. Tempering these predictions, however, are studies of drought frequency and intensity during this century, which suggest that at least during the early stages of global warming other factors have overwhelmed the drying effects of warmer weather. For example, in the U.S. and the former U.S.S.R., increases in cloud cover during the past several decades have led to reduced evaporation. In western Russia, in fact, soil moisture has increased.

WARREN E. FAIDLEY *Weatherstock*

HURRICANES, a kind of tropical cyclone, will probably occur in different global patterns as a result of warming; their overall incidence, however, may not change. Hurricane Andrew, shown raging here in Miami, struck the southeast coastal U.S. in 1992, causing $30-billion worth of damage.

Stormy Weather

Great as they are, the costs of droughts and heat waves are less obvious than those of another kind of weather extreme: tropical cyclones. These storms, known as hurricanes in the Atlantic and as typhoons in the western North Pacific, can do enormous damage to coastal areas and tropical islands. As the climate warms, scientists anticipate changes in tropical cyclone activity that would vary by region. Not all the consequences would be negative; in some rather arid regions the contribution of tropical cyclones to rainfall is crucial. In northwest Australia, for example, 20 to 50 percent of the annual rainfall is associated with tropical cyclones. Yet the damage done by a single powerful cyclone can be truly spectacular. In August 1992 Hurricane Andrew killed 54 people, left 250,000 homeless and caused $30-billion worth of damage in the Caribbean and in the southeast coastal U.S.

Early discussions of the possible impacts of an enhanced greenhouse effect often suggested more frequent and more intense tropical cyclones. Because these storms depend on a warm surface with unlimited moisture supply, they form only over oceans with a surface temperature of at least 26 degrees C. Therefore, the reasoning goes, global warming will lead to increased ocean temperatures and, presumably, more tropical cyclones.

Yet recent work with climate models and historical data suggests that this scenario is overly simplistic. Other factors, such as atmospheric buoyancy, instabilities in the wind flow, and the differences in wind speed at various heights (vertical wind shear), also play a role in the storms' development. Beyond enabling this rather broad insight, though, climate models have proved of limited use in predicting changes in cyclone activity. Part of the problem is that the simulations are not yet detailed enough to model the very intense inner core of a cyclone.

The historical data are only slightly more useful because they, too, are im-

perfect. It has been impossible to establish a reliable global record of variability of tropical cyclones through the 20th century because of changes in observing systems (such as the introduction of satellites in the late 1960s) and population changes in tropical areas. Nevertheless, there are good records of cyclone activity in the North Atlantic, where weather aircraft have reconnoitered since the 1940s. Christopher W. Landsea of the NOAA Atlantic Oceanographic and Meteorological Laboratory has documented a decrease in the intensity of hurricanes, and the total number of hurricanes has also followed suit. The years 1991 through 1994 were extremely quiet in terms of the frequency of storms, hurricanes and strong hurricanes; even the unusually intense 1995 season was not enough to reverse this downward trend. It should be noted, too, that the number of typhoons in the northwestern Pacific appears to have gone up.

Overall, it seems unlikely that tropical cyclones will increase significantly on a global scale. In some regions, activity may escalate; in others, it will lessen. And these changes will take place against a backdrop of large, natural variations from year to year and decade to decade.

Midlatitude cyclones accompanied by heavy rainfall, known as extratropical storms, generally extend over a larger area than tropical cyclones and so are more readily modeled. A few studies have been done. A recent one by Ruth Carnell and her colleagues at the Hadley Center of the U.K. Meteorological Office found fewer but more intense storms in the North Atlantic under enhanced greenhouse conditions. But the models do not all agree.

Analyses of historical data also do not give a clear conclusion. Some studies suggest that since the late 1980s, North Atlantic winter storm activity has been more extreme than it ever was in the previous century. Over the past few decades, there has also been a trend toward increasing winds and wave heights in the northern half of the North Atlantic Ocean. Other analyses by Hans von Storch and his colleagues at the Max Planck Institute for Meteorology in Hamburg, Germany, found no evidence of changes in storm numbers in the North Sea. In general, as with the tropical cyclones, the available information suggests that there is little cause to anticipate global increases in extratropical storms but that regional changes cannot be ruled out.

The Future

Although these kinds of gaps mean that our understanding of the climate system is incomplete, the balance of evidence suggests that human activities have already had a discernible influence on global climate. In the future, to reduce the uncertainty regarding anthropogenic climate change, especially on the small scales, it will be necessary to improve our computer modeling capabilities, while continuing to make detailed climatic observations.

New initiatives, such as the Global Climate Observing System, and detailed studies of various important climatic processes will help, as will increasingly powerful supercomputers. But the climate system is complex, and the chance always remains that surprises will come about. North Atlantic currents could suddenly change, for example, causing fairly rapid climate change in Europe and eastern North America.

Among the factors affecting our predictions of anthropogenic climate change, and one of our greatest uncertainties, is the amount of future global emissions of greenhouse gases, aerosols and other relevant agents. Determining these emissions is much more than a task for scientists: it is a matter of choice for humankind. SA

The Authors

THOMAS R. KARL, NEVILLE NICHOLLS and JONATHAN GREGORY were all members of the Intergovernmental Panel on Climate Change, which assessed and reported on the human impact on global climate. Karl is a senior scientist at the National Oceanic and Atmospheric Administration's National Climatic Data Center. Nicholls is a senior principal research scientist at the Australian Bureau of Meteorology Research Center. Gregory is a climate modeler at the Hadley Center of the U.K. Meteorological Office.

Further Reading

CHANGING BY DEGREES: STEPS TO REDUCE GREENHOUSE GASES. U.S. Congress, Office of Technology Assessment, 1991.

POLICY IMPLEMENTATION OF GREENHOUSE WARMING: MITIGATION, ADAPTATION, AND THE SCIENCE BASE. National Academy of Sciences. National Academy Press, 1992.

GLOBAL WARMING DEBATE. Special issue of *Research and Exploration:* A Scholarly Publication of the National Geographic Society, Vol. 9, No. 2; Spring 1993.

GLOBAL WARMING: THE COMPLETE BRIEFING. John T. Houghton. Lion Press, 1994.

CLIMATE CHANGE 1995: THE SCIENCE OF CLIMATE CHANGE. Contribution of Working Group I to the *Second Assessment Report of the Intergovernmental Panel on Climate Change.* Edited by J. T. Houghton, L. G. Meira Filho, B. A. Callendar and N. Harris. Cambridge University Press, 1996.

INDICES OF CLIMATE CHANGE FOR THE UNITED STATES. T. R. Karl, R. W. Knight, D. R. Easterling and R. G. Quayle in *Bulletin of the American Meteorological Society,* Vol. 77, No. 2, pages 279–292; February 1996.

New Chemical Tools to Create Plastics

Small molecular machines called metallocene catalysts have revolutionized the industrial synthesis of valuable plastics

by John A. Ewen

Alchemists of medieval times tried to convert lead into gold. Today's chemists know that quest is a lost cause, but many of us still earn our living discovering how to convert inexpensive raw materials into valuable products. The modern equivalent of the alchemist's magical potion is a catalyst—a substance that initiates a desirable chemical reaction or speeds up a reaction that would otherwise be too slow to be economical. The impact on society of this latter-day magic can hardly be overestimated.

Catalysts have made possible—among other things—the development of several modern plastics. Plastics are synthesized by persuading small organic molecules to join together in long chains known as polymers. These materials have become so widely used it is hard to imagine life without them. Light, waterproof and resistant to corrosion, they are the designer's first choice for such disparate items as water pipes, trash bags, hair combs, fibers for clothing and road construction, and packaging for food and medicine. The billions of pounds of plastic sold every year have made the industry a vital sector of the economy.

Despite their ubiquity, some important synthetic polymers have, since the 1950s, been made using catalysts that offer only limited control over the range of lengths of the polymer chains, as well as over other structural features that influence the properties of the bulk material. Chemists have long realized that improved catalysts could provide better control of polymerization, making it possible to create plastics with physical properties finely tuned to particular uses. Plastic for a gearwheel, for example, might be made especially durable, and material for piping might be designed to withstand high temperatures. But without a clear understanding of the traditional catalysts, researchers could make only slow improvements by trial and error.

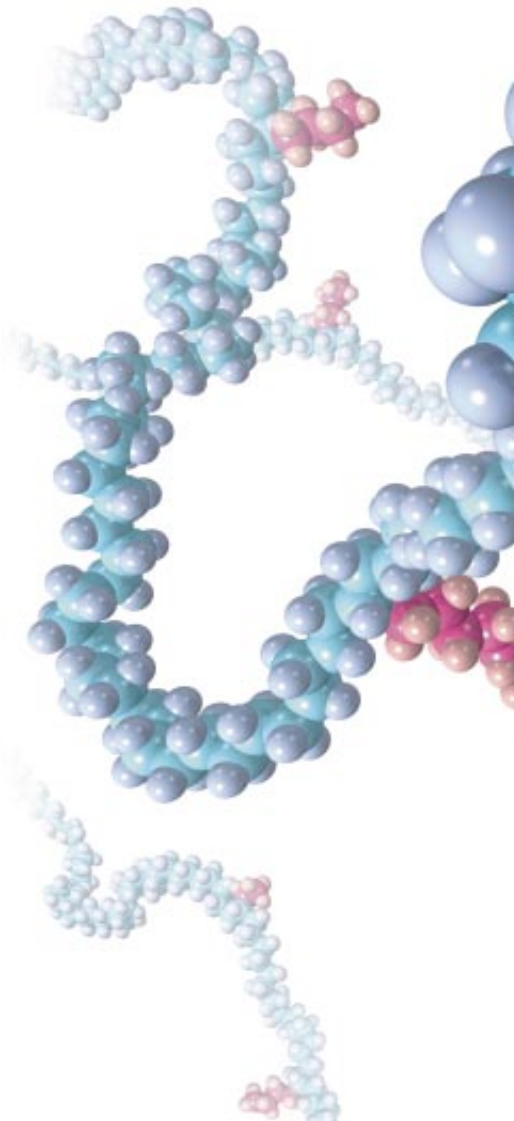
In the past two decades major discoveries have shown how to create catalysts that provide superb control over polymer growth. The plastics industry is already producing hundreds of millions of pounds of material using these catalysts, called metallocenes. Business analysts predict that within a few years, plastics synthesized with metallocenes will be found in most homes.

Old-Time Catalysts

The traditional catalysts got their start back in the mid-1950s, when Karl Ziegler and Giulio Natta in Europe invented processes for polymerizing ethylene and propylene to make polyethylene and polypropylene. Ziegler and Natta shared a Nobel Prize for Chemistry in 1963 for their work, and these simple plastics are now mainstays of the industry. Ethylene and propylene are small molecules of carbon and hydrogen produced in large quantities by the petroleum industry; ethylene has two carbon atoms and propylene three. Both have a reactive double bond connecting two adjacent carbon atoms. The Ziegler-Natta catalysts are mixtures of solid and liquid compounds containing metal atoms that attack the double bond, converting it into a single one. In doing so, the catalysts allow the affected carbons to bond to neighboring ethylene or propylene molecules.

The result is a rapidly lengthening polymer. Polyethylene consists of a zigzag carbon backbone studded with hydrogen atoms. Polypropylene differs in having attached to its backbone regularly spaced groups consisting of a car-

POLYETHYLENE



bon atom and three hydrogen atoms. Polypropylene is stiffer than polyethylene and has more diverse uses.

One commercial impetus for improving the process for making polyethylene arose from the need to have better control over side chains. Pure long-chain (or linear) polyethylene can be brittle. The polymer can better withstand tearing forces, such as those exerted on a garbage bag, if the chains incorporate an occasional short side chain. Side chains result in the binding together of microcrystals of the plastic. But traditional Ziegler-Natta catalysts are not efficient at incorporating side chains in the desired manner, so manufacturers employing the old technology have difficulty making highly tear-resistant polyethylene.

The challenge with polypropylene was that, unlike ethylene, propylene can be added to a chain in two different mirror-image orientations. An analogy might be the different ways to stack books. We could make sure all the spines were on the same side of the stack, or we could alternate spine-left with spine-right. The stability of the stack would be rather different in each case. Likewise, the properties of polypropylene vary depending on how the monomers—the individual molecules forming the polymer—are inserted. This complicates the use of traditional Ziegler-Natta catalysts for making polypropylene.

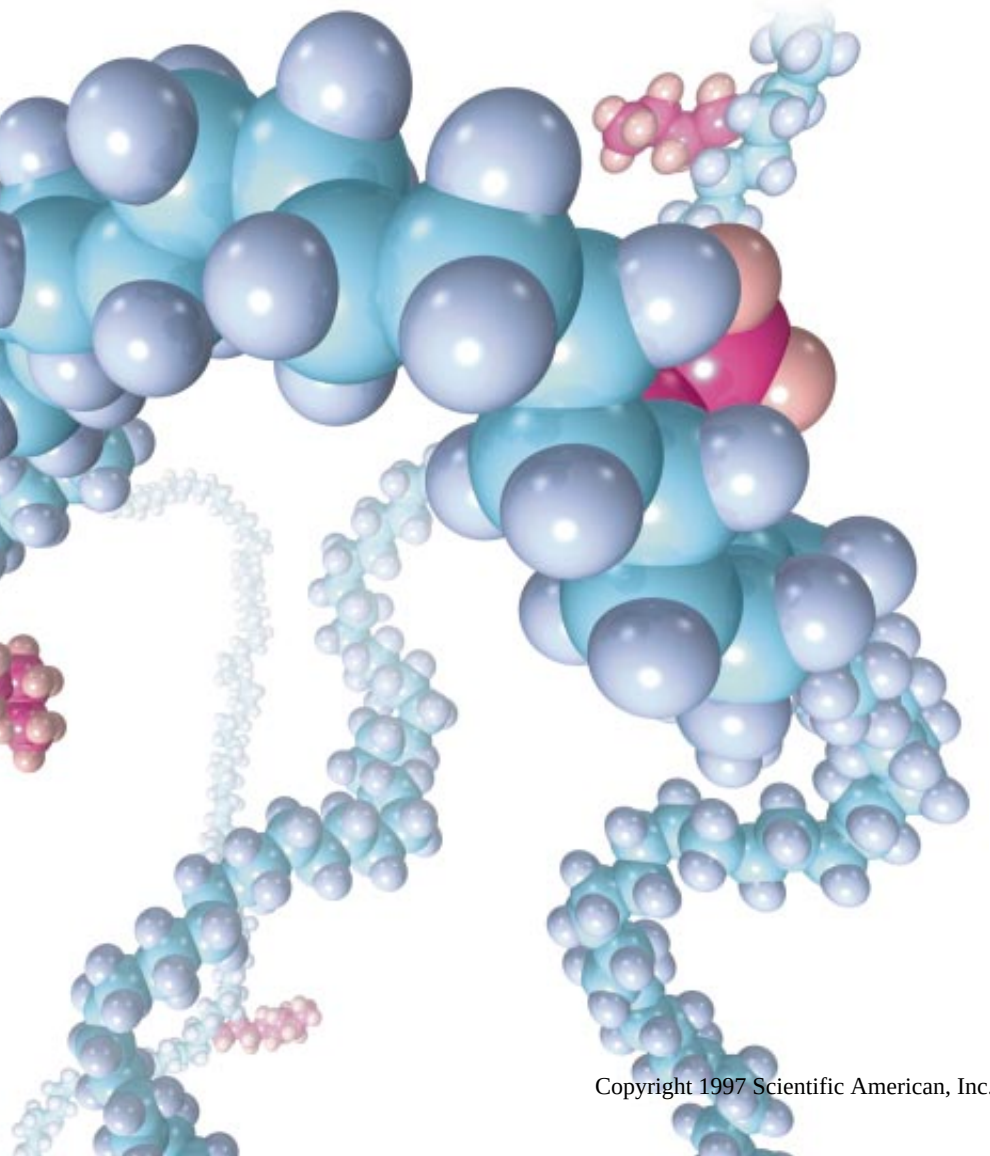
Molecular Machines

My involvement in this story goes back to 1980, when I was hired by Exxon as a catalyst expert in a long-range research group based near Houston. Catalysts work like machines, and it is essential to understand how they function before we can think about designing better ones. Because the formula for making Ziegler-Natta catalysts had been modified through trial and error, nobody knew what structures surrounded the metal atoms. Yet the structure of a catalyst dictates what it does and thus predetermines the properties of the polymers it will make. My first task was therefore to learn something about the mechanisms at work in the traditional catalysts.

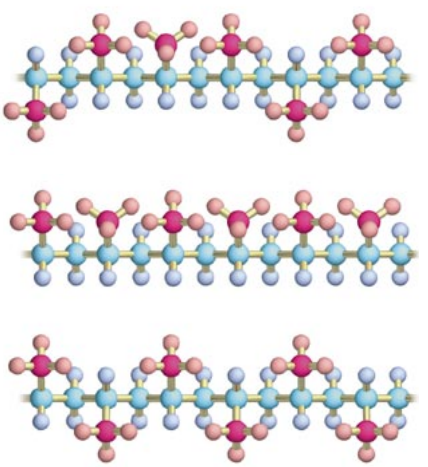
Although chemists find it hard to manipulate very large molecules, they can exercise fine control in the synthesis of small ones. The structure and symmetries designed into minute quantities of a catalyst are reflected and amplified millions of times over in the structure of the polymers they produce. This is essentially the principle that nature uses when enzymes synthesize the complex molecules found in living organisms. My idea was to construct small-molecule catalysts having well-defined, stable structures and a single catalytic site. We could then determine their mode of action by studying the polymers they made.

A major influence on my thinking was Piet Cossee's theories about Ziegler-Natta catalysts. In the 1960s Cossee, at the Royal Shell Laboratory in Amsterdam, proposed that a metal atom forms temporary bonds simultaneously with a growing polymer chain and with the double bond in the monomer. These bonds allow the chain end to be electrically attracted toward the monomer. Cossee theorized that the chain end migrates and fuses with the monomer. The chain thus becomes a little longer, and its old position becomes vacant. As the process repeats, the chain extends.

I chose to investigate metallocenes,



FORMS OF POLYPROPYLENE



BIOGRAFX

KEY MODERN PLASTICS that can be made with metallocenes include polyethylene (*left*) and polypropylene (*three chains above*). Random placement of carbon side chains (*pink*) gives a weak structure (*top*), but isotactic and syndiotactic polypropylene (*middle* and *bottom*) are valuable.

FINA OIL AND CHEMICAL COMPANY

DOW CHEMICAL COMPANY

SYNDIOTACTIC polypropylene (*left and center blocks in left photograph*) has many advantages over conventional polypropylene (*right block*), including increased transparency. Polyethylene (*right photograph*) synthesized by a variant of a metallocene catalyst is in commercial production; it is more resistant to tearing than polymer made with a conventional catalyst.

molecules first synthesized in 1953 by John Birmingham, who was working in the laboratory of the British Nobel laureate Geoffrey Wilkinson. Like traditional Ziegler-Natta catalysts, metallocenes contain a metal but just a single atom, usually titanium or zirconium. The metal atom is linked to two rings of five carbon atoms and to two other groups, often single carbon atoms with attached hydrogens. The rings play a key role: they partly enclose the metal atom like a half-open clamshell. Electrons associated with the rings influence the metal, modifying its propensity to attack carbon-carbon double bonds.

Chemists knew in 1957 from the work of Natta and others, including David Breslow, Wendell P. Long and Norman Newburg of Hercules Powder Company, that the titanium metallocene catalyzed the polymerization of ethylene. The reaction was, however, too slow to be of commercial interest. But in 1976 Hansberg Sinn and Walter Kaminsky of the University of Hamburg made a significant observation: adding controlled amounts of water mysteriously made metallocenes catalyze far more quickly.

I recalled that back in 1967, F. Dyachovskii, A. K. Shilova and A. E. Shilov of the U.S.S.R. Academy of Sciences had written a paper in English in which they subscribed to the theory of the Hercules workers that the metal atom in a metallocene has a positive charge and works basically like the metal in a traditional Ziegler-Natta catalyst. Sinn and Kaminsky's results therefore suggested to me that modified metallocenes might have commercial potential.

Another reason for investigating metallocenes was that the two five-carbon rings in the basic metallocene molecule have between them 10 hydrogen atoms. It seemed likely that replacing some of these hydrogens with other atoms or groups would alter the metallocenes' properties. So there was a world of potential to be exploited, and such modifications had, surprisingly, never been studied systematically.

From Theory to Practice

We soon confirmed that milligram quantities of a zirconium metallocene in a high-pressure reactor could yield hundreds of grams of plastic in an hour. But to make practical catalysts, further improvements would be necessary. The trick, I suspected, would be to vary the amount of crowding around the five-carbon rings surrounding the metal atom: a lot of chemistry is a matter of how much room there is for things to happen.

With a reluctant and skeptical assistant in tow, I first experimented with polyethylene. We started simply, preparing and testing metallocenes in which we replaced one hydrogen atom at a time with a simple hydrocarbon group. Eventually a pattern emerged. If we added a group that donated negative charge to the rings, the catalyst would become more reactive toward ethylene and, by working faster, would be able to make longer polymers. Too many groups, though, crowded the rings and made the catalyst less reactive, so chains were shorter. These observations gave us previously unavailable direct control over the properties of the polymers. Nowadays metallocenes can make polyethylene chains containing just a few or hundreds of thousands of carbon atoms.

At one point we tested metallocenes in which we had substituted single-carbon groups for all 10 of the hydrogen atoms on the two rings. This catalyst, which is much more crowded than the parent molecule, exhibited a significant property. In mixtures of ethylene and propylene, it was 250 times more reactive to ethylene than it was to the bulkier propylene. That bias toward the smaller monomer is much larger than that of the parent metallocene. So pronounced was the preference that this catalyst could produce long-chain crystalline polyethylene even in mixtures containing large amounts of propylene or of bigger molecules. This discovery had a practical consequence. It meant that by mixing together crowded and more open catalysts, industrial chemists could perform several different, well-controlled reactions at once in a single vessel. They could, for example, synthesize in one step blends of crystalline polyethylene and more flexible polymers containing side chains incorporated by the uncrowded catalyst.

We next attempted a more radical step. We made a bridge with a silicon atom joining the two five-carbon rings on one side. This modification, we figured, should partly peel the rings back, as though pulling open the clamshell. The effect was dramatic: the resulting metallocene was far more reactive toward propylene and larger molecules than the

parent metallocene was. Catalysts with this structural feature are now employed in industry to incorporate larger molecules into a polymer as side chains.

Although it was by now clear that metallocenes had valuable properties, a severe technical problem limited their use. Unless the catalysts were attached to a supporting matrix, they tended to produce a polymer dust that agglomerated and stuck to everything. Serendipity solved this problem, but only several years later: in 1990 James C. Stevens of Dow Chemical tested a titanium catalyst that was based on a metallocene but had a nitrogen atom replacing one of the five-carbon rings. This catalyst rapidly made polyethylene at the relatively high temperatures that Dow was using for its plastic manufacturing. (Dow uses temperatures of more than 120 degrees Celsius, whereas 50 to 70 degrees C is more typical.) At these high temperatures, the polymer in the Dow process is soluble, and so the dust problem was avoided.

Dow quickly adopted the catalyst. Another pleasant surprise was that the very open structure of this compound enabled it to polymerize so avidly that it could link chains with side chains to make polymers with a controlled amount of interlinking of chains. Such materials are flexible, making them especially suitable for insulation around electrical cables.

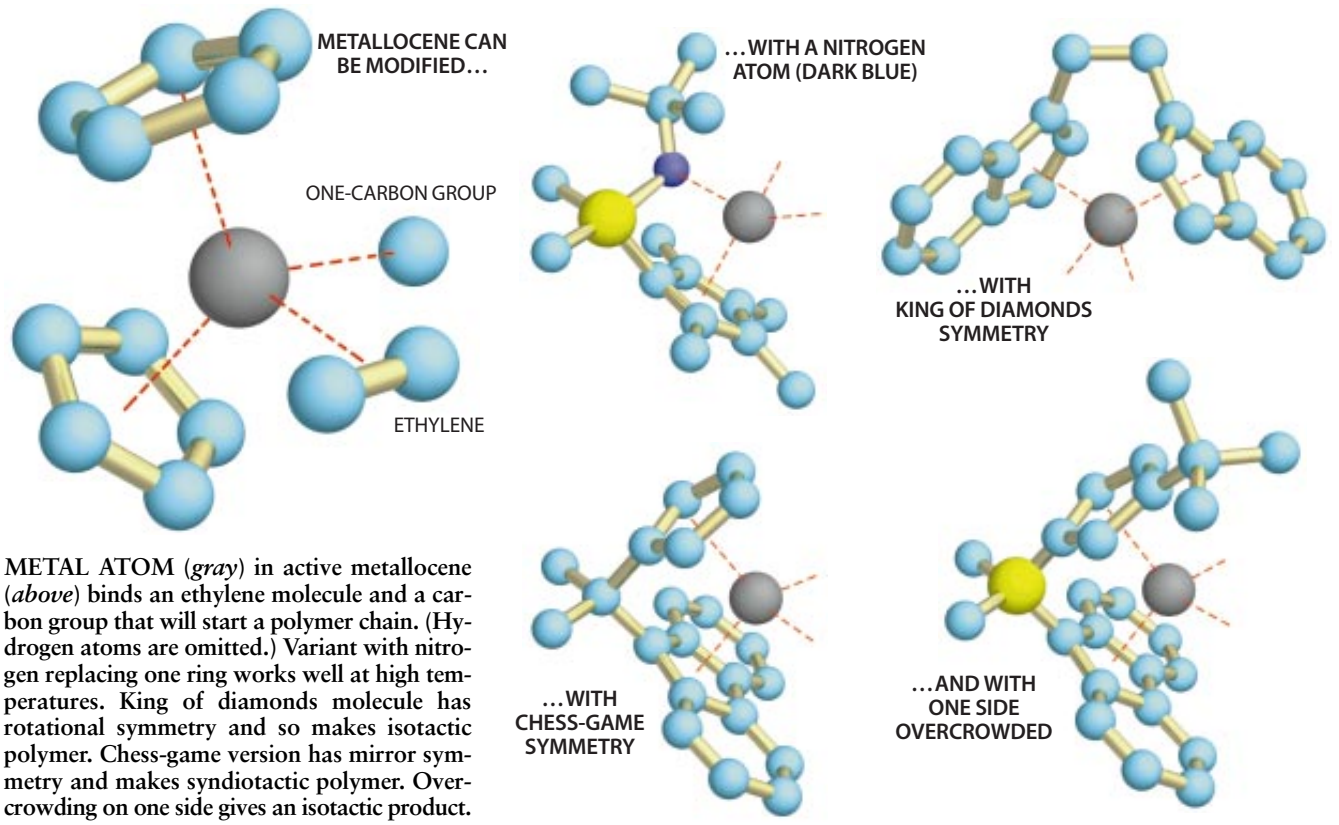
Metallocenes of different types are now being used in a variety of commercial processes for making polyethylene. Their products are more uniform and so have strikingly better properties than plastics made with the traditional Ziegler-Natta catalysts. Manufacturers are selling polymer mixtures with optimum chain lengths and side-chain content for specific uses.

Fearful Symmetry

In 1982 I moved on to investigate how to adapt metallocenes to the manufacture of polypropylene, which presented a more difficult challenge because of the two mirror-image ways to add a monomer. Polypropylene in which the great majority of monomers have the same orientation as their neighbors is called isotactic. Isotactic polypropylene chains fit together better than polypropylene chains consisting of random monomers, making the polymer harder and more crystalline.

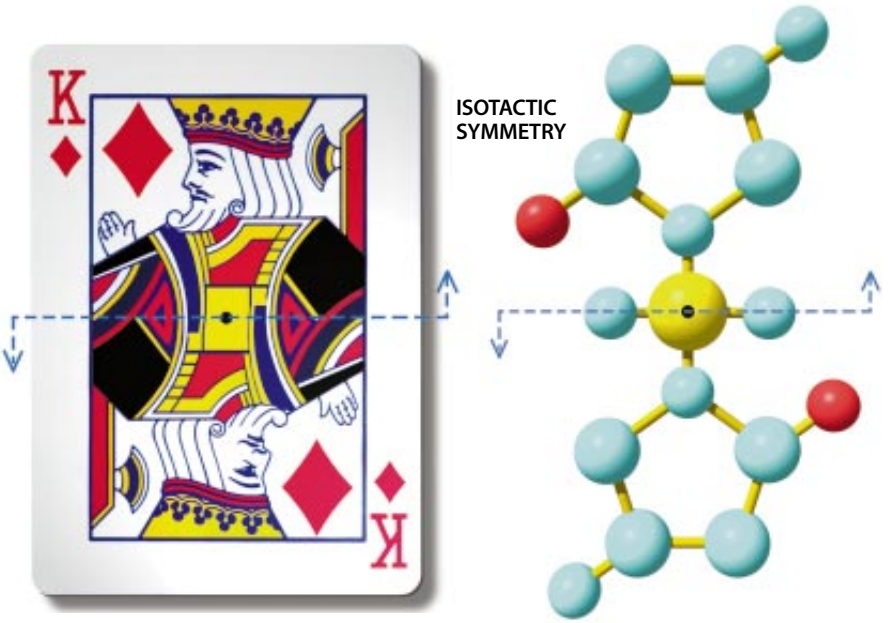
Producing high-quality isotactic polypropylene boils down to ensuring that 99 percent of the time the monomer fits into the active site of a catalyst in only one orientation. I knew from the work of Adolpho Zambelli of the University of Salerno that there were two ways to achieve this control with Ziegler-Natta catalysts. One, called chain-end control, occurs when some catalysts are at –30 to –80 degrees C. The low temperature slows the rate at which the growing chain moves in relation to the catalyst, and blocks of identically oriented monomers can result. But errors are frequent, and when I tried to apply this technique using a metallocene, the blocks were too short. The resulting polymer was a less than useful, stretchy elastic substance with a low melting point.

The second approach exploits the symmetry of the catalyst itself, so a monomer can fit only in the desired orientation, like a key in a lock. As I turned my efforts to this method, I was most concerned about something called the back-skip reaction, which Cossee had proposed in his theory on traditional Ziegler-Natta catalysts. Those catalysts can produce isotactic polypropylene. Cossee reasoned that this must mean that after migrating and fusing to a monomer, a chain skips back to the position where it was attached before the addition. Without such a back-skip, Cossee intuitively recognized, the chain's geometry would have alternated the orientations of the monomers.



BIOGRAFX

METAL ATOM (*gray*) in active metallocene (*above*) binds an ethylene molecule and a carbon group that will start a polymer chain. (Hydrogen atoms are omitted.) Variant with nitrogen replacing one ring works well at high temperatures. King of diamonds molecule has rotational symmetry and so makes isotactic polymer. Chess-game version has mirror symmetry and makes syndiotactic polymer. Overcrowding on one side gives an isotactic product.



METALLOCENE CATALYST variants are shown as if flattened, with the central metal atom removed. The molecule at the left has the rotational symmetry of a playing card. Looking the same from left or right, it always adds monomers to a chain in the same orientation, creating isotactic polypropylene. The molecule at the right, in con-

I was not sure whether back-skipping would occur in metallocenes. Therefore, in order to make a isotactic polymer, I needed a catalyst shaped to ensure that the orientation of added monomers would be identical from the viewpoint of either bond, just as a face-up king of diamonds card looks identical to players on both sides of a table. The orientation of each monomer would then be the same whether back-skipping occurred or not. But there was another complication. The five-carbon rings of metallocenes rapidly rotate and change places. I needed a bridge between the two rings to make the structure more rigid.

An organometallic chemist, Hans Brintzinger of Konstanz University in Germany, had synthesized titanium metallocenes that met my requirements. The rings were bridged with a two-carbon linkage, and organic groups attached to the rings imposed constraints that were identical viewed from either attachment position. This was a molecule with the symmetry of a playing card. The resulting polymer should be isotactic whether or not Cossee's back-skip reactions occurred.

The chances of success seemed minuscule, but the gamble worked. I can still recall the joy of seeing the first few grams of crystalline, isotactic polypropylene formed from a metallocene swirling around in a Pyrex vessel of toluene solvent. Paolo Corradini's group in Milan then calculated that the modified five-carbon rings jointly force the growing chain away from them, which pushes the propylene into place to form the isotactic polymer. I made another catalyst that was identical except that it lacked the playing-card symmetry. As expected, it produced material with randomly oriented monomers, because the two rings could no longer jointly control the orientation of monomers.

The isotactic polypropylene made with the Brintzinger compound still had too many erroneous insertions, however, to compete with commercial Ziegler-Natta material. I tested a slightly more crowded version, which improved the product somewhat but still had problems: the titanium catalysts I was using did not work for very long at practical temperatures.

At this point, Exxon displayed no interest in pursuing metallocenes for making polypropylene, and I left the company in 1984. Exxon permitted me to publish my work in an article later that year. The article attracted a lot of scientific interest, and in 1985 I joined Fina Oil and Chemical.

Five years later, following the advances I describe below, metallocenes had become the hottest area in catalyst chemistry. I and others were by then very close to commercializing polypropylene manufacture with metallocenes. Exxon sued me, claiming to own substantially all my work on metallocenes. The lawsuit and my counterclaim for libel lasted five years. In 1995 a jury fully vindicated me, and an Exxon scientist was found guilty of anonymously mailing copies of Exxon's claims to other scientists around the world. Many of these scientists helped me at the trial. I will always be grateful for their support.

Mirror Molecules

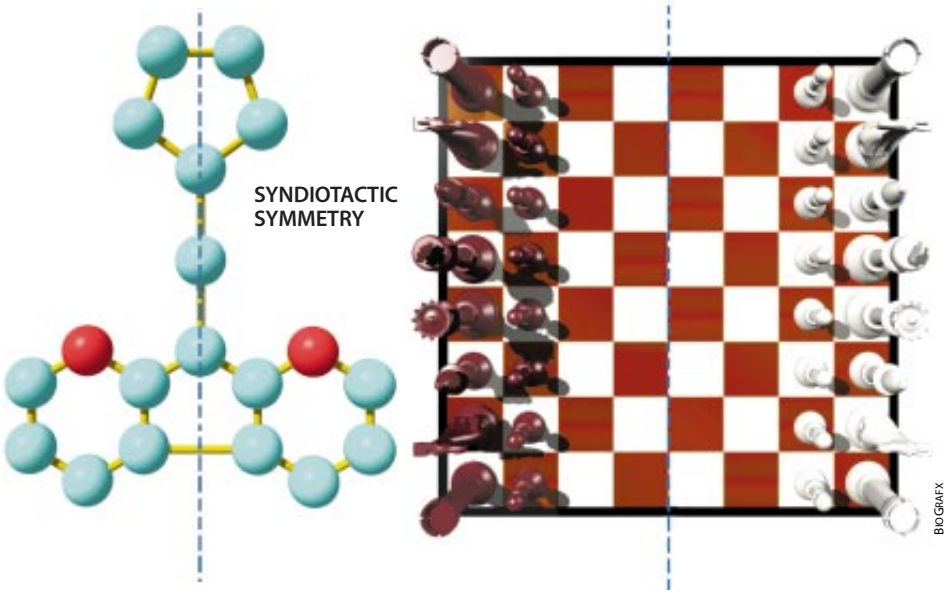
At Fina, my small metallocene research group included Michael J. Elder and Larry Jones, who did nearly all the even more exciting work I recount next. We joined in the worldwide effort to improve the performance of metallocenes in the commercial manufacture of polypropylene.

In 1985 Kaminsky and Brintzinger, hoping to produce better isotactic material, tested the zirconium version of Brintzinger's titanium metallocene that I had studied earlier. Although the zirconium compound worked quickly, the resulting polymers were still only one tenth of the length needed: they formed a brittle wax of low melting point. I substituted the rare metal hafnium for zirconium, but it failed to increase the melting point significantly.

We finally succeeded in increasing the polymer's melting point by using yet more rigid catalysts in which the two five-carbon rings were bridged by a silicon atom. Adding large groups of atoms to the side of the rings facing away from the silicon bridge further improved performance by reducing the space for erroneous insertions. Commercial metallocenes all make use of these innovations.

In 1987 I put to a simultaneous acid test, as one might say, both Cossee's chain-migration/back-skip theory and the Dyachovskii-Shilova-Shilov extension of it to the positively charged metal in metallocenes. If the chain-migration reaction was indeed occurring in metallocenes, it should have been possible to exploit the effect to make polypropylene that had alternating orientations of the monomer. This form is known as syndiotactic, and it has a host of valuable properties. Compared with the isotactic form, it has better impact strength, is more transparent and is unusually resistant to degradation by the gamma rays used to sterilize some food and medical equipment.

At the time, this polymer could be made only in tiny amounts in laboratories. My idea was that we could make it in quantity if the two alternating bonding positions for propylene imposed constraints that were mirror images of each other. This is the symmetry of chess



trast, has the mirror-image symmetry of a chess game. As a growing polymer migrates from side to side, this catalyst adds monomers in alternating orientations, creating syndiotactic polymer. The atoms colored red have been discovered to play an important role in preventing premature termination of polymer chains.

pieces at the start of a game. For white, the queen is to the left of the king. For black, the opposite is true.

My hunch was a long shot, but I selected a compound that had two six-carbon rings built onto each side of one of the five-carbon rings. A bridge connected the two five-carbon rings. The success of this experiment provided the first proof that both Dyachovskii and Cossee were correct: the charged metal atom was bringing about a chain migration that allowed this metallocene to produce syndiotactic material. Some veteran chemists at first refused to believe it. But again, the end of the chain, together with the constraints of the site's chess-game symmetry, was dictating the orientation of the monomer. As the chain migrated from one side of the catalyst to the other with each addition, new monomers inserted themselves in alternating orientations.

This catalyst was also the first zirconium-based metallocene catalyst to make any kind of polypropylene with chains as long as those of commercial polymers. The key, it turned out, was that the substituents on the modified five-carbon ring included two carbon atoms in critical positions that prevented premature termination of the chain. The practical problem of chain length, which had bedeviled zirconium metallocenes, had thus been solved inadvertently. And syndiotactic polypropylene has now become the preferred plastic for various critical applications, especially in the medical field, where it commands a high price and is used for blood bags and sterile garments.

One additional experiment provided more evidence for Cossee's theory. When the catalyst that made syndiotactic polypropylene was starved of propylene, the melting point of the product decreased. This turned out to be because some insertions were omitted under these conditions; the chain had back-skipped to its previous position before a new monomer could be added, so some monomers were incorporated in the same orientation as their neighbors. The imperfection had disrupted the crystal structure.

Researchers have since made some informative structural variations of this valuable catalyst. We designed one modified version so that when the chain was in one of its two possible positions it would give isotactic insertions, but in the alternate position the catalyst was effectively symmetric and so gave random insertions. The result was the elegant molecule hemi-isotactic polypropylene.

In another variation, the crowding on one side of the catalyst was more severe, so that when the chain was in that position it was forced to back-skip without adding a monomer. The predictable result was isotactic polypropylene. Thirty years after Cossee daringly proposed a back-skip mechanism to explain how the puzzling Ziegler-Natta compounds could produce isotactic polymer, we had developed a well-understood catalyst that relied on back-skipping to produce isotactic material with the same chain length and melting point.

Exxon had hired me to unravel the mysteries of Ziegler-Natta catalysts. I found no mysteries. Most of the questions had already been answered by Zambelli, Cossee and Dyachovskii and his colleagues. In demonstrating their foresight, I gratefully used their theories to design new catalysts. These catalysts then obediently made polymers that rejuvenated a 40-year-old industry. SA

The Author

JOHN A. EWEN earned a Ph.D. in chemistry at Tulane University in 1979. He was a catalyst chemist for Exxon Chemical Company in Baytown, Tex., from 1980 to 1984. The following year he joined Fina Oil and Chemical in Deer Park, Tex., to research production of polypropylene and the chemistry of zirconium-based catalysts. Since 1991 Ewen has been president of Catalyst Research Corporation in Houston, which advises companies on commercial aspects of catalysis.

Further Reading

LIGAND EFFECTS ON METALLOCENE CATALYZED ZIEGLER-NATTA POLYMERIZATIONS. J. A. Ewen in *Catalytic Polymerization of Olefins*. Edited by Tominaga Keii and Kazuo Soga. Proceedings of the International Symposium on Future Aspects of Olefin Polymerization, 1985. Elsevier Science, 1986.

CHEMISTRY OF CATIONIC DICYCLOPENTADIENYL GROUP 4 METAL-ALKYL COMPLEXES. Richard F. Jordan in *Advances in Organometallic Chemistry,* Vol. 32, pages 325–387; 1991.

METALLOCENE/PROPYLENE STRUCTURAL RELATIONSHIPS: IMPLICATIONS ON POLYMERIZATION AND STEREOCHEMICAL CONTROL MECHANISMS. J. A. Ewen, M. J. Elder, R. L. Jones, L. Haspeslagh, J. L. Atwood, S. G. Bott and K. Robinson in *Die Makromolekulare Chemie: Macromolecular Symposia,* Vols. 48–49, pages 253–295; August 1991.

STEREOSPECIFIC OLEFIN POLYMERIZATION WITH CHIRAL METALLOCENE CATALYSTS. H. H. Brintzinger, D. Fischer, R. Mulhaupt, B. Rieger and R. M. Waymouth in *Angewandte Chemie International Edition in English,* Vol. 34, No. 11, pages 1143–1170; June 16, 1995.

Galaxies in the Young Universe

By comparing distant primeval galaxies with older ones nearby, astronomers hope to determine how galaxies form and evolve

by F. Duccio Macchetto and Mark Dickinson

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION; ROBERT WILLIAMS AND THE HUBBLE DEEP FIELD TEAM *Space Telescope Science Institute*

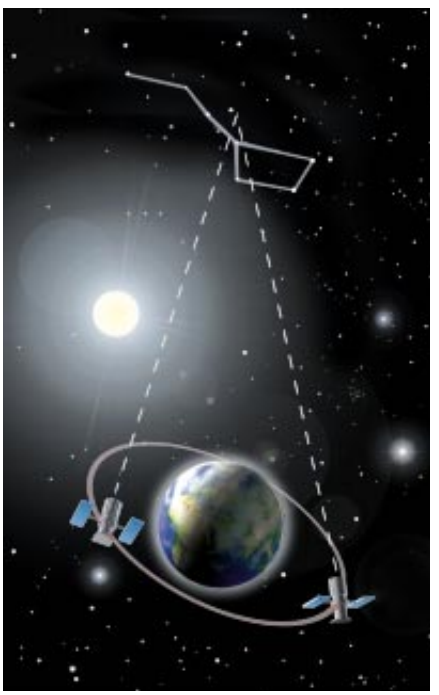
Our conception of the universe has changed radically during this past century, as the powers of astronomical observation have steadily improved. Edwin Hubble's pioneering work in the 1920s led to the idea that a "big bang" gave birth to a universe that has been expanding ever since. Later studies revealed that the universe changed in other ways over time as well. Initially it was filled with exceedingly hot, dense, nearly uniform material. Now it is relatively empty. As astronomers look across millions of light-years, the matter they see is collected into a sparse handful of seemingly isolated galaxies. How this transformation occurred, and why the galaxies formed as they did, remains a central question in cosmology today.

Considerable effort has gone into examining galaxies nearby—those that are the product of some 10 billion years of evolution. In recent years, however, astronomers have made enormous progress in studying galaxies at cosmological distances—namely, ones that existed when the universe was young. Thanks to the travel time imposed by the finite speed of light, scientists can peer directly into the past by looking out to greater distances. At this time, the light cosmologists can see left the most distant galaxies when the universe was less than one fifth its present age. With new instruments and techniques, the hope is to view distant "primeval" galaxies in the process of forming and to trace their evolution to the present day.

Of course, such views do not come easily. Young galaxies are so far away that they appear small and faint to even the most powerful telescopes. A galaxy the size of our own Milky Way, observed as it was when the universe was half its current age, would span an angle of mere arc seconds on the sky. Seen through earthbound telescopes, such a galaxy would lose most of its distinguishing structural features; even under excellent conditions, cosmologists can often do little more than distinguish remote galaxies from faint stars. In recent years, however, the Hubble Space Telescope—which has extremely sharp vision because it is positioned above the earth's atmosphere—has provided detailed images of the distant universe.

Many Hubble research programs are dedicated to studying young galaxies, but one in particular has taken center stage. In December 1995, Hubble was trained on an unremarkable patch of sky, $^1/_{140}$ the apparent size of the full moon, near the Big Dipper. The spot was chosen simply because it afforded a clear view out of our own galaxy and an efficient place to park the telescope. Over the course of 10 full days, Hubble took hundreds of exposures through four filters, covering the spectrum from near ultraviolet to near-infrared radiation. These images, known as the Hubble Deep Field (HDF) observations, have given us the best view of the distant universe. (In astronomical parlance, "deep" refers both to faintness and distance.)

The HDF images reveal about 3,000 faint galaxies, which take on a bewildering variety of shapes and col-



BRYAN CHRISTIE

HUBBLE DEEP FIELD (HDF) observations (*mosaic at left*) offer the clearest view of the distant universe to date. To make these images, researchers trained the orbiting telescope on a patch of sky, $^1/_{140}$ the apparent size of the full moon, near the Big Dipper for 10 days in December 1995 (*diagram*). Astronomers hope that by examining faint primeval galaxies and older ones, they can learn how galaxies mature.

ors. Many are more than a billion times fainter than what can be seen with the naked eye. The challenge astronomers face is interpreting these two-dimensional pictures of a four-dimensional universe. After all, everything along the line of sight—near and far, young and old—is projected onto the same plane of the sky. The goal is not only to identify primeval galaxies among this mix but to compare their characteristics with those of older galaxies nearby and at all intermediate distances. In doing so, we hope to determine how galaxies form and evolve.

Judging by size or brightness alone, it is not easy to tell how old any one galaxy actually is. Faint objects in our neighborhood and intrinsically bright ones farther away can look very similar. But we do have other ways to determine age. For instance, because the universe is continuously expanding, the distance to a galaxy is proportional to the velocity of its motion away from us. This recession causes a Doppler shift in the light a galaxy emits. As a consequence,

DETAILED VIEWS from HDF observations show countless galaxies. After all, everything along the line of sight—both near and far, young and old—is projected onto the same plane of the sky. In all, the HDF images reveal some 3,000 faint galaxies, which assume an array of unusual shapes and colors. Many are more than a billion times fainter than what can be seen with the naked eye.

its characteristic spectral features are displaced toward longer, redder wavelengths. By measuring this so-called redshift, denoted by z, we can determine a galaxy's relative distance and youth.

The timeline of cosmic history is ordered according to redshift: larger values of z represent earlier epochs when the universe was smaller, younger and more dense. Since any particular redshift z, the universe has expanded by a factor of $(1 + z)$. The relation between z and age is more complex. But generally speaking, at a redshift of z, the universe was at most $1/(1+z)$ times its present age. Thus, at a redshift of 1, the universe was at most half as old as it is today; at a redshift of 3, it was less than 25 percent as old and perhaps as little as 12.5 percent as old.

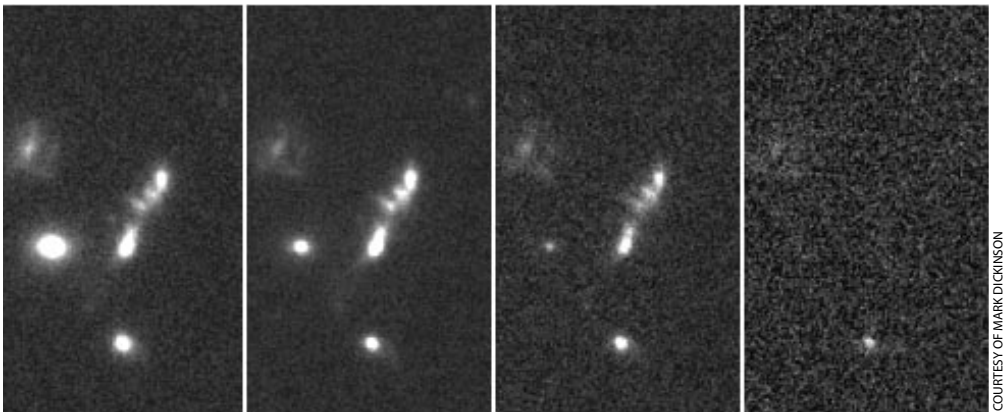
Extensive redshift measurements have now determined distances for thousands of faint galaxies out to $z = 1$, the practical limit of current surveys. The Hubble images let us classify the galaxies, compare them with ones nearby and evaluate their evolutionary state. Many of these galaxies seem to have had a relatively quiescent past: the Hubble pictures reveal bright spiral and elliptical objects, not unlike those nearby, out to redshifts of at least 1. Even at relatively early times, these normal-looking galaxies seem to have existed in numbers comparable to the total found in the universe today. We imagine, then, that many galaxies have remained largely the same for billions of years.

Many others, though, have undergone dramatic changes, according to the redshift surveys and the Hubble images. Take the simple exercise of counting galaxies in the sky. There are simply far too many faint ones. At the limits of today's observations, there are at least 10 times as many galaxies as there are in the local universe. The blue colors of these galaxies, and the strong emission features seen in their spectra, suggest that in comparison to galaxies today, they formed stars quite rapidly—an activity that made them brighter and thus more easily seen in the surveys.

Moreover, many of these galaxies have irregular, convoluted morphologies, suggesting that galaxy interactions and perhaps even mergers were common long ago. Although irregular and interacting galaxies can be found closer to home, they seem to have been far more prevalent when the universe was young. At the extreme limits of the HDF, the galaxy count is dominated by a vast number of extremely compact objects, barely resolved even by Hubble's sharp eye. Based on these data, astronomers have concluded that the overall star formation rate in the universe has declined dramatically during the latter half of the universe's history and that most action occurred in irregular galaxies.

Cosmologists do not yet understand what physical mechanisms drive this evolution or where these galaxies have gone. Perhaps the rate of galaxy interactions used to be higher only because the universe was smaller and galaxies were closer together. Maybe frequent interactions triggered the star formation researchers now see. Or maybe early galaxies exhausted their gas supplies, stopped forming stars and faded away to near invisibility. Whatever the case, the findings also tell scientists that the formation of "ordinary" spiral and elliptical galaxies is apparently still out of reach of most redshift surveys; they are found in abundance all the way out to $z = 1$. So to complete this history, astronomers must push the search further and also probe the nearby universe for remnants of the apparently disappearing "faint blue galaxy" population.

It is not yet possible to ascertain directly distances for the vast majority of galaxies found at the limits of the HDF observations. They do not provide enough light for even the largest telescopes to measure their redshifts. Thus, astronomers employ other techniques to search for galaxies beyond $z = 1$. One method is to rely on distant objects, such as radio sources, as markers or beacons. Certain galaxies generate powerful emissions at radio wavelengths. This radio emission is presumed to originate from an active core, or nucleus, within the galaxy, such as a hidden quasar. Radio galaxies are not common today, but their signature emissions can attract attention from far off in the universe. Indeed, some of the brightest radio sources in the sky are located at vast distances and have redshifts rivaling those of the most remote quasars.



COURTESY OF MARK DICKINSON

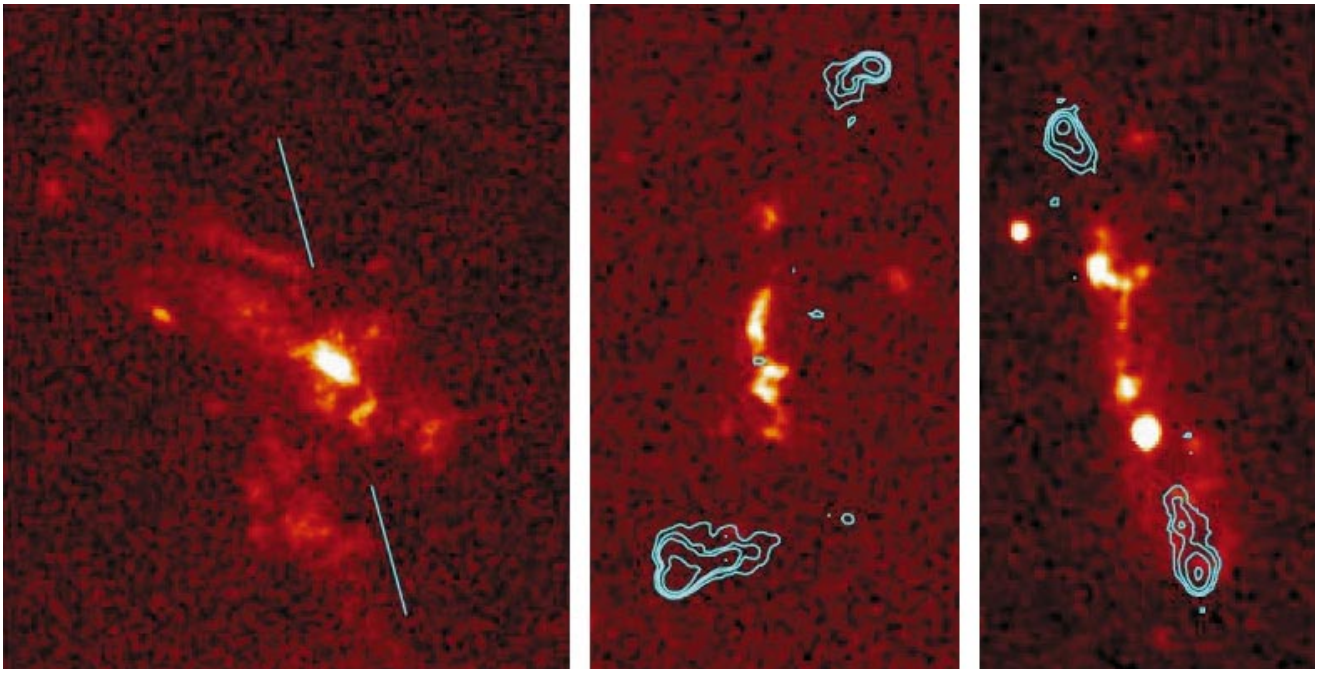
ULTRAVIOLET DROPOUT GALAXIES are common in the deepest astronomical images. Viewed through several filters, these distant galaxies are visible at blue, red and green wavelengths but disappear in near-ultraviolet images (*above*). Hydrogen, which is ubiquitous in galaxies and the spaces between them, absorbs all UV light that is bluer than a certain wavelength.

Relying on Radio Galaxies

Powerful radio sources originate most often from elliptical galaxies, which are now generally thought to be quite old. The hope, therefore, has been that the distant galaxies in which radio sources are generated seeded today's elliptical galaxies. On closer inspection, though, distant radio galaxies exhibit highly unusual morphological and spectral traits. New Hubble images display these peculiarities, which include bizarre and complex forms. It seems that a strong radio source can alter a galaxy's appearance and perhaps its evolution as well. Some radio galaxies may well be true primeval galaxies. But at present, because they seem so abnormal, their properties are difficult to interpret. Thus, their pedigree as progenitors of normal modern-day galaxies is suspect.

Fortunately, galaxies are gregarious, and where one is found, others are often lurking. Radio galaxies sometimes

NASA; ROBERT WILLIAMS AND THE HUBBLE FIELD TEAM *STScI*



NASA; NATIONAL RADIO ASTRONOMY OBSERVATORY; M. LONGAIR *University of Cambridge*

RADIO GALAXIES 3C265 (*left*), 3C324 (*center*) and 3C368 (*right*) exhibit unique morphological and spectral traits. For these composite images, the Hubble Space Telescope captured the galaxies (*red emissions*), and the Very Large Array Radio Interferometer recorded the radio maps (*blue lines*). Powerful radio sources can very likely alter a galaxy's appearance and perhaps its evolution. Some radio galaxies may well be true primeval galaxies, but at present they are difficult to interpret.

inhabit clusters, in which they are surrounded by many other faint, more ordinary galaxies. Using powerful new observational tools and techniques, we have sought out and studied these prosaic companions. Such efforts have located rich collections of galaxies around radio sources as distant as $z = 2.3$, when the universe was at most 30 percent of its present age. A few of these distant clusters have been studied in detail using the Hubble, such mighty ground-based telescopes as the Keck 10-meter telescope in Hawaii, and such orbiting x-ray telescopes as ROSAT.

One cluster around the radio source 3C324, at $z = 1.2$, shares many common features with rich clusters nearby. It contains hot gas that shines brightly at x-ray wavelengths. This grouping reveals that some young galaxy clusters were extremely massive—a strong challenge to some theories of cosmic structure formation. Moreover, the cluster around 3C324 contains galaxies remarkably similar to the giant ellipticals that populate closer clusters: they have very red colors and simple, spheroidal forms. Such features indicate that the stars in these distant cluster galaxies were already mature when they emitted the light we observe. Evidently they must have formed much earlier, at some higher redshift, where cosmologists must now extend their search for the stars' birth.

The properties of these galaxies fuel the debate on the age of the universe itself. Recent efforts to determine the rate of cosmic expansion, called the Hubble Constant, have suggested that the universe may be younger than was previously thought. Some observations imply that the universe is perhaps less than 10 billion years old, yet astronomers find stars in our own Milky Way they believe to be older than 10 billion years—an impossible contradiction if both the Hubble Constant data and the stellar ages are correct! If cosmologists are to believe the elliptical galaxies near 3C324 were already old at a redshift of 1.2, the problems become more pronounced.

Questing for Quasars

In the search for primeval galaxies, some astronomers have turned to quasars, the brightest objects in the universe, as beacons. As the light from a distant quasar travels through space to reach the earth, it encounters clouds of gas that imprint characteristic features on the spectrum in the form of absorption lines. Most of these spectral lines are quite weak and are probably produced by tenuous gases unrelated to normal galaxies. Occasionally, however, the lines are broad and deep and totally absorb the quasar radiation at that wavelength. The inferred mass and size of these absorbers suggest they are parts of the disks or halos around galaxies. These strong absorption features can be found easily at redshifts of 3 and beyond, suggesting that there must have been galaxies present in the young universe to produce them.

According to theory, young galaxies should form out of clouds of hydrogen, in which many hot blue stars are constantly being born. As generations of these stars cycle through their short lives, the process of nuclear fusion transforms hydrogen into the heavier elements, which astronomers generally lump together under the name "metals." These early stars then explode as supernovae, ejecting the metals into the surrounding gas clouds. In doing so, they also shock and compress the gas clouds, triggering the birth of new generations of stars. Based on this model, the characteristic spectrum for a protogalaxy would be one dominated by blue starlight, with traces of metals. In addition, it might show a strong Lyman-alpha emission line, a feature produced by the plentiful hydrogen that the hot blue stars heat.

Searching for Lyman-alpha emission from young galaxies at high redshift has become a cottage industry, but few

good examples have been found. This may be because Lyman-alpha radiation can be easily reabsorbed, especially when dust is present. The mechanisms that produce metals from the hot stars also produce dust, and so young protogalaxies might quench their own Lyman-alpha emission. Without the characteristic Lyman-alpha emission line, it can be quite difficult to recognize young galaxies or measure their redshifts. Nevertheless, astronomers have had some success using special filters tuned to detect hydrogen Lyman-alpha emission.

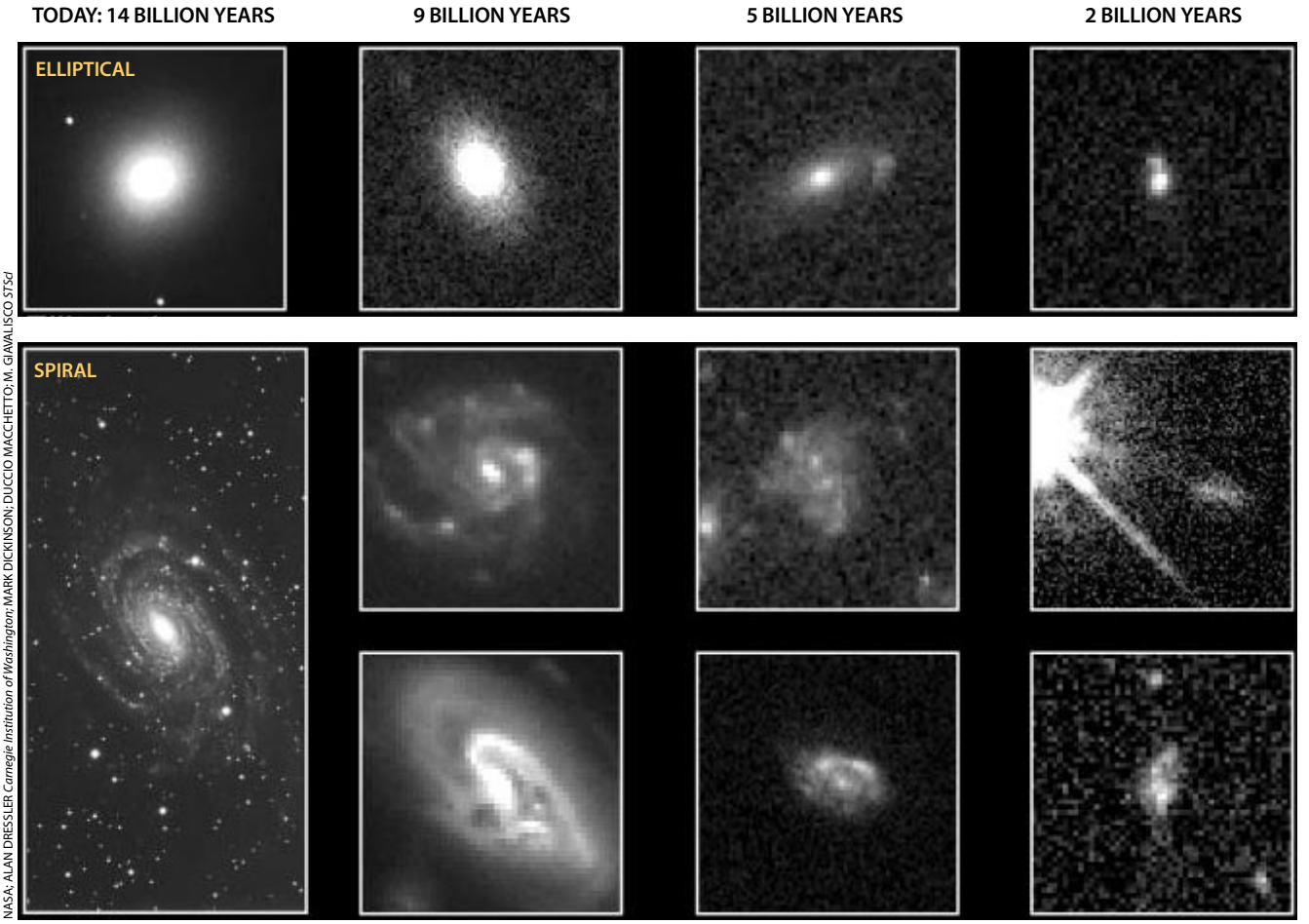
In the field around the distant quasar 0000-263, so named for its coordinates, this technique yielded an exciting find, dubbed with the rather uninspiring name "G2." It was one of the first apparently normal galaxies found at a redshift greater than 3. Spectroscopic measurements subsequently confirmed its distance, and further observations identified several other galaxies in the field at similarly large redshifts; one of these is probably responsible for the Lyman-alpha absorption in the quasar spectrum. Deep Hubble images offered the first clear portrait of galaxies when the universe was somewhere between 10 and 25 percent of its present age. G2 itself appears spheroidal, like a younger blue counterpart to the ellipticals found in the 3C324 cluster discussed earlier. Galaxies such as G2 may well be the progenitors of today's elliptical galaxies.

A Cosmic Baby Boom

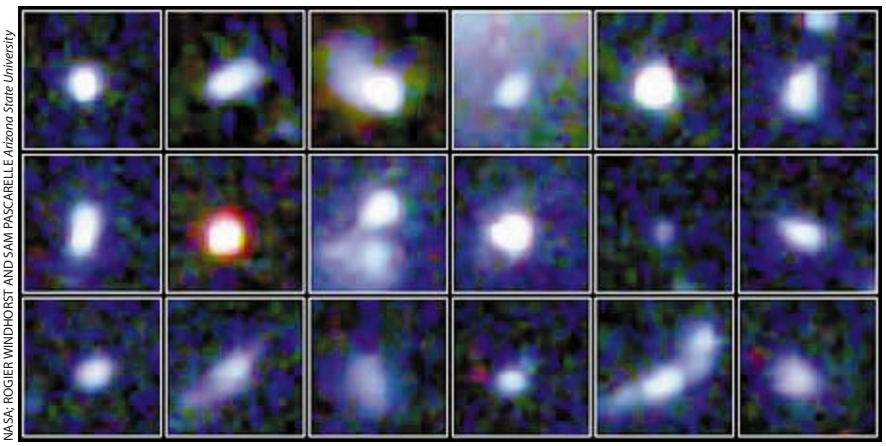
Most recently, scientists have uncovered a treasure trove of galaxies in the early universe, thanks to new search strategies. The most effective method does not capitalize on Lyman-alpha emission but instead takes advantage of a particular color signature that all distant galaxies share. The cause is again hydrogen gas, but the mechanism is different. Hydrogen, which is ubiquitous both within galaxies and in intergalactic space, strongly absorbs all ultraviolet light that is more blue than a certain wavelength. The effect on the starlight a young galaxy emits is dramatic: there is a sharp "cutoff" in its spectrum, producing an unmistakable color signature. Viewed through several filters, a distant galaxy should be visible at red or green wavelengths but disappear in the bluest images.

New work shows that these "ultraviolet dropout" galaxies are remarkably common in the deepest astronomical images: they appear everywhere in the sky, and systematic surveys have identified hundreds of them in just the past two

GALAXIES EVOLVED as the universe aged. Today (*left column below*) most galaxies assume either elliptical (*top*) or spiral (*bottom*) forms. HST images from earlier cosmic epochs, when the universe was at most a third its current age, reveal objects resembling mature ellipticals. Their presence suggests that such galaxies formed rapidly after the big bang. Spiral galaxies, in contrast (*middle and bottom rows*), tend to appear increasingly less regular as one looks back in time, perhaps suggesting that their formation and evolution was more gradual.



NASA; ALAN DRESSLER *Carnegie Institution of Washington*; MARK DICKINSON; DUCCIO MACCHETTO; M. GIAVALISCO *STScI*



NASA; ROGIER WINDHORST AND SAM PASCARELLE *Arizona State University*

FAINT BLUE SUBGALACTIC CLUMPS, some 11 billion light-years from the earth, may well be the stuff from which today's galaxies were made. The Hubble telescope found these 18 objects in the northern part of the constellation Hercules, near Draco. Astronomers believe such small clumps collided and merged with one another to generate giant modern-day galaxies.

years. Their blue colors again suggest that they are forming stars rapidly. Very few of the galaxies, though, exhibit strong Lyman-alpha lines, supporting the notion that these emissions can be easily extinguished by dust. Confirming redshifts without the benefit of Lyman-alpha emission is extremely difficult. Even so, spectra taken with the Keck telescope have now done so for more than 100 galaxies having redshifts between 2 and 3.8, including about 20 (so far) within the HDF.

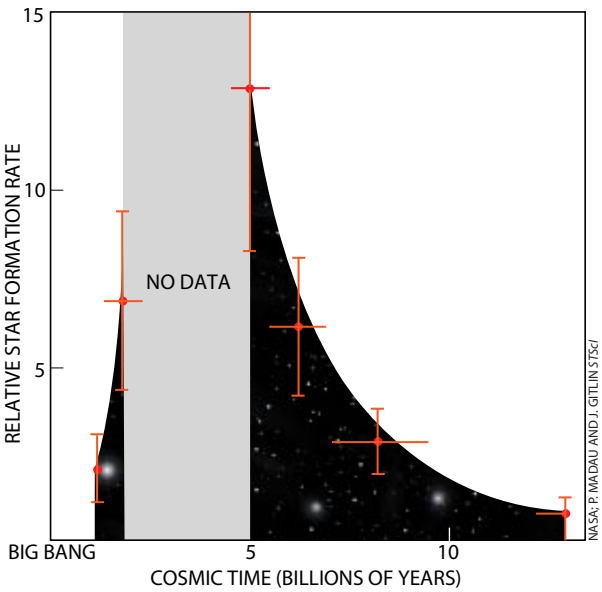
It is clear that a substantial population of galaxies was already present in the universe when it was only a few billion years old. How these objects relate to galaxies such as the Milky Way is much less certain. Are these the elusive primeval galaxies, the direct ancestors of today's spirals and ellipticals? Are they collapsing to form their first generations of stars?

The HDF images show that many of these objects were much smaller than galaxies like our own Milky Way. Many have bright knots and condensations thousands of light-years across—features resembling huge star-forming regions in some nearby galaxies. Many have close companions, suggesting that they are forming from the merging of small galaxies or even subgalactic fragments.

One theory is that such fragments, with sizes around one tenth that of the Milky Way, formed early on and that most of today's galaxies assembled from the merging of these smaller clumps. The mean density of the universe at redshift $z = 3.5$ was 90 times higher than it is now, and the chances for encounters and mergers were correspondingly much larger. Through this process of frequent mergers, and helped by the clumping of matter caused by gravitational attraction, galaxies may have built up gradually until they reached the sizes and masses that are typically found now.

Combining data on nearby galaxies with results from the deep redshift surveys and the ultraviolet dropout techniques, astronomers have sketched the global history of star formation, starting with the first few billion years of the universe's history. The picture that is emerging suggests that the rate of star formation climbed steeply as the universe expanded during the first 20 to 30 percent of its age. It then peaked, perhaps somewhere between redshifts of 2 and 1—an epoch still poorly explored by today's observations. And it has gradually declined again ever since.

Currently, averaging over all galaxies, stars are forming at a rate less than 10 percent of what it was at its peak. The universe has apparently settled into a quiet maturity. This scenario remains incomplete, and new observations will likely revise it, but the fact that it exists demonstrates the astonishing advances made in observational cosmology during the past decade.



NASA; P. MADAU AND J. GITLIN *STScI*

STAR FORMATION was swift 12 billion years ago when the universe was young, as shown on this graph, based on observations of distant galaxies using the Hubble telescope and earthbound telescopes. About nine billion years ago, the star formation rate—for reasons yet to be discovered—began to decline dramatically.

New Directions

How will astronomers study the high-redshift universe in the future? One exciting possibility takes advantage of a phenomenon known as gravitational lensing. Large masses, such as dense galaxy clusters, can bend, amplify and distort the image of objects beyond them. A faint galaxy behind such a cluster can be magnified and stretched out into a giant arc, revealing morphological details beyond the limits of the best telescopes.

Gravitational lensing also makes distant galaxies appear much brighter and thus easier to analyze spectroscopically. Using these natural lenses to further boost the resolution of Hubble, astronomers have obtained stunning images of gravitational arcs, some of which have been confirmed to be intrinsically faint galaxies at redshifts comparable to those of the ultraviolet dropout population discussed earlier.

Also, telescopes and instruments now in the making will extend the work described here to larger samples and greater redshifts. A suite of giant telescopes is being built on mountains around the world. This past February U.S. astronauts returned to Hubble and installed two instruments that will en-

NASA; W. N. COLLEY AND E. TURNER *Princeton University*; J. A. TYSON *Bell Labs, Lucent Technologies*

GRAVITATIONAL LENSING has enabled cosmologists to see greater detail within faint galaxies. The many blue, looped objects in the picture above are multiple images of the same distant galaxy. The yellow elliptical and spiral galaxies clustered at the center of the photograph, named 0024 + 1654 for their position in the sky, possess a tremendous gravitational field. This field bends light and so distorts, magnifies or duplicates the images of objects behind it.

hance the telescope's capabilities and open new windows on the distant universe. And the European Infrared Space Observatory, a satellite probing the longer wavelengths of light where distant, dusty galaxies may emit much of their energy, has recently surveyed the HDF.

The U.S. is preparing its own infrared mission, an advanced telescope called the Space Infrared Telescope Facility. Looking further ahead, the National Aeronautics and Space Administration has begun considering designs for a Next Generation Space Telescope, a successor to Hubble that could probe the earliest epochs of galaxy history. With these new instruments, and most especially with continued creativity in developing novel techniques for finding and studying objects in the distant universe, astronomers can dig deep enough to characterize the evolution of galaxies and perhaps reveal at least a few of the secrets of their formation. SA

The Authors

F. DUCCIO MACCHETTO and MARK DICKINSON are both affiliated with the Space Telescope Science Institute in Baltimore, Md. There Macchetto is associate director for science programs for the Hubble Space Telescope (HST) project. As a member of the European Space Agency (ESA), he has served as the ESA Hubble project scientist and as the principal investigator for ESA's Faint Object Camera on board the telescope. Dickinson is currently an Alan C. Davis Postdoctoral Fellow at John Hopkins University and at the Space Telescope Science Institute, where he previously held an Aura Fellowship.

Further Reading

THE FARTHEST THINGS IN THE UNIVERSE. Jay M. Pasachoff, Hyron Spinrad, Patrick S. Osmer and Edward Cheng. Cambridge University Press, 1994.

A SHORT HISTORY OF THE UNIVERSE. Joseph Silk. Scientific American Library, W. H. Freeman and Company, 1994.

FAINT BLUE GALAXIES. Richard S. Ellis in *Annual Review of Astronomy and Astrophysics*, Vol. 35, 1997 (in press).

Hubble Deep Field (STScI) site on the World Wide Web available at http://www.stsci.edu/ftp/science/hdf/hdf.html

SALLY MANN, *Untitled* (1988), from the exhibition *Hospice: A Photographic Inquiry*, courtesy of the Corcoran Gallery of Art

Seeking a Better Way to Die

by John Horgan, *staff writer*

One of the nagging ironies of modern medicine is that while it has enormously extended life spans, it has also stretched out the dying process. In his 1908 book *Science and Immortality,* the physician William Osler reported on his study of 486 deaths at Johns Hopkins Hospital in Baltimore. He found that only about one in five of the deceased seemed to be suffering in their final days. For "the great majority," he stated, "death was a sleep and a forgetting."

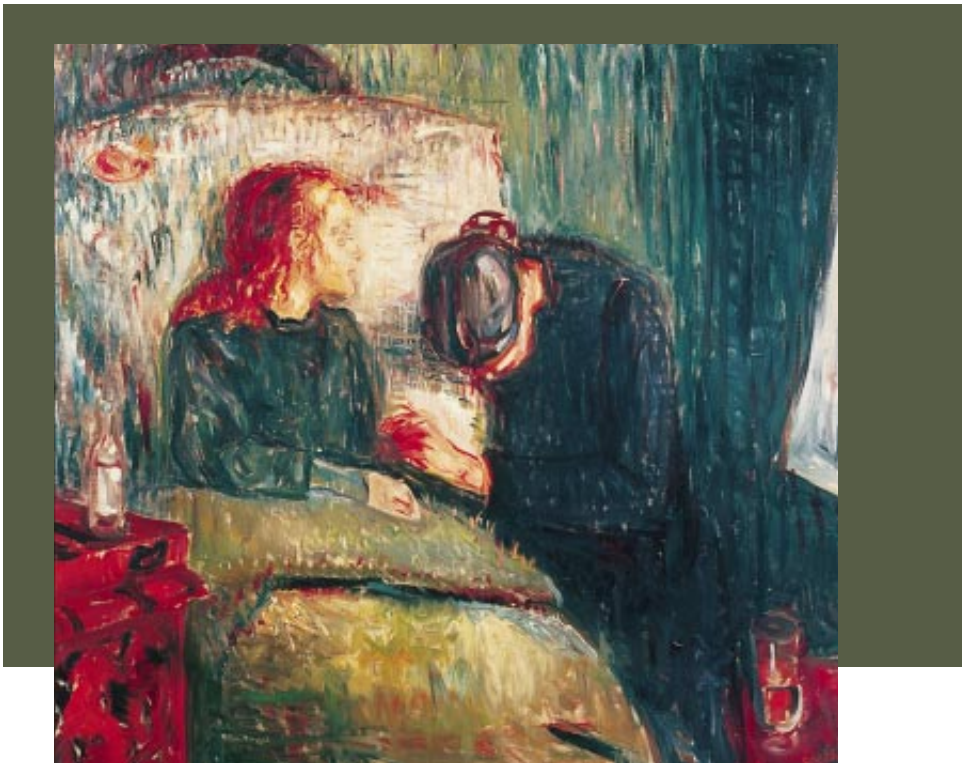
Of the 2.5 million people who now die in the U.S. annually, roughly two thirds succumb after protracted struggles with chronic illnesses such as heart disease and cancer. The dying are also increasingly "sequestered" from the rest of society, says David Rothman of Columbia University, who is writing a book on the history of death in America. The percentage of deaths occurring in hospitals, nursing homes or other institutions has risen steadily since the 1930s to nearly 80 percent today. And recent research has found that most of those who are conscious while dying feel distress.

Given this situation, it is hardly surprising that euthanasia is winning so many adherents. Polls show that a majority of Americans support the right of patients to receive a lethal overdose from their doctors. Last year a Michigan jury acquitted Jack Kevorkian, alias "Doctor Death," of charges related to his having helped two women kill themselves. Meanwhile appellate courts in California and New York struck down state prohibitions against physician-assisted suicide. The U.S. Supreme Court is expected to rule this summer on the constitutionality of such prohibitions.

The issue has riven the health care community. The American Medical Association (AMA), the American Nursing Association, the National Hospice Organization and dozens of other groups have filed briefs with the Supreme Court opposing physician-assisted suicide. Supporters of legalization include the American Medical Student Association, the Coalition of Hospice Professionals and Marcia Angell, editor of the *New England Journal of Medicine.*

Nevertheless, this dispute masks a deep consensus among health care experts that much can and should be done to improve the care of the dying. While the media have focused on euthanasia, a diverse collection of physicians' groups, foundations, hospitals and other organizations has quietly begun seeking alternative solutions. The avenues being explored include treating

As the U.S. Supreme Court ponders physician-assisted suicide, health care providers strive to improve care of the dying



TATE GALLERY, LONDON; ART RESOURCE, NEW YORK

PEACEFUL WAY OF DYING (*left*) **is the goal of hospices. Although hospice care is increasing in the U.S., most people in the industrial world still die in hospitals or nursing homes. When Edvard Munch painted *The Sick Child* (*right*) in 1907, most deaths occurred at home.**

physical and psychological distress more aggressively, relaxing restrictions on the use of opioids, educating health care workers and the public about the needs of the terminally ill, and expanding the use of hospices, which emphasize comfort rather than cure.

Some observers draw an analogy between euthanasia and abortion: the ideal situation in each case would be to reduce the need for such drastic solutions. "Even though people disagree on physician-assisted suicide," says Rosemary Gibson of the Robert Wood Johnson Foundation, a major supporter of programs addressing the needs of the dying, "they are working together on this."

The Lessons of SUPPORT

The largest investigation to date of the problems posed by end-of-life care is the Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments. Called SUPPORT, it followed more than 9,000 severely ill patients admitted to five teaching hospitals during the early 1990s. This past January the SUPPORT team published data on 3,357 deceased patients whose final days had been observed by their relatives.

According to the reports of the family members, 40 percent of those patients who were conscious experienced severe pain "most of the time," and more than 25 percent were anxious or depressed. Overall, almost two thirds of the patients "had difficulty tolerating" their condition, the researchers wrote in the *Annals of Internal Medicine*.

On the other hand, the study contradicted the widespread belief that tyrannical physicians are keeping many patients alive against their will. Nine out of 10 patients approved of their medical treatment, even invasive procedures that seemed contrary to desires indicated in living wills and other "advanced directives."

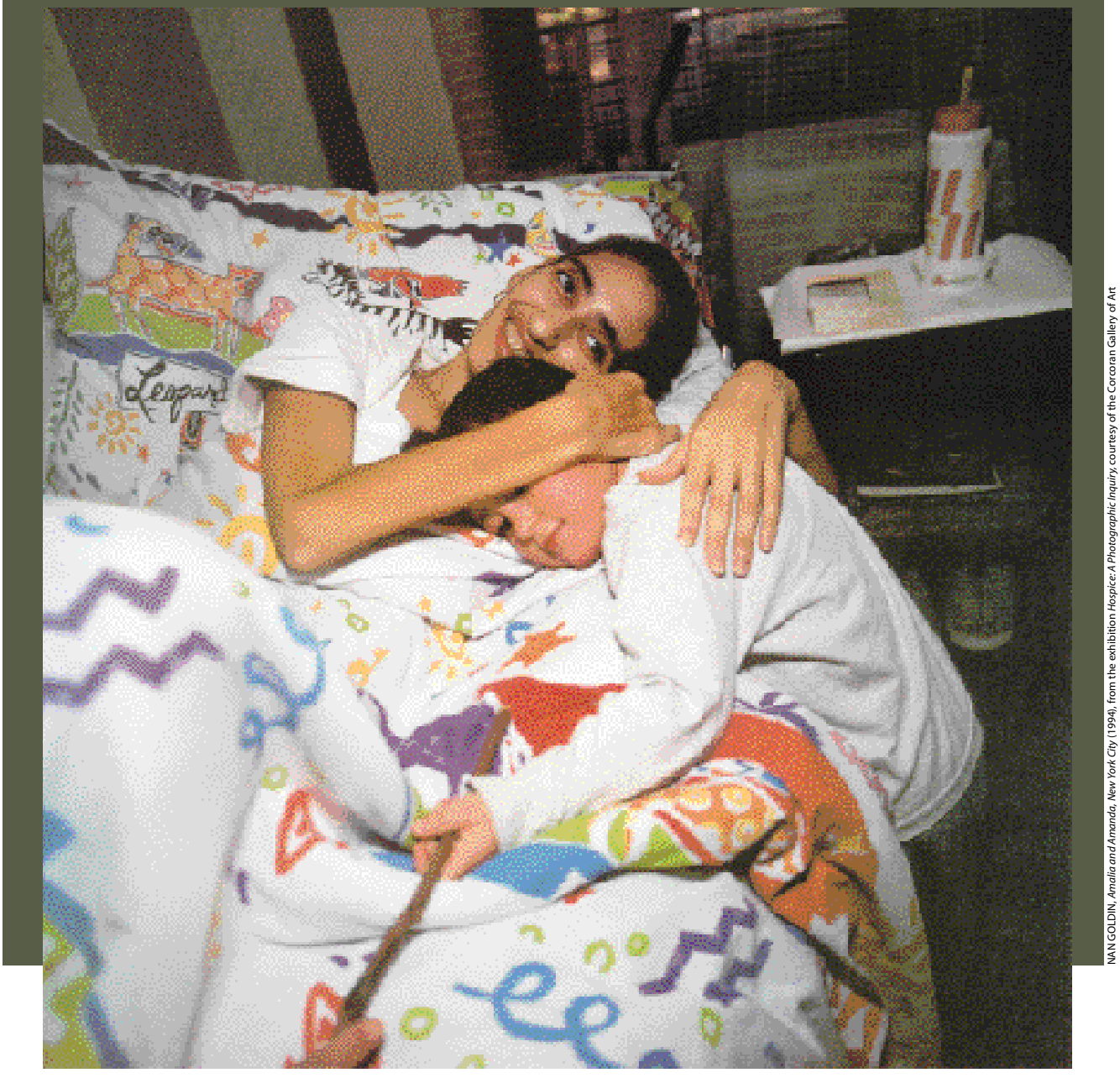
None of the patients requested or received lethal overdoses of drugs, at least to the knowledge of the researchers. "It is very rare for people this sick to be looking to die faster than they have to," says Joanne Lynn, the lead author of the paper and head of the Center to Improve Care of the Dying at George Washington University.

The tendency of even extremely ill patients to cling to life makes sense, Lynn adds, given the intractable uncertainties of medical prognoses. In a recent paper in the *Duquesne Law Review,* the SUPPORT group notes that of those subjects with congestive heart failure, the second most common cause of death (after cancer) in the U.S., 28 percent who were expected to live less than six months survived for at least another year.

Although predictions for cancer patients were more accurate, 13 percent of the lung cancer patients with prognoses of less than six months survived a year or more. Conversely, after examining prognoses calculated by a statistical model from the day before patients actually died, the researchers found that 17 percent were expected to live for at least two more months and 7 percent for six months.

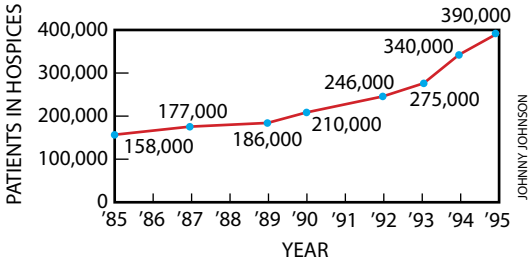
Expanding Hospice Care

Lynn thinks it is inappropriate to legalize assisted suicide when so little effort has been expended on exploring alternatives. "Having been involved with some wonderful hospices and nursing homes," she says, "I'm very impressed with what can be achieved with relatively modest investments." Indeed, many experts consider the most successful model for the delivery of palliative care to be hospices, which are often termed "the gold standard in care of the dying." Only a few decades old, the hospice concept involves providing comfort—includ-



NAN GOLDIN, *Amalia and Amanda, New York City* (1994), from the exhibition *Hospice: A Photographic Inquiry*, courtesy of the Corcoran Gallery of Art

CANCER PATIENT receiving hospice care in New York City embraces her daughter. The number of people benefiting from hospice services in the U.S. has surged to nearly 400,000 since Medicare coverage began in 1983, according to the National Hospice Organization.



JOHNNY JOHNSON

ing medical, psychological, social and even spiritual services—for those who are beyond cure and approaching death.

Since 1985, two years after the federal Medicare program began covering hospice care for people expected by physicians to live less than six months, the number of hospice providers in the U.S. has surged from 500 to more than 2,500. Some 400,000 patients—about one sixth of all those dying—now receive some hospice care every year in the U.S., according to the National Hospice Organization, based in Arlington, Va. More than 80 percent must be served in their homes, under the terms of the Medicare legislation; the rest are in hospitals, nursing homes or facilities dedicated to hospice care.

Studies have estimated the median duration of hospice care at just over one month. Medicare provides hospice coverage for indefinite periods, but payments lasting much more than six months are reviewed; Medicare officials can retroactively deny payment or require co-payment for patients whose prognoses are deemed to have been too positive.

Most hospice patients are dying of cancer, although an increasing number are afflicted with AIDS. John J. Mahoney, president of the National Hospice Organization, says many more people should be eligible for hospice treatment, and they should be admitted at earlier stages in their disease.

Hospice care is also economical, according to Mahoney. A 1995 study sponsored by the hospice organization examined records of all Medicare beneficia-

ries who died of cancer in 1992. The investigation—which was not controlled for independent variables—found that Medicare spent almost 50 percent less on hospice patients during their last month of life than on those receiving standard care. Medicare currently spends well over $1 billion annually on hospice reimbursements, and the average daily cost for individual patients is about $100, roughly the same as for nursing home care.

The effort to expand the hospice population has run into a roadblock recently as a result of an investigation by the U.S. Department of Health and Human Services, which oversees Medicare. Last year department inspectors charged hospices in Puerto Rico, Florida and elsewhere with admitting patients who were not terminally ill and thus not eligible for Medicare funding; some patients had allegedly remained at hospices for three or four years. Mahoney fears that the accusations, which have been disputed, will make doctors and hospices even more conservative in deciding when patients should be enrolled in hospice care.

Combating Pain

One possible solution to this problem—proposed by Lynn of the Center to Improve Care of the Dying—is a program called MediCaring. It is intended to extend hospicelike services to patients who have incurable diseases but do not qualify for hospice care because they have unpredictable conditions or are expected to live longer than Medicare guidelines allow. Unlike hospice patients, participants in MediCaring would in some cases be eligible for aggressive, expensive treatments such as chemotherapy, bone marrow transplants and surgery, although access to such treatments would be restricted. Pilot projects employing MediCaring approaches are now under way in four states and the District of Columbia.

All patients, not just those in hospices, should be able to receive expert palliative care, according to Kathleen M. Foley of the Memorial Sloan-Kettering Cancer Center in New York City. An authority on the treatment of pain, Foley is also director of the Project on Death in America. Created by the investor and philanthropist George Soros in 1994, the New York–based project has already dispensed more than $15 million for programs on death and dying. They range from studies of how different ethnic groups view death to an initiative of the United Hospital Fund to improve end-of-life care at 12 New York hospitals.

Foley opposes legalizing physician-assisted suicide, which she denigrates as "treating suffering by eliminating the sufferer." In the course of her career, she says, she has repeatedly encountered patients who asked to be put out of their misery. In almost every case, she says, the requests abated after the patients had received supportive care, including analgesics, antidepressants or counseling.

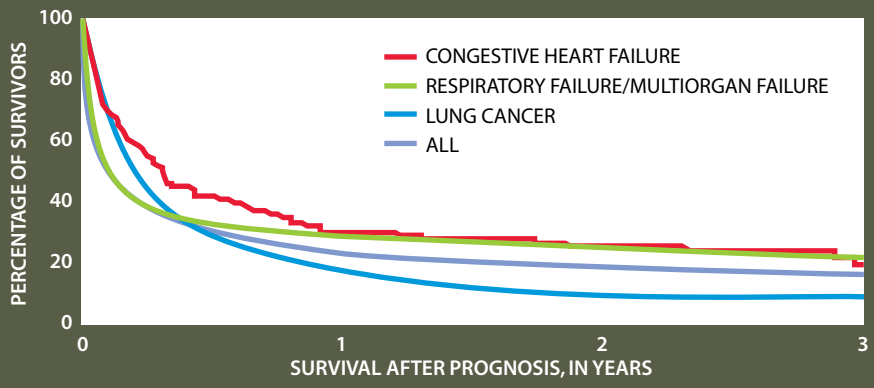
There have been enormous advances in the management of pain, Foley asserts. Drugs can be delivered through skin patches, topical creams and implanted pumps as well as intravenously. New automated delivery systems, which measure levels of medication in the blood, can stop pain before it starts while minimizing side effects such as nausea, grogginess and constipation. Too few health care professionals, Foley adds, are integrating these advances into practice—in large part because they have not been trained to view end-of-life care as an important part of medicine. Searching the medical literature for articles on "death," she notes, will yield more articles on cell death than on human death.

Foley advocates making pain a "vital sign" that is monitored along with other important parameters of health by nurses and physicians. Pain should also be treated as early as possible in the trajectory of a disease, according to Neil MacDonald, an oncologist at the Clinical Research Institute of Montreal. "Preventing a problem is preferable to reacting to it," he states. Recent animal studies, he says, reveal that chronic pain can alter the central nervous system, making the discomfort more severe and intractable through a "kindling effect."

Clinical observations, MacDonald notes, suggest that the seriously ill may also become trapped in a vicious cycle of intensifying distress. Chronic pain often becomes more severe over time, he explains, leading to psychological distress that in turn makes the physical pain harder to endure. In the same way, shortness of breath, which is common among the dying, may trigger panic that exacerbates the breathing problem.

Major obstacles to proper treatment of pain, MacDonald points out, are the laws and social attitudes that hinder prescriptions of analgesics—and opi-

How Long Do "Terminal" Patients Really Live?



JOHNNY JOHNSON

Prognostic uncertainties were underscored by SUPPORT, the Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatment. A significant percentage of a group of 3,693 patients given no more than a 50 percent chance of living for six months survived well beyond that period.

RESTRICTIONS ON OPIOIDS and unwarranted fears about the dangers of addiction hamper the treatment of pain.

ALEXEI HAY

oids such as morphine in particular. This problem is being addressed by the Pain and Policy Studies Group at the University of Wisconsin. In 1991 the researchers surveyed state medical boards—which monitor prescriptions of drugs and discipline physicians deemed to be overprescribing—on their attitudes about painkillers. The survey determined that board members had too little appreciation of the importance of pain management and excessive anxiety about the dangers of addiction. Many members wrongly equated addiction with physical dependency on a drug, says David E. Joranson, head of the study group. In fact, Joranson explains, addiction is a psychological condition characterized by obsessive craving. Many patients become physically dependent on opioids—requiring increased doses to bring about the same effect and displaying withdrawal symptoms when medication ceases—but psychological addiction is rare.

The researchers have held six workshops to educate medical board members in different states about the importance of pain treatment and the true risks of opioids. They plan to conduct a second survey of state medical boards in 1999 to determine how much the boards' attitudes and practices have changed. The Wisconsin investigators are also working with the World Health Organization and other groups to reform international regulations hampering the prescription of opioids for pain treatment.

Physicians, Educate Yourselves

This summer the American Medical Association plans to launch a program to educate physicians about palliative care. Spearheading the program is Linda L. Emanuel, director of the AMA's Institute for Ethics. Emanuel was the primary author of the brief opposing physician-assisted suicide that the AMA submitted to the Supreme Court. In her past 12 years as an internist, she says, only two patients have asked her to kill them; both changed their minds after she made it clear that they had other options for treating their distress.

Nevertheless, "the need to improve things is enormous," Emanuel says. The AMA educational project, she explains, will encourage physicians to discuss patients' attitudes toward dying in advance of terminal illness and to learn how to treat common afflictions of the severely ill, including pain, bedsores, incontinence, depression and psychosis.

During the program's first phase, panels of experts will discuss palliative care and other issues with health care leaders from many states at a few large, regional conferences; those meetings will be followed by many smaller assemblages for separate states and communities. "We hope to directly train about half of all physicians" in the U.S. "and indirectly reach all the others," Emanuel says. The goal, she adds, is to "change the culture of medicine."

The Faculty Scholars Program could be another instrument of change. Funded by the Project on Death in America, it provides fellowships of up to $70,000 a year to "outstanding clinicians and academic leaders who want to make a career-level commitment to care of the dying," says Susan Block of Harvard Medical School, director of the program. Twenty-six scholars have already been chosen, and 10 more selections are pending.

Scholars are pursuing a variety of projects. These include revamping curricula for medical schools and residency programs and developing new models of service delivery, such as one in which teams of clinicians provide hospicelike care for terminally ill patients in hospitals. Such units have already been established at Massachusetts General Hospital, Sloan-Kettering and elsewhere.

Unlike Emanuel, Foley and others who oppose physician-assisted suicide, Block believes that for a very small number of people it may be the best way to die "on their own terms." Block nonetheless agonizes over whether such acts

RICK RICKMAN *Matrix*

MEDICAL TECHNOLOGIES provided by hospitals save lives, but they can also prolong the suffering of the terminally ill.

JOSEPH PLUCHINO

BETTER PALLIATIVE CARE will reduce demand for physician-assisted suicide, asserts Kathleen M. Foley of the Memorial Sloan-Kettering Cancer Center.

should be legalized. "I can't find a stance that feels right," she remarks. "I can't say it's always wrong, but I am afraid of what might happen with legalization." Moreover, she adds, "the vast majority of patients who want physician-assisted suicide have problems like depression that we can help them with."

But changing the practice of medicine is not enough, according to Ira Byock, a veteran hospice physician based in Missoula, Mont., and president of the American Academy of Hospice and Palliative Medicine. Several years ago Byock created the Missoula Demonstration Project to identify new solutions to end-of-life care. Although dying patients must receive competent medical care, Byock says, "this is too important to be left to the medical experts. We are exploring nonmedical ways to support patients and their families."

One goal of the project—which is supported by the Robert Wood Johnson Foundation, the Project on Death in America and other foundations—is to create what Byock calls a "very detailed, high-definition snapshot" of death and dying in Missoula. The study will compile information on all the deaths occurring within Missoula, including medical histories of the deceased and surveys of the attitudes of surviving family members. The project will also sponsor discussions in churches, schools and workplaces with the goal of finding innovative ways to make death less agonizing.

Premortem Eulogies

Possibilities under consideration are training volunteers to spend time with terminally ill people; allowing dying patients to express themselves through art or by recording their life histories for a permanent archive; combining child care facilities with hospices so that the very young and very old can mingle; using pets to raise patients' spirits; having choirs of neighbors sing songs outside the homes of the dying; having friends and relatives give eulogies to a dying person before death.

Researchers will then track how attitudes and behavior change over time. Comparison of Missoula to a control population in Pocatello, Idaho, will help determine if any of the interventions is working. Byock hopes the project will become a model for similar ones across the country. Eventually, the project may do for end-of-life care what the famous Framingham Heart Study—which for almost half a century has tracked the habits and cardiac histories of the residents of Framingham, Mass.—has done for the treatment of heart disease.

The problems posed by caring for those near death will surely become more pressing as society ages. The percentage of the U.S. population age 85 or older is expected to grow from just over 1 percent, its current level, to 5 percent by the middle of the next century, according to the Alliance for Aging Research, a nonprofit group in Washington, D.C. Several studies nearing completion should soon inject more data into the discussion of end-of-life care. The Institute of Medicine, a branch of the National Academy of Sciences, has been examining the issue for more than three years and is expected to publish a major report on the subject this summer.

A group led by Diane E. Meier of the Mount Sinai Medical Center is also about to make public the largest survey to date of physicians' attitudes toward assisted suicide and related issues. The preliminary data suggest that doctors with the most experience prescribing lethal doses of drugs to patients tend to be more comfortable with their decisions. Meier finds that tentative result disturbing. Ideally, she says, physicians committing such acts would never "get over their fear and trembling."

Many health care workers have already privately admitted to having helped patients die. According to a report in the May 23, 1996, issue of the *New England Journal of Medicine,* 20 percent of a group of 850 nurses working in intensive care units acknowledged having deliberately hastened the death of a patient. A survey of 118 San Francisco–based doctors, published in the *New England Journal of Medicine* this past February, found that half had prescribed lethal doses of drugs to patients suffering from AIDS.

Such acts are likely to become still more common if the Supreme Court rules in favor of physician-assisted suicide this summer. Sloan-Kettering's Foley contends that, given its lack of knowledge about palliative care, the medical profession is simply not prepared for such a responsibility. "Doctors don't know enough to kill," she declares. SA

by Shawn Carlson

When Hazy Skies Are Rising

Haze is a vital indicator of our atmosphere's health. Most haze is natural; it is created by water vapor, wind-borne dust, forest fires and volcanic eruptions. Unfortunately, a lot of what humans do also ends up in the atmosphere—smog is the best-known example of artificial haze. Although many cities monitor the clarity of their local atmosphere, surprisingly little is known about how the amount of haze is changing globally because no one is coordinating haze observations over widely dispersed areas.

That may change with the latest design from Forrest M. Mims III. (Mims may be familiar to readers from his columns in this section in 1990.) He has invented an atmospheric haze sensor that costs less than $20 and is so simple to construct that even the most hardened technophobe can put it together in under an hour. Mims's instrument could revolutionize this important area of study by opening the field to all comers, that is, to amateur scientists.

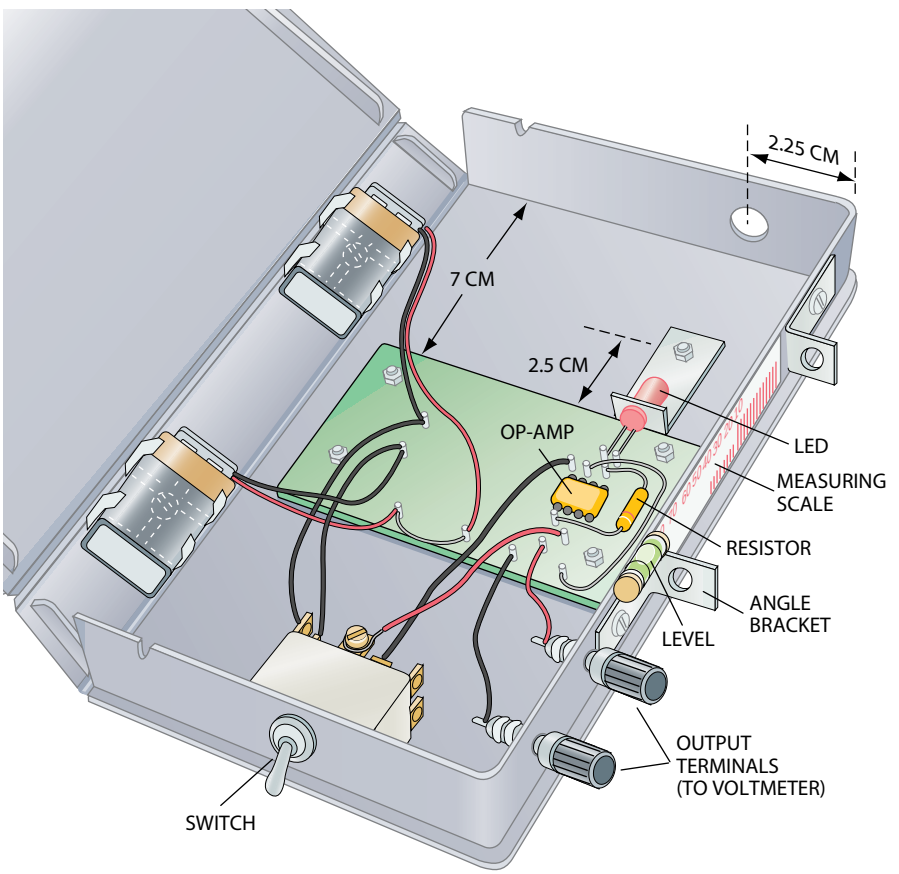
The Visible Haze Sensor exploits the fact that sun is an ideal probe for measuring haze. The intensity of sunlight striking the top of the earth's atmosphere is essentially constant. As a result, by measuring the intensity of the sunlight at the ground and knowing the thickness of the atmosphere it has passed through, we can determine how much light has been scattered or absorbed and hence how much haze there is.

There is one complication. Air molecules themselves scatter light, in a phenomenon known as Rayleigh scattering. Although Rayleigh scattering creates beautiful blue skies and brilliant sunsets, it also complicates measurements of haze. Because the efficiency with which air molecules scatter light depends on wavelength, scientists restrict their measurements to a narrow sliver of the solar spectrum so that the effects of Rayleigh scattering can be easily corrected for.

Most professional haze instruments employ a broad-spectrum photodetector, coupled with an expensive narrowband filter to achieve the necessary selectivity. Mims realized that a narrowband detector would serve the same purpose. The perfect device for the amateur scientist is a light-emitting diode (LED), which generates light of a sharply defined color when a current passes through it. But this process is easily reversed: light falling on the LED creates a small current that can be readily detected. Furthermore, just as LED emissions appear only in a narrow wavelength band, the diode generates current only when stimulated by light within a small range of colors.

Mims's device uses a green LED (Radio Shack 276-022A), which emits light at around 555 nanometers and detects light at around 525 nanometers. It costs less than a dollar. The rest of the circuit consists of one resistor, an operational amplifier, two nine-volt batteries and a voltmeter. Mims houses his instrument inside a plastic VHS videocassette case with a hole drilled in one end to admit sunlight [*see illustration below*].

Two angle brackets on the side of the case let you align the instrument directly with the sun. (If possible, attach them with a hot-glue gun to make a rigid connection.) Place a small piece of white tape over the bottom bracket. Open the case and point the instrument toward the sun. Move it around until the bright sun spot is centered on the LED. If the outside brackets on the side are approximately aligned, a second image of the sun should appear somewhere on the tape. Lightly mark its center in pencil.



IAN WORPOLE

HAZE SENSOR
fits neatly into a videocassette case. Angle brackets align the case to the sun. They can also be used with the level to calculate the sun angle.

The case may flex slightly when closed, so you may need to adjust your mark. Close the case and align the instrument to the sun using the outside brackets. Then tilt the instrument slightly in different directions while watching the voltmeter and find the orientation that produces the largest voltage. Make a permanent mark at the center of the sun's image, and you're ready to go.

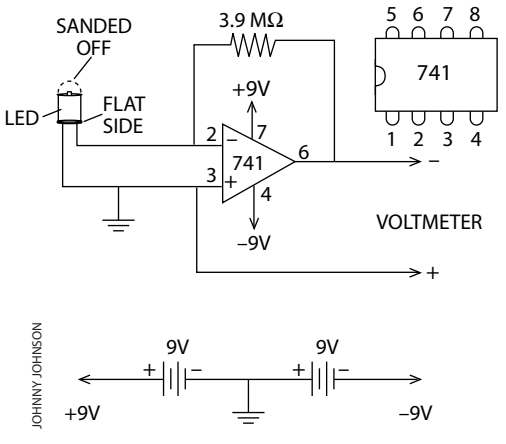
Although you can begin collecting data right away, you will eventually need to calibrate your photometer. To do that, you'll have to set aside half a day (either early morning to solar noon or solar noon to early evening) when the sky is clear blue and there are few or no clouds. First, you'll need to record the dark signal, that is, the voltage produced when no light strikes the detector. Cover the sun port and record the voltage. You'll need to subtract this number from every measurement you make.

Measure the voltage every 20 minutes when the sun is high overhead and every seven minutes when it is low on the horizon. Always record the time, the sky condition near the sun, and the angle of the sun above the horizon. (Keep your data reliable by recalibrating your instrument at least once a year.)

To complete the calibration, you'll need to plot your data. First, calculate the air mass for each measurement. The air mass is the depth of the air column between you and the sun; by convention it is 1.0 when the sun is straight overhead (a sun angle of 90 degrees) and infinite as the sun sets. The formula is 1/sin(sun angle). After compensating for the dark voltage, plot the natural logarithm of your photodetector measurements against the air mass. The result is called a Langley plot and should be a straight line out to an air mass value of about 10.

If you project that line back to an air mass of zero [*see upper illustration in box above*], the result will be the logarithm of the voltage that your instrument would read if you could use it to measure the sun's brightness just above the atmosphere—the so-called extraterrestrial constant, or ET. This constant is the starting point for calculations of aerosol optical thickness, which quantifies haze.

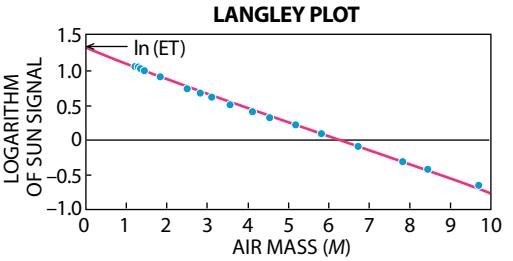
I can't think of a better way for students and adults alike to become more aware of their environment than to get involved in monitoring it. Committing to a regular program of measurements teaches young people responsibility as well as science. You can even become part of a worldwide network of haze observers by submitting observations via the Internet. Visit the TERC site for detailed instructions. SA



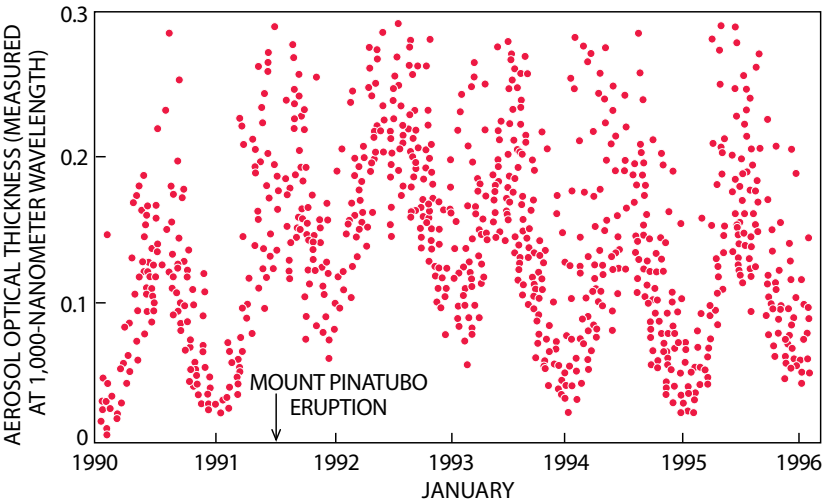
***SENSOR CIRCUIT** is simple to assemble.*

Measuring Haze

Aerosol optical thickness (AOT) is the number that ultimately expresses haze measurements. To calculate it, you must subtract out your photodetector's dark voltage and compensate for the effect of Rayleigh scattering and the angle of the sun. The equation is $AOT = [\ln(ET) - \ln(V_s - V_d) - (M \times P)/8660]/M$, where ET is the extraterrestrial constant (the signal that would be measured above the atmosphere), V_s is the measured voltage, V_d is the dark voltage, M is the air mass (as calculated from the sun angle), and P is the atmospheric pressure in millibars. The constant 8,660 is an empirical factor for Rayleigh scattering at the wavelength detected by the LED.



The graph below shows the AOT taken by Forrest Mims from his home in Seguin, Tex. The seasonal variation, with summer skies more hazy than winter ones, is clear.



The Visible Haze Sensor photometer was developed through a grant from the National Science Foundation to the Global Lab program at TERC in Cambridge, Mass. TERC has set up a haze monitoring site on the World Wide Web at http://www.concord.org/haze/

For more information about this project, check out the Society for Amateur Scientists's World Wide Web site at http://www.thesphere.com/SAS/ The SAS can be reached at 4951D Clairemont Square, Suite 179, San Diego, CA 92117 or at (619) 239-8807 or (800) 873-8767. I gratefully acknowledge informative conversations with Forrest Mims.

by Ian Stewart

Big Game Hunting in Primeland

This is my 50th "Mathematical Recreations" column, and so it seems appropriate to take a look at big numbers. These have risen to unexpected prominence in cryptography, where the central concerns nowadays are primes and factors. For example, the RSA coding system—named after its inventors Ronald L. Rivest of the Massachusetts Institute of Technology, Adi Shamir of the Weizmann Institute of Science in Israel and Leonard M. Adleman of the University of Southern California—is based on numbers that are the product of two large primes—say, around 100 digits each. I won't describe the cryptosystem itself but the two underlying mathematical issues, which are:

- Primality testing: Given a big number, how can you decide whether or not it is prime?
- Factoring: Given a big number, how can you find its factors?

Recall that a factor of a number is any number that divides it exactly and that a prime is a number having no factors other than itself and 1. Obviously, any solution to the factoring problem solves the primality testing problem, but the factoring problem seems considerably harder. Right now it is a feasible task to test an arbitrary 200-digit number for primality but not to find its factors (unless it has special features). This makes the exploration of Primeland, the world of prime numbers, a lot easier than that of Factorland, the world of all integers. Nevertheless, the big game hunters of number theory have made extensive inroads into both. This month's column is mostly about primality testing; next month's will be about factoring.

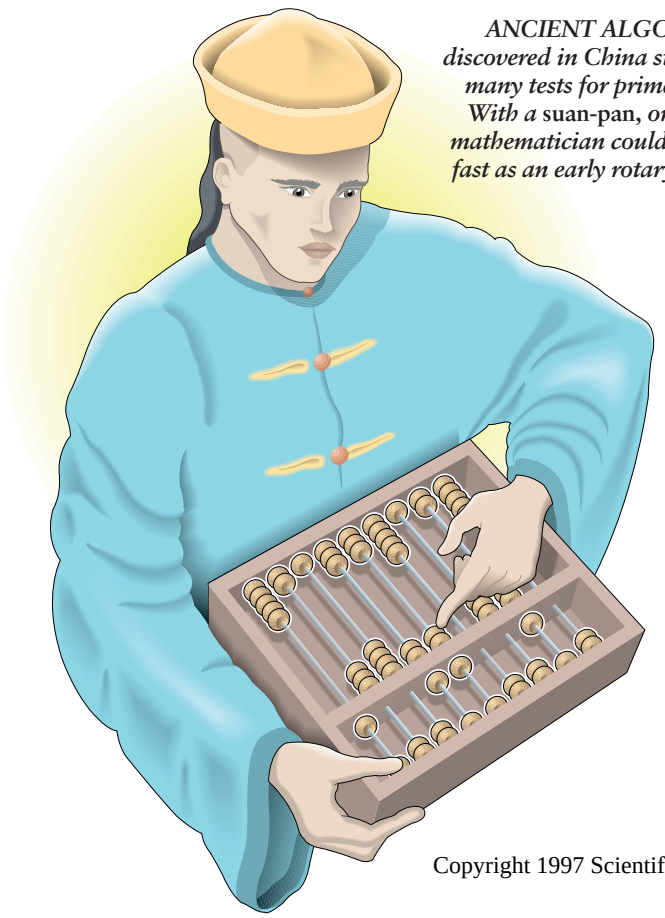
At school we are taught one foolproof method, or algorithm, for finding factors: trial division. Given a number such as 1,997, try in turn all possible divisors up to its square root (which in this case is 44 and a bit). It is enough to try all prime divisors; but in general that requires access to a list of primes, so the usual compromise is to try 2 and all odd numbers up to the square root. Here we would try 2, 3, 5,…, 43—a total of 22 divisors. None of them divides 1,997 exactly, so 1,997 is prime. Trying the same method on 1,921, we do find a factor: 17. Then we work out $1,921 \div 17 = 113$ and repeat the method on 113, finding that to be prime. So the complete factorization is $1,921 = 17 \times 113$.

For relatively small numbers, this approach works fine, but it fails utterly in the more distant reaches of Factorland. For a number such as 98,765,432,123,456,789, with a mere 17 digits, the number of trial divisions required is around 157 million. In general, for an n-digit number the method requires about $\sqrt{n}/2$ trials. If n has 100 digits, then the number of trials is of the order of 10^{50}, and the fastest supercomputer, running for the current lifetime of the universe, would barely have gotten started.

Estimating the running time of algorithms belongs to a relatively new area of mathematics known as Complexity Theory. Such estimates focus on the number of decimal digits and find no essential difference in complexity between $\sqrt{n}/2$ and $\sqrt{n}$. In fact, $\sqrt{n}$ either has the same number of digits as $\sqrt{n}/2$ or just one more. You might call this kind of argument the Principle of Permanence of Large Numbers. In gauging the efficiency of various methods, I will simplify expressions in this manner, if the number of digits doesn't change much.

Euclid proved that there are an infinite number of primes, so they never run out. They do, however, thin out a bit: Carl Friedrich Gauss conjectured and Jacques Hadamard and Charles de la Vallée Poussin proved that the number of primes less than n is approximately $n/\log n$ (where the logarithms are "natural" ones, to the base $e = 2.71828\ldots$). The primes are scattered in apparently random fashion. It can't really be random, but it sure looks unpredictable.

There are many ways to demonstrate that a number is prime without seeking its factors. One example is Wilson's



ANCIENT ALGORITHM discovered in China still motivates many tests for prime numbers. With a **suan-pan**, *or abacus, a mathematician could compute as fast as an early rotary calculator.*

JUAN VELASCO *Computer Graphics*

Theorem from 1770, which says that n is prime if and only if it divides $(n-1)! + 1$. Here $m! = 1 \times 2 \times 3 \times \ldots \times m$ is the factorial function. For example, $10! + 1 = 3{,}628{,}801$ is divisible by 11, so 11 is prime. Unfortunately, this method fares even worse than trial division in a complexity analysis, because it requires roughly n multiplications to be carried out (as opposed to $\sqrt{n}$ trial divisions).

Modern primality tests center on a similar discovery that goes back to ancient China: if n is prime, then it divides $2^n - 2$. It seems the Chinese thought that the converse was also true, in which case it would be possible to use their criterion as a primality test. But wouldn't it also take about n multiplications to compute $2^n - 2$ and test it for divisibility by n? Curiously, the answer is no, and it is here that we begin to make serious inroads into the Principle of Permanence of Large Numbers.

Imagine, for the sake of argument, that we are interested in the status of the number 107 and want to check whether it divides $2^{107} - 2$. It certainly looks as if we need 106 multiplications: 2, 2×2, $2 \times 2 \times 2$,. . . . But there is another way. First we work out only the powers 2^2, 2^4, 2^8 and so on, up to 2^{64} in this case, where the exponents are the powers of 2. We do this by repeated squaring: $2^2 = 2 \times 2$, $2^4 = 2^2 \times 2^2$, $2^8 = 2^4 \times 2^4, \ldots, 2^{64} = 2^{32} \times 2^{32}$. That takes only six multiplications altogether.

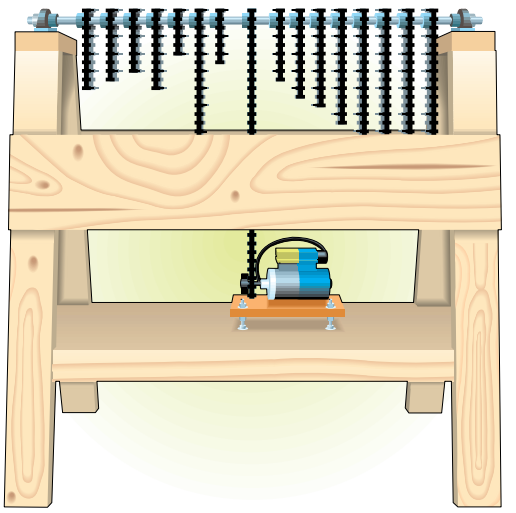
Next we write 107 in binary notation, $107 = 1101011$. This means that $107 = 1 \times 64 + 1 \times 32 + 0 \times 16 + 1 \times 8 + 0 \times 4 + 1 \times 2 + 1 \times 1$, which implies that $2^{107} = 2^{64} \times 2^{32} \times 2^8 \times 2^2 \times 2^1$, requiring only four more multiplications. In short, we can calculate 2^{107} using only 10 multiplications, not 106.

The improvement is even greater as the number gets larger. A complexity analysis shows that calculating 2^n by this method needs at most $2\log_2 n + 1$ multiplications. Here $\log_2 n$ is the logarithm of n to base 2, which is roughly 3.4 times the number of decimal digits in n. So to compute 2^n when n has 100 digits requires a mere 340 multiplications, not the 10^{100} or so needed if you do it by repeatedly multiplying by 2.

You might be worried that the numbers involved get rather large, so that the multiplication steps take longer and longer. In fact, we don't need to calculate $2^n - 2$ in its full glory: all we want

to know is whether it is divisible by n. The trick is to carry out all the arithmetic modulo n—meaning that we disregard any integer multiples of n. All numbers, at any stage of the calculation, can be replaced by their remainders on division by n.

For large n, this simplification makes a huge difference. Suppose n has 101 digits; then the "raw" value of $2^n - 2$ has roughly $10^{99.5}$ decimal digits. Not 99.5 digits: 10 to that power. The universe is too small to hold this number, even if we could inscribe one digit on every electron in existence. In contrast, carrying out the arithmetic modulo 10^{100} implies that no number ever exceeds 100 digits.



JUAN VELASCO *Computer Graphics*

NUMBER MACHINE
built in 1926 aided the search for primes.

Although the Chinese observation is not a foolproof test for primality, its exceptions—nonprimes n that divide $2^n - 2$, otherwise known as pseudoprimes to base 2—are rare. (One example is $n = 341 = 11 \times 31$.) Many of these exceptions can be dealt with by employing a generalization of the Chinese test, proved by Pierre de Fermat and known as Fermat's Little Theorem: if n is prime, then n divides $a^n - a$ for any a. So instead of testing with $2^n - 2$, we can use $3^n - 3$, $5^n - 5$ and so on. Again there are exceptions that pass the new divisibility test but are not prime: they are called pseudoprimes to base a. Numbers can be pseudoprimes to several bases: for instance, $2{,}701 = 37 \times 73$ is a pseudoprime to bases 2 and 3. But it is very rare for a number to be a pseudoprime to a variety of different bases. The only number less than 25 billion that is a pseudoprime to bases 2, 3, 5 and 7 is 3,215,031,751—and that's not a pseudoprime to base 11. So by combining a small number of Chinese-type tests and knowing this single exception, we can rapidly test any number up to 25 billion.

FEEDBACK

The most unusual response to the November 1996 column on calendars was a compact disc from Bob Gramann of Fredericksburg, Va.: *Mostly True Songs,* whose fourth track speculates about a decimalized calendar.

John S. Young of Scottsdale, Ariz., described a system to maintain a mental calendar. Start with a year code, the number of years after 1900 plus the number of leap days since (and including) 1900, taken modulo 7. Add a month code, derived by forming a running total mod 7 of each month's excess over 28. Add the date, again mod 7. Then 0–6 represent Sunday through Saturday. Paul A. Delaney of Laurel, Md., and Robert Eisenberg of Joliet, Ill., suggested similar methods.

In addition, Eisenberg described the fascinating theorem that in any year, leap year or not, all the dates from 1 to 31 will be matched with every possible day of the week. (That is, the first of the month will be a Tuesday for at least one month of the year and so on.) He also showed how to find the years that have the maximum number of Friday the 13ths: 1998 will be such a year. (I am reminded of Walt Kelly's Pogo cartoons, in which one running gag was "Friday the 13th falls on a Tuesday this month.") Charles D. Reed of Washington, D.C., and Ken Wood of Blackwood in Wales sent methods for finding the day of a given Gregorian or Julian date. And Peter Baum of Onset, Mass., explained how to find the "Rata Die" simply, by basing the calculations on a "year" that begins March 1. For details, contact him by e-mail at pbaum@capecod.net —*I.S.*

Unfortunately, every base has infinitely many pseudoprimes, so we can't just find a really good base and use only that. Indeed, matters are rather worse: there exist so-called Carmichael numbers, which are pseudoprimes to any base that does not have a factor in common with the number itself. The smallest Carmichael number is $561 = 3 \times 11 \times 17$; the next few are 1,105; 1,729; 2,465; and 2,821. In 1912 Robert D. Carmichael of Indiana University found 15 such numbers and opined that "this list might indefinitely be extended." His conjecture was finally proved in 1992: there are at least $n^{2/7}$ Carmichael numbers smaller than n.

To get around the pseudoprime difficulty, Gary L. Miller, now at Carnegie Mellon University, discovered a more complicated variant of the Fermat test, whose exceptions he called strong pseudoprimes. Every nonprime has at least one "witness," a base to which it is not a strong pseudoprime. If it can be shown that these witnesses are sufficiently small, one may test a number for primality by trying out all potential witnesses. By appealing to a notorious unsolved problem in number theory, the Riemann Hypothesis, Miller showed that every nonprime n has a witness of size at most $70(\log n)^2$. Because $\log n$ is much smaller than n, this estimate leads to a very efficient primality test.

In 1980 Adleman and Robert S. Rumely of the University of Georgia found a way to modify Miller's test to avoid appealing to the Riemann Hypothesis. The running time is now a little longer, but their method is entirely feasible for, say, 200-digit numbers. Subsequently, other, more sophisticated primality tests have been developed, notably the elliptic curve algorithm of Hendrik W. Lenstra, Jr., of the University of California at Berkeley.

These discoveries put mathematicians in a frustrating position. Given a 200-digit number, for example, they can rapidly carry out a test on an ordinary desktop computer to see whether the number is prime. If the answer is no, it must have factors. But the method offers no clue regarding what they are. Finding them, as we shall see next month, is decidedly difficult; still, there are some clever new ideas, and the factoring problem may yet succumb. If it does, many cryptosystems will be in trouble. SA

SCIENTIFIC AMERICAN CORRESPONDENCE

Reprints are available; to order, write Reprint Department, Scientific American, 415 Madison Avenue, New York, NY 10017-1111, or fax inquiries to (212) 355-0408. E-mail: info@sciam.com

Back issues: $8.95 each ($9.95 outside U.S.) prepaid. Most numbers available. Credit card (Mastercard/Visa) orders for two or more issues accepted. To order, fax (212) 355-0408.

Index of articles since 1948 available in electronic format. Write SciDex®, Scientific American, 415 Madison Avenue, New York, NY 10017-1111, fax (212) 980-8175 or call (800) 777-0444.

***Scientific American*-branded** products available. For free catalogue, write Scientific American Selections, P.O. Box 11314, Des Moines, IA 50340-1314, or call (800) 777-0444. E-mail: info@sciam.com

Photocopying rights are hereby granted by Scientific American, Inc., to libraries and others registered with the Copyright Clearance Center (CCC) to photocopy articles in this issue of Scientific American for the fee of $3.50 per copy of each article plus $0.50 per page. Such clearance does not extend to the photocopying of articles for promotion or other commercial purposes. Correspondence and payment should be addressed to Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923. Specify CCC Reference Number ISSN 0036-8733/96. $3.50 + 0.50.

Editorial correspondence should be addressed to The Editors, Scientific American, 415 Madison Avenue, New York, NY 10017-1111. Manuscripts are submitted at the authors' risk and will not be returned unless accompanied by postage. E-mail: editors@sciam.com

Advertising correspondence should be addressed to Advertising Manager, Scientific American, 415 Madison Avenue, New York, NY 10017-1111, or fax (212) 754-1138. E-mail: advertising@sciam.com

Subscription correspondence should be addressed to Subscription Manager, Scientific American, P.O. Box 3187, Harlan, IA 51537. The date of the last issue of your subscription appears on each month's mailing label. For change of address notify us at least four weeks in advance. Please send your old address (mailing label, if possible) and your new address. We occasionally make subscribers' names available to reputable companies whose products and services we think will interest you. If you would like your name excluded from these mailings, send your request with your mailing label to us at the Harlan, IA address. E-mail: customerservice@sciam.com

Visit our Web site at http://www.sciam.com/

PORTRAIT OF A PATHOGEN

Review by Paul Ewald

Yellow Fever Black Goddess: The Coevolution of People and Plagues
BY CHRISTOPHER WILLS
Addison-Wesley, 1996 ($24)

Much of biology and medicine has been a race to understand life processes at ever smaller scales. Researchers are now arriving at the finish line: the parts of molecules that regulate and encode life processes. The rapid pace of advances and the narrow focus of this effort, however, appear to have left in the dust the grander challenge that faces physicians and others concerned with public health. Somehow we must apply to the health sciences the entire spectrum of biology, from the small scale of molecular biology and biochemistry through superorganismal scales of ecology and evolutionary biology. In *Yellow Fever Black Goddess,* Christopher Wills provides an engrossing if flawed glimpse of the impending syntheses.

The majority of the book summarizes the views of the epidemiologists, population biologists, parasitologists and the like who are providing fresh perspectives on the intertwined fates of people and disease. Wills sometimes accomplishes these summaries with great skill, sometimes not. His historical analyses, for example, draw on the 1976 book *Plagues and Peoples,* by William H. McNeill, then a historian at the University of Chicago. But where McNeill often pushed the edge of knowledge in this branch of human ecology, Wills occasionally lags behind it. Both McNeill and Wills mistakenly advance the idea that evolution will uniformly favor benign pathogens over the long term. (Because parasites need hosts to survive, the reasoning goes, virulent pathogens would be at an evolutionary disadvantage.) But McNeill wrote before the problems with this notion were generally recognized. Wills, in contrast, continues to advocate long-term selection for stable coexistence, in spite of the now well-accepted arguments to the contrary.

GIRAUDON/ART RESOURCE, NEW YORK

***MALARIA** devastated Marseilles in 1720. The disease remains virulent today.*

The past 15 years of research have made it clear that evolutionary processes do not mandate benign associations between parasite and host when short-term forces of survival and reproduction push in the opposite direction. Just as with predators and prey, natural selection can lead to an unending series of population expansions and crashes. Yet Wills sees plagues as temporary blips in an otherwise balanced natural landscape:

> Plagues that afflict us do so because we have upset the balance of nature.... The conditions that engender plagues always end sooner or later. This always, *always* means that plague pathogens will not survive in the long term unless some of them manage subsequently to lose that virulence and revert to the habits of their low-profile relatives.

This conclusion roams beyond the realm of evidence. Plagues involve the presence of virulent pathogens in a large number of hosts. The end of a plague may or may not be associated with selection for lowered virulence. Many plagues burn themselves out because enough hosts die so that their population density decreases and the pathogen can no longer be transmitted effectively. If the selective pressures on the competing strains of pathogens favor the more vicious variants, however, the pathogens may maintain high virulence. A low profile may result from a low rate of infection rather than a loss of virulence.

Wills employs colorful metaphors to make some complex evolutionary ideas comprehensible, but his use of analogy also leads him astray. He engagingly casts disease organisms as Uriah Heep, the Charles Dickens character who was ingratiating when subordinate but overbearing when dominant. At the outset of the book, he considers each parasite to be a tiny Uriah Heep. So far so good, because what is best for individual parasites within a host should have important effects on the parasite's evolved characteristics. But then his anthropomorphism gives way to organismorphism, and he casts plagues as Uriah Heep as well.

This extension of the metaphor is not so good, because natural selection does not treat a population of plague organisms in a population of people (say, Europeans during the 15th century) as a single organism. Whether the pathogens that survive between outbreaks are par-

ticularly virulent or benign depends on the details of how they are transmitted, as well as on the competition among different strains of the pathogen within the same host. By treating populations and species as individual organisms, Wills fails to distinguish the population- and species-level effects of selection from the process of natural selection.

It is in this context that a central problem with Wills's characterization of plagues becomes apparent. He defines a plague as something that eventually subsides—there is no room in his description for diseases that kill in large numbers for millennia. According to the modern theory, however, we should find perpetual plagues when the particulars of transmission favor both high virulence and high frequency of infection. One of the most damaging plagues of human history fits this picture. In areas where mosquito density is perpetually high, *Plasmodium falciparum* malaria kills countless thousands of people annually, year after year. Its effects are more sporadic where mosquito density sometimes falls.

When highly virulent pathogens cannot infect hosts frequently over ecological timescales, plagues wax and wane. When highly virulent pathogens can continue to infect hosts often, we see dangerous endemic disease. Because both outcomes are feasible, we need to determine the conditions that favor the various outcomes rather than trying to fit all kinds of virulence into a transient plague model. *P. falciparum* contradicts the contention that "plagues are sporadic" and afflict us "because we have upset the balance of nature." If we are going to dismiss perpetually apparent diseases such as *P. falciparum* malaria as something other than plagues because they are not characterized by sporadic occurrences, then the phrase "plagues are sporadic" is tautological.

Wills argues that certain malaria parasites are severe because they have more recently entered humans from other species and therefore have had insufficient time to evolve toward benignity. Yet there is no reliable evidence that the most severe *Plasmodium* species—the previously mentioned *P. falciparum*—has been in humans for a shorter time than more benign species. Recent comparisons of *Plasmodia* DNA by A. A. Escalante and his colleagues at the University of California at Irvine, indicate that *P. falciparum* may have afflicted the human lineage since before humans were humans, perhaps even before humans and chimpanzees diverged.

Indeed, the *P. falciparum* lineage has proved its success by the fact that it is still here after thousands and probably even millions of years of battle with human and prehuman immune systems and more recently with human cultural interventions, all in the face of intense competition among *Plasmodia*. Wills has identified one way in which the more benign malaria parasites are so-

THE ILLUSTRATED PAGE

Nuclear Weapons of the United States: An Illustrated History
BY JAMES N. GIBSON
Schiffer Publishing, Atglen, Pa., 1996 ($49.95)

Over the half-century of the cold war, the U.S. produced some 97 types of nuclear weapons, 11 varieties of nuclear-capable bomber aircraft and five classes of ballistic-missile submarines. Now that arsenal is gradually being reduced, although nuclear weapons are still very much with us. This book consists of short descriptions of nearly every piece of U.S. atomic hardware, along with material on tactical aircraft, cruise missiles and antiaircraft, antimissile and antisubmarine weapons. It offers a chilling reminder of the days when a bomb test might rattle the windows in Las Vegas or vaporize an island in the central Pacific.

With its oversize format and high-quality paper, *Nuclear Weapons of the United States* is the sort of thing you might expect to find on Dr. Strangelove's coffee table. Author James N. Gibson's focus is more on delivery systems rather than the workings of the gadgets themselves; for example, there is no separate chapter on warheads. But many of the photographs are striking, a showcase of the perverse beauty of destructive technology. Among them is an image (*right*) taken of a test of two Talos supersonic surface-to-air missiles, each of which could carry a conventional or a five-kiloton nuclear warhead. *—Glenn Zorpette*

phisticated: their potential for dormancy. But he has glossed over the equally sophisticated characteristics of *P. falciparum*, such as its high rate of antigen switching (a phenomenon that lets it hide from the immune system) and adherence mechanisms that allow it to avoid destruction by the spleen. Such characteristics allow *P. falciparum* to be specialized for life in the virulent fast lane rather than for the more secure and sometimes moribund lifestyles of other *Plasmodia*, much to the dismay of its victims.

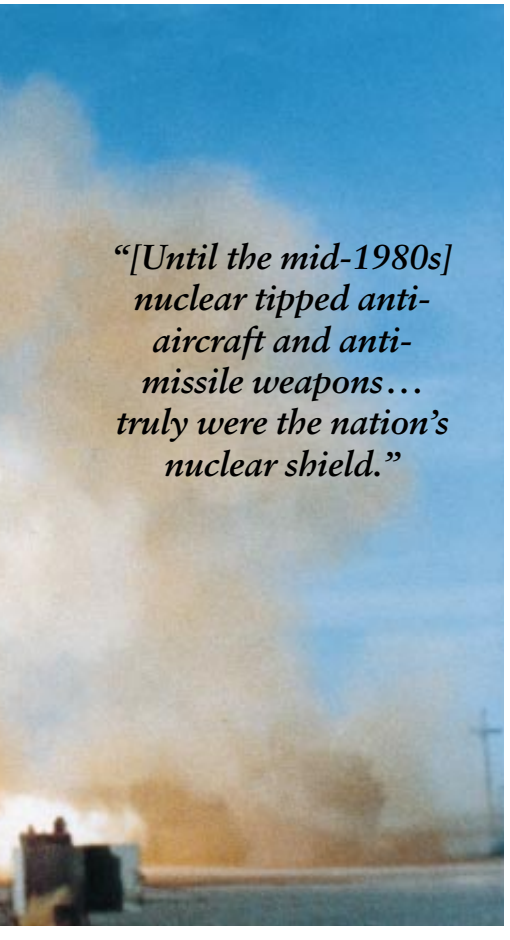
This fast-lane specialization accords with a high potential for mosquito-borne transmission throughout the year. Calculations indicate that *P. falciparum* spreads much more rapidly than other malaria species when first introduced into a population of susceptible people. If *P. falciparum* is often selected for effectiveness in relatively continuous transmission, competitive pressures may favor traits for getting in and out of the hosts quickly to avoid immune defenses.

The end result may be a well-adapted parasite that produces large numbers of progeny that do great amounts of damage to their human hosts and yet are transmissible from these hosts for only a short time because of host immunity. Such are the vagaries of evolution. A well-adapted parasite, like a well-adapted free-living organism, is not necessarily one that has the greatest longevity. Are aphids less "sophisticated" than 17-year cicadas? Cicadas appear to be more sophisticated at invoking dormancy as a survival strategy, whereas aphids seem to be more sophisticated at altering characteristics (such as sexual versus asexual reproduction) that allow effective short-term competitive success.

The author's characterization of cholera and typhoid organisms similarly illustrates this common confusion. Because typhoid organisms can generate carrier infections (in which the host maintains but does not transmit the pathogen) more frequently than cholera organisms can, Wills considers the former more sophisticated. Similarly, he concludes that the typhoid organisms in the temperate zone tend to be more sophisticated than those of the tropics. If the carrier state is such a sophisticated solution, then why does it occur in so few infected individuals?

Wills resurrects with approval Lewis Thomas's reassuring contention that "pathogenicity may be something more frightening to [the pathogens] than to us." Although it is essential to consider virulence from the pathogen's perspective, I think Thomas's quote clouds rather than clarifies. Natural selection does not tally the number of pathogens left to die in a dead person. Rather it tallies the copies of genes that make it to the succeeding generation after the immunological battle.

To appreciate this point, ask whether our body's immunological defenses are more frightening to us than they are to the pathogens. Billions of our own cells routinely die as a programmed part of our immune system's activities, but this death evolved because it fostered the propagation of our genes (including the ones responsible for that cell death). When pathogens are growing in a host, the death of the pathogens resulting from the exploitation of that host may likewise be favored.

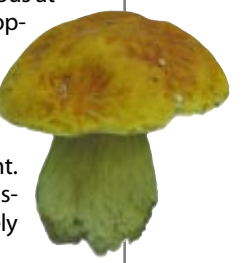


BENDIX

BRIEFLY NOTED

THE SUPERMEN: THE STORY OF SEYMOUR CRAY AND THE TECHNICAL WIZARDS BEHIND THE SUPERCOMPUTER, *by Charles J. Murray. John Wiley & Sons, New York, 1997 ($24.95).*
The world's fastest computer is now probably the Internet—or whatever subset of its connected machines can be harnessed to the same computing job. For the better part of 30 years, however, the speed title was held by whatever came out of the laboratories of the enigmatic genius Seymour Cray (who died last year in a traffic accident) and his colleagues. Although investors and customers grew unwilling to pay the price for those one-of-a-kind marvels, the ride was amazing while it lasted.

IN THE COMPANY OF MUSHROOMS, *by Elio Schaechter. Harvard University Press, Cambridge, Mass., 1997 ($24.95).*
Mushrooms are everywhere. They are beautiful, noisome, delicious or deadly, depending on how you choose them. Elio Schaechter has spent his career as a mycologist partaking of the phallic stinkhorn (delicious at the right stage in its development) and climbing out of bed at 6 A.M. to identify a poisonous amanita that had allegedly been ingested by a hospital patient. His love for his subject suffuses his writing in this relatively brief text.



LANDMARKS IN MECHANICAL ENGINEERING, *by the ASME International History and Heritage Committee. Purdue University Press, West Lafayette, Ind., 1996 ($39.95; paperbound, $19.95).*
The devices documented in this volume, ranging from a hand-carried outboard motor to mine elevators and hydraulic presses more than six stories tall, have shaped modern life so thoroughly that it is difficult to imagine where we would be without them. Yet the originals were often sold for scrap or today lie in pieces in corporate attics. This book from the American Society of Mechanical Engineers not only gives succinct descriptions but reveals where some of the surviving wonders of human invention can be seen and touched.

Continued on page 116

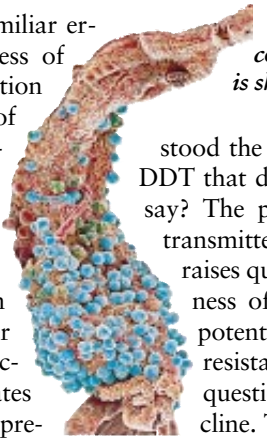
BRIEFLY NOTED

Continued from page 115

THE COMPLEAT COCKROACH, *by David George Gordon. Ten Speed Press, Berkeley, Calif., 1996 ($11.95).*
They are unpleasant, disease-ridden, prolific and infuriatingly hard to kill. David George Gordon tells more than most people will want to know about one of humanity's oldest pests. His lighthearted text is informative and enjoyable; it also teaches a great deal about insects in general.

SELF-SIMILAR MELODIES, *by Tom Johnson. Two-Eighteen Press (P.O. Box 218, Village Station, New York, NY 10014), 1996 ($24).*
An appreciation of the connection between mathematics and music dates back at least to Pythagoras, but rarely does one have a chance to explore that connection in such intimate detail as in these bracing pages. Composer Tom Johnson weaves back and forth between notes and numbers, delighting in patterns that lead up to self-similarity (the basis of fractal geometry). Some familiarity with both sides of the topic is a prerequisite; committed readers will, however, receive a unique aesthetic education.

HORN OF DARKNESS, *by Carol Cunningham and Joel Berger. Oxford University Press, New York, 1997 ($25).*
A short letter to *Nature* about the population biology of the rhinoceros in southern Africa was the fruit of a much larger story of three years of trekking through the Namibian desert, small child in tow. The authors, who are both academic and domestic collaborators, effectively recount a complicated personal and political story. Rhino "management" is such a controversial subject—even as poachers continue to erode the tiny remnants of wild populations—that useful data are sparse and hard to come by. Even harder when scientists whose work might cast doubt on official policy find themselves barred from the country, the sad end to this tale.



MOSQUITO GUT covered by malaria parasites is shown in this micrograph.

LONDON SCHOOL OF HYGIENE *SPL/Photo Researchers, Inc.*

Wills also makes some familiar errors in describing the process of evolution. He confuses evolution itself with the generation of variation on which natural selection acts and with the reduction in variation that may occur as a result of natural selection. When discussing the genetic changes in pathogen populations that occur over time in response to human activities, for instance, he states that these changes do not represent, "in most cases, a real evolutionary process."

This blurring of evolution with the generation of variation leads Wills to overlook some potentially powerful predictive aspects of current theory that operate at the interface of evolution and epidemiology. He writes, "It is not at all obvious why cholera strain 0139, which is apparently at least as vicious as the 01 strain that it is replacing, is spreading through eastern India and Bangladesh." But modern evolutionary principles actually make the reasons for this diffusion fairly obvious: wherever cholera can be rapidly transmitted through contaminated water supplies, virulent fast-lane strains will take advantage of this niche. For years, I had been predicting such a spread, and when the 0139 outbreak began, UNICEF's senior adviser, Jon Rohde, dashed me off a note from India. His opening sentence was, "You must be sitting there saying, 'I told you so.'"

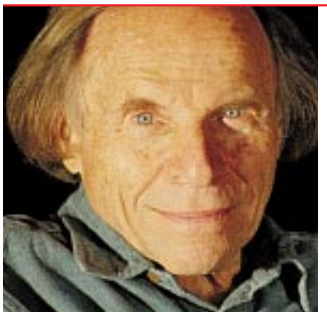
The power of the evolutionary approach to epidemiology lies in just such predictions. Particularly harmful diarrheal variants can be expected to expand their territory in areas where pathogens can be transmitted through the water supply. New mild variants are predicted to gain ground in areas that have safe drinking-water supplies. By providing uncontaminated drinking water, we should be able to prevent the spread of virulent competitors whether or not we have specifically identified them in the population.

Another shortcoming of *Yellow Fever Black Goddess* is that Wills sometimes compromises the rigor of his argument for dramatic effect. When discussing the 1994 outbreak of bubonic plague in India, for example, he contends, "No plague-like outbreak could have withstood the deluge of tetracycline and DDT that descended on it." Who can say? The potential for plague to be transmitted in the absence of fleas raises questions about the effectiveness of DDT. The demonstrated potential for pathogens to evolve resistance to antibiotics raises questions about that of tetracycline. This logical flaw is dangerous: reliance on massive interventions with antibiotics could leave us shorthanded if resistance evolves.

Some of the author's most innovative ideas appear in the book's final chapters, where he intertwines medicine with ecology. Focusing on the causes of biological diversity, he explains how parasites should suppress a host population in proportion to the commonness of the host, thus reining in host species that might otherwise dominate the ecological community. This effect, he argues, should also favor immunological diversity among humans. People with common immunological genotypes provide the biggest populations for pathogens to exploit, and so disease will eventually reduce their numbers as specialized pathogen variants arise. People with rarer immunological types will have a reproductive advantage, and so their population will increase, repeating the pattern of change indefinitely.

Although there is bathwater to be thrown out, there is also a vital baby that needs to be kept and fostered from Wills's book. With appropriate caveats, his highly accessible synthesis will allow scientists, nonscientists and, perhaps most important, young scientists in the making to see the potential for discoveries and insights in the fertile territory between the disciplinary boundaries. Like the tropical vines he describes, Wills's lively intertwining of health sciences, history, biology and socioeconomic considerations should help form a nexus for a greater diversity of ideas and insights.

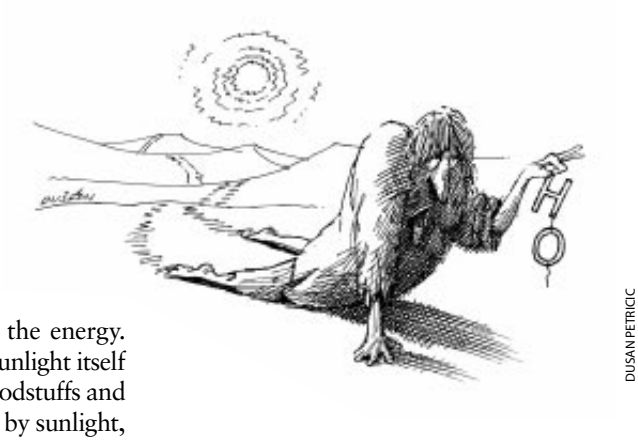
PAUL EWALD is a professor in the department of biology at Amherst College. He is the author of Evolution of Infectious Disease *(Oxford University Press, 1994).*



WONDERS

by Philip Morrison

HOH and Life Elsewhere



DUSAN PETRICIC

For once, write the familiar H_2O as HOH; it stands for the same molecular species, two hydrogen atoms linked to one oxygen. Although the symbol HOH is ramrod straight, the real molecule is a bent, symmetric clump of three near-spherical atoms. Nearly two thirds of all the atoms in an adult human are stacked in just this form. To be sure, water is but the stage and setting of life; life's actors are mainly carbonaceous molecules, more complex by far. We can now make out most of the molecular devices—those intricate electromechanical contrivances similar to ratchets, clamps, tubes and electrodes—at work in every cell.

Destructive collisions, whether with overenergetic photons or with fast-moving atoms, can sever and dismember big molecules. Too high temperature is fully hostile. Too low temperature does not damage the molecules, but their reactions slow or even cease. And all that watery machinery has to work in its watery ambience. Wherever water cannot persist, biomolecules cannot function. So liquid water marks the regime we know as hospitable.

The chemical bonds between atoms require intimate contact. But out in the darkness of our galaxy, a random hydrogen atom may collide with another only once every few thousand years. If on our planet the evolution of life were paced so slowly, we could not hope for wonders from such lonesome molecules. Life can evolve within its maximum allotted time—the life span of the stars—only where atoms and molecules fairly jostle one another. This state centers somewhere on our very own unit of density, that of liquid HOH.

Another necessity is a long-term supply of available energy to drive change; our sun offers us the light that leaks steadily out of the thermonuclear furnace in its core, burning hydrogen into helium. Photons transport the energy. We ourselves thrive not by sunlight itself but by proxy; our starchy foodstuffs and free oxygen are both woven by sunlight, from atoms of air and water on molecular looms in green plants.

Life makes one further claim on hydrogen that is much less of a commonplace. The linkages that unite molecules are strong and generally created by the transfer or sharing of electrons between bound atoms. But some type of chemical bond gentle enough to permit assembly and disassembly by the atomic environment without disruptive energies is also a necessity for the unceasing transformations of the life stream. That essential bond is provided not by electrons alone but by entire hydrogen atoms.

While a hydrogen atom is isolated, the single electron remains spherically distributed around the host proton. The bond that links the hydrogen atom to any other in a molecule is usually an example of electron sharing, the so-called covalent bond. But when the lone electron has been drawn into such a bond, the proton is no longer central. The H atom, now no simple sphere, exhibits a positive-charge excess on the side away from the covalent bond. Such an atom may form a second, much weaker bond, drawing a negatively charged atomic neighbor to the proton-favoring side.

No other atom is as abundant in our still-maturing cosmos as hydrogen.

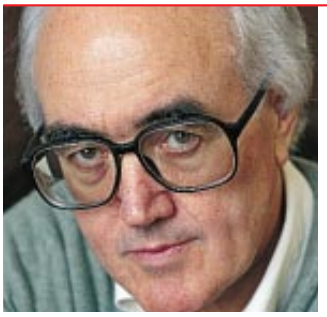
That link, mediated by both constituents of a hydrogen atom, is called a hydrogen bond. Longer and less intimate than the electron bonds, it binds with one tenth the energy. Most crucial biomolecules, even DNA itself, are replete with hydrogen bonds, thus offering the opportunity for moderate change. It is probable that only hydrogen-rich fluid solvents, where hydrogen bonds can readily form, are a medium fit for ever changing life.

HOH has no peer as a hydrogenous solvent. Molecules that make up the alcohols, for example, are more complicated. Other candidates, such as cold liquid ammonia (NH_3) and hydrofluoric acid (HF), are cosmic rarities. Insistent life is where we will find it, but hydrogen bonds appear for now as indispensable. For this reason, too, the medium of choice is liquid HOH. We know its unusual secondary properties as well, like expansion on freezing, but these seem minor. It is on the supply side that HOH is exceptional.

No other atom is as abundant in our still-maturing cosmos as the lightest and simplest of all, hydrogen. (Excepting dark matter; it is gravitationally dominant, but because it is not atomic, it is otherwise uninvolved.) The second most plentiful atom is the inert helium atom; less common than hydrogen by about 10-fold, it forms no easy chemical links. Those two atoms, H and He, are the primary material bequest of the old universe of hot and uniform gas before the stars formed. No other atomic species now numbers as many as one in 1,000 of the hydrogen population. All the others, common or rare, are products of some complex alchemy in the stars, whence some atoms are driven explosively into space, as others drift slowly out.

The third most common atom of all is the oxygen we breathe, the sun's legacy from stars gone by, whose debris of dust and gas was gathered up once into the

Continued on page 119



CONNECTIONS

by James Burke

Stones and Bones

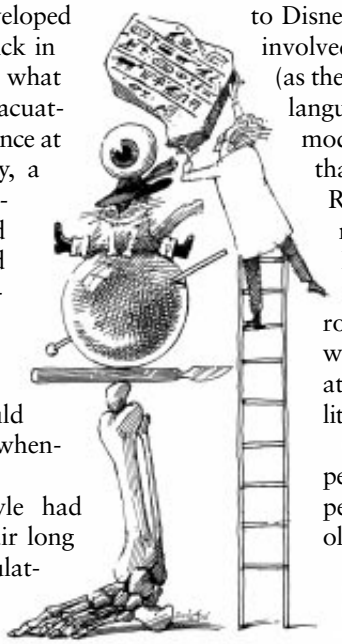
Things went internationally ballistic when the author of Mother Goose *said Plato was boring.*

I was in the British Museum last week, working in the great Reading Room and having a silent lament at its imminent demise. In 1998 we must move to new premises, whose layout is sure not to encourage octogenarian emeriti (like the one sitting next to me) to snore in that endearing way so familiar to users of the room. It was an Italian who did us all the favor. Antonio Panizzi, a Keeper of Books at the BM in 1831, got the Reading Room built.

At the time, plans were still embroiled in various scams associated with the lottery that was supposed to pay for setting up the museum in the first place. It must be said, however, that whatever money ended up in somebody's pocket, what remained was enough to do the job, and the museum had been opened in 1759, thanks above all to the single-minded obstreperousness of Sir Hans Sloane (inventor of hot chocolate). When he died in 1753, Sloane had the greatest collection of "curiosities" anybody had ever seen, and he left it to the nation on condition there was a museum to house it, plus £20,000 for his heirs. Hence the lottery and in turn the British Museum.

A buddy of Sloane's in all this was a surgeon called William Cheselden. He was famous for removing gallstones in 54 seconds (well, without anesthetic it hurt), for being the queen's physician and for knowing Newton well. In 1733 he had produced a massive book on bones (*Osteographia*), which included illustrations done with a camera obscura (a mirror reflects the pinhole image of an object onto a sheet of paper where it can be faithfully copied). A similar magnum opus had been done, but with *no* pics, a few years earlier, in Scotland. Point being, the Scots author had been Cheselden's student in London and heard his mentor's book would have pictures. Wise fellow didn't want to steal the thunder of someone who was also a big cheese and a friend of the great and good. Mind you, Alex Monro, the anatomist-author in question, was no shrinking violet himself. In 1726 he was partly responsible for getting Edinburgh its first medical school. And for setting a new fashion in body snatching, done by students who wanted to ensure they got their money's worth at his dissection classes. Monro finally calmed public outrage by doing a deal with the local law, which involved some dubious arrangements to do with regular delivery of the corpses of recently executed criminals.

Another hotshot Monro had learned his stuff from, down south in London, had been Francis Hauksbee, the man who invented the amazing "influence machine." This consisted of little more than a large glass globe on a spindle. If you cranked a handle (that turned the spindle and spun the globe), then held a hand against the globe, you got an electrified hand. The mystery influence also migrated down threads and attracted stuff like lint and feathers. Hauksbee had developed this spinning-globe trick in an effort to find out what you could do with evacuated vessels, because science at the time was, literally, a lot of fuss about nothing. Everybody wanted to see what you could do with a vacuum, especially given that Hauksbee's boss, Robert Boyle, had come up with a pump that would make you a vacuum whenever you wanted one.



By that time, Boyle had played around with air long enough to have formulated the law (about the behavior of a gas at constant temperature) known as Boyle's Law, or in French, *la loi Mariotte*. You'll notice the strange way the French have of saying "Boyle." That's because Edme Mariotte did it at roughly the same time as Boyle (or before, if you're French). Fact is, in 1679 Mariotte relied heavily on Boyle's 1662 work but never mentioned him. Mariotte spent most of his life doing similarly similar things. Kind of his blind spot. Ironic that one of the few *original* things Mariotte did do was to discover the blind spot, that place on the eye's retina insensitive to light. Anyway, on one occasion he "confirmed" the work of a Pierre Perrault, who had recently measured the rainfall in the Paris Basin and concluded that the Seine (and rivers in general) got most of its water from the sky.

In 1697 Perrault's brother, Charles, who was a dashing type and sort of French minister of culture, compiled a collection of folk stories called *Tales of Mother Goose* (which included the characters Little Red Riding Hood, Puss in Boots, Cinderella and others, all handed to Disney on a plate). He also got involved in near-fatal arguments (as the French will do about their language) over whether or not modern writers were better than the ancient Greeks and Romans. Things went internationally ballistic when Perrault said Plato was boring. You could hear the row as far away as Ireland, where it engaged the passing attention of one of English literature's more acerbic wits.

Jonathan Swift's life is a perfect example of what happens when you ignore that old adage: "Be nice to people on the way up; you may need them on the way down." If it hadn't been

DUSAN PETRICIC

for care packages from the Berkeley family (to whom, uniquely, he had been nice), he might have starved. And never have become pals with son George, who later became a bishop and a power in American education (the place in California is named after him) and who published, in 1709, an essay on vision. In which he offered the revolutionary idea that what you see is not what you get. That the brain interprets what the eye sees by associating the signals received by the sense organs with things already known in the brain. This brand of associationism eventually excited one of those overachievers you love to hate. This medical type could read at age two and had read the Bible twice by age four. By age 20 he knew Hebrew, Arabic, Persian, Turkish and seven more. Get the point?

So it's not surprising that by 1801 (age 28), Thomas Young was already professor of natural philosophy at the Royal Institution and lecturing only on acoustics, optics, gravitation, astronomy, tides, the nature of heat, electricity, climate, animal life, vegetation, cohesion and capillary attraction of liquid, the hydrodynamics of reservoirs, canals and harbors, techniques of measurement, common forms of air and water pumps, and new ideas on energy. Had enough?

Young also published a new theory about how light was probably a wave, did the famous experiment in which he sent light through adjacent holes to produce the now familiar interference patterns and announced that the retina responded to all colors as variable amounts of the three primaries. Then, ho hum, he turned to (and helped to crack the problem of) hieroglyphics. He had a relatively easy time, thanks to a sample of hieroglyphics the Brits had snitched from Napoléon when they drove him out of Egypt, in 1799. What made this object particularly useful was that the hieroglyphic text was carved right next to its Greek equivalent. So he was halfway there (*of course,* he also knew Greek).

Funnily enough, when I get bored with snoring scholars or the indecipherable hieroglyphs in the scientific monographs I wade through in my job, I can wander out and enjoy looking at how much easier it was for Thomas Young. Because straight out the library door and at the second right is what he used to crack the code: the Rosetta Stone. SA

Wonders, *continued from page 117*

forming sun. Our sun is no chemist—it has few molecules because its forte is nuclear fusion. In the hot plasma of the interior of the sun and every common star—the seat of almost nine tenths of the atomic matter in the galaxy—only fragmented atoms and their freed electrons endure. Molecules are all but absent.

The unfailingly original Sir Fred Hoyle once imagined for his novel *The Black Cloud* a whole new unknown branch of life: life whose complexity unfolded out of changes inside stars, rather than among molecules. That prodigious organism had a mind structured by repeated but fleeting pulses and waves moving within starry plasma.

Our favored model for life is much less imaginative. We see life as carbonaceous molecular structures, in HOH solvent, open to change. The rocky earth, dense and partly solid, is atomically mostly oxygen, two atoms out of every three. Some rocky moons of the planets are similar. So with reason we look for hydrogenic life on a small oxygenic planet or moon near a lasting hydrogenic star, never too close or far from star glow.

Whence come the molecules themselves? Most are related to star birth. Chief among them is HH, or H_2, hydrogen gas. About a tenth of the hydrogen atoms of the galaxy are paired off just so. A birthing star draws some local space mixture of gases—mostly molecular hydrogen—to spin down into its surface. Much of that flow will be reejected at the poles. These little known supersonic flows are hot; like flames on Earth, they can synthesize a range of molecules from the mix, soon to discharge them. In such places we have found gassy molecular disks around many young stars. Might the end point sometimes be a single star with its planetary cortege, a seat of watery life, a true sun?

How should we look for nearby life? We live on one wet, animated surface, and now we drill life up from warm rock layers as well. Maybe some form survives within the icy shell that encrusts Europa, a tide-warmed moon of big Jupiter? Meteorite samples tease us with hints of past life on Mars; perhaps riverbanks were common there long ago. The microwaves may yet bring us a fuller report from unseen observers far away. Till then, sun-seeking astronomers, *cherchez l'eau.* SA

ANTISHOPLIFTING LABELS

by Joseph Ryan, Jr.

Plastic labels attached to virtually any consumer product are tiny sentinels in retailers' unceasing efforts to reduce the nation's $10-billion bill for shoplifting. Originally the idea of a store manager who could not outrun a shoplifter, antishoplifting labels attached to ordinary store goods, from book spines to meat packages, serve as marks of retail ownership.

The principal components of these antishoplifting systems are the labels themselves and a transmitter and receiver rigged to an alarm that creates an "interrogation zone" at the store's entrance or at the checkout aisle. In many early-generation systems still in use, the labels contain semiconductor diodes that modulate a microwave signal from the transmitter, then "answer back" to a receiver, sounding an alarm. Shoplifters can defeat these units by placing store merchandise in foil-lined coats or bags. Shoppers, moreover, can often set off false alarms when carrying consumer electronics into the store that contain components that mimic the transmission characteristics of the diode labels.

A new generation of labels based on "acoustomagnetic" principles has been designed to operate at a carefully chosen narrow band of low radio frequencies. The selected frequencies minimize any interference from other electronics. In these systems the labels contain a resonator made of an amorphous magnetic metal of a precise length and thickness, which is excited by low-frequency radio signals generated by a transmitter at the store's entrance. The resonator vibrates at the same frequency, thus transmitting an identical signal. A receiver at the entrance that detects the resonator's signal then initiates an alarm.

Similar to the low radio frequencies used by the U.S. Navy to communicate with underwater submarines, the electromagnetic fields from the transmitter can penetrate the foil shielding inside a coat pocket that might hide a labeled item. But the labels, if not enclosed in hard plastic casing, can still be disabled if they are bent or removed by a shoplifter, one reason they are beginning to be affixed inside the packaging or the product, whether it be a CD jewel box or a pair of blue jeans.

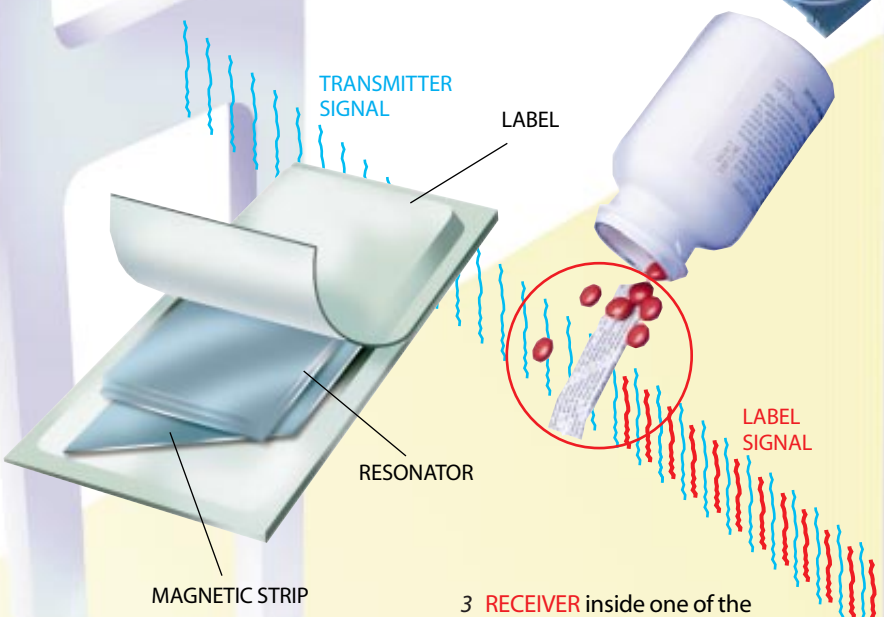
Antishoplifting labels might find other uses as well. Some items may soon be able to hold tiny "smart" chips that can be read and written to by radio waves, so that they can carry records of when and where an item was built and purchased. If you bought a pair of sneakers, your receipt would be in your shoes if you ever wanted to return them.

JOSEPH RYAN, JR., is vice president for global source tagging at Sensormatic Electronics Corporation in Boca Raton, Fla.

STORE PRODUCTS can carry an antishoplifting label (*in red circle*) that has been sewn into fabrics, affixed to the packaging, or sanitary-wrapped and placed among the contents.

1 PEDESTAL that brackets a store entrance or a checkout aisle contains a unit that transmits 58-kilohertz low radio-frequency pulses (*blue pulses*).

2 RESONATOR inside a product label passing through the pedestals begins to vibrate at 58 kilohertz, identical to the transmitted frequency.



3 RECEIVER inside one of the pedestals turns on during the 11 milliseconds between each transmitted pulse so that it can pick up the identical signal (*red pulses*) emitted by the label. If it receives a signal at least four times, it sets off an alarm.

4 MAGNETIZED STRIP adjacent to the resonator ensures that the oscillations of the resonator remain precisely at 58 kilohertz. A scanning device at the checkout counter can turn off a label when the merchandise is sold by demagnetizing the strip or by altering its magnetic properties. The resonator then either will not vibrate or will do so at the wrong frequency.

LAURIE GRACE (*drawing*); SENSORMATIC (*photographs*)