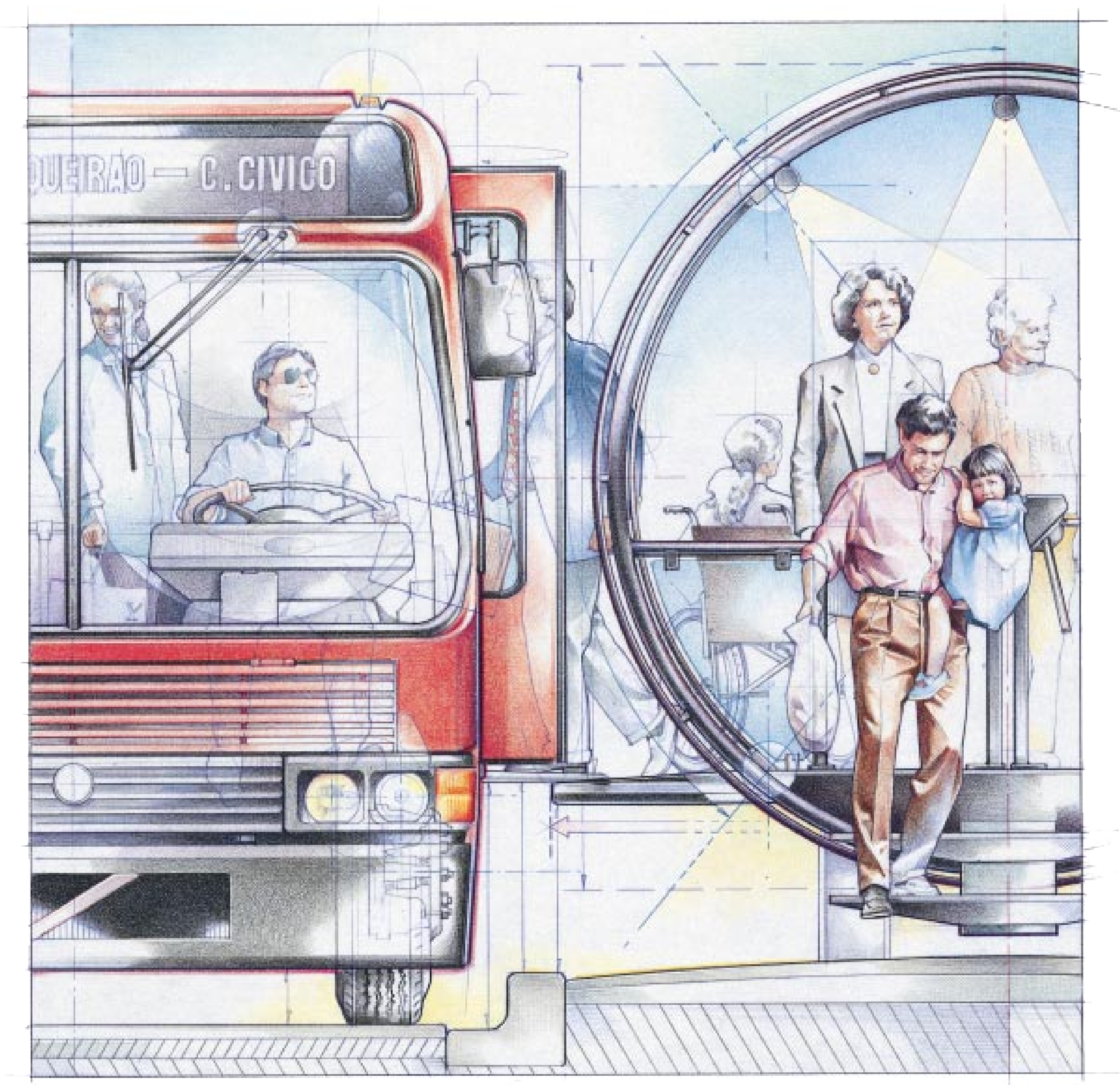
SCIENTIFIC AMERICAN

MARCH 1996
$4.95

Revealed: spy photo secrets.

Gene-testing nightmares.

Dangerous comets and asteroids.



Designing mass transit *that works puts a city on the road to success.*

SCIENTIFIC AMERICAN

March 1996 Volume 274 Number 3

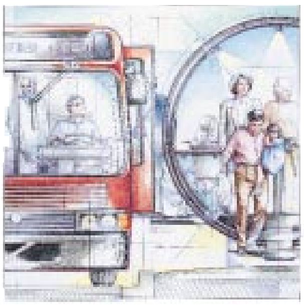
Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111. Copyright © 1996 by Scientific American, Inc. All rights reserved. No part of this issue may be reproduced by any mechanical, photographic or electronic process, or in the form of a phonographic recording, nor may it be stored in a retrieval system, transmitted or otherwise copied for public or private use without written permission of the publisher. Second-class postage paid at New York, N.Y., and at additional mailing offices. Canada Post International Publications Mail (Canadian Distribution) Sales Agreement No. 242764. Canadian GST No. R 127387652; QST No. Q1015332537. Subscription rates: one year $36 (outside U.S. and possessions add $11 per year for postage). 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 SCAinquiry@aol.com. **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 247-7631.**

DEPARTMENTS

Letter from the Editor

This issue, *Scientific American* runs the gamut on technology. We open with an article on how mundane, low technology can still have a terrific positive impact on a community. We close with a report on how one of the hottest high-tech areas can cause new headaches for society, even when deployed with the best of motives. Together these pieces make the point that technology is only as good or bad as what you do with it.

"Low technology," in the first case, really means first-rate civil engineering. Buses, artificial lakes and efficient roadways aren't glamorous. They don't have the show-biz appeal of virtual-reality interfaces for the Internet, or robots performing surgery, or "stealth" aircraft. But as the Brazilian city of Curitiba discovered, and as Jonas Rabinovitch and Josef Leitman recount, beginning on page 46, these unglamorous creations have greatly improved the quality of life for its two million inhabitants. As in a fine antique watch, the gears of this city mesh together exactly right, thanks to smart urban planning.



COVER ART by Bruce Morser

Aspects of a Curitiba-style solution will strike many readers as suspect. Public transportation as the key to solving a city's ills is anathema to many Americans. (It is noteworthy, however, that car ownership is exceptionally common among the Curitibans—they just avoid using cars where they would be a hindrance.) Some may wonder whether the lessons of Curitiba hold much relevance for more mature metropolises. But even if Curitiba's methods cannot be directly cloned for Los Angeles or Paris, its example should spur inventive thinking.

Tim Beardsley's "Vital Data" (see page 100) is provocative, too, if grimmer. Work on mapping human DNA is paying off speedily in tests for defective genes. People have never before had such sophisticated tools for making informed choices about having healthy children and for anticipating the state of their own future health.

The cloud over this silver lining is that precious few physicians, let alone members of the public, know what to do with all this genetic information. Progress in treatment lags behind that in diagnostics. Moreover, a genetic thumbs-up or thumbs-down is not the same as a diagnosis. Because misapplications of this knowledge are easy, some of society's responses are offsetting the marvelous potential good. Readers can become part of the solution by learning more about this technology and the ethical stakes.

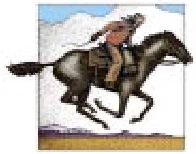
JOHN RENNIE, *Editor in Chief*

SCIENTIFIC AMERICAN®

Established 1845

SCIENTIFIC AMERICAN, INC.
415 Madison Avenue,
New York, NY 10017-1111

  PRINTED IN U.S.A.



LETTERS TO THE EDITORS

Get Smarty

Thanks for John Horgan's thought-provoking news story on the long-term rise in IQ scores, "Get Smart, Take a Test" ["Science and the Citizen," SCIENTIFIC AMERICAN, November 1995]. Could it be that the modern person uses a different aspect of intelligence than his counterpart of a century ago? And if the average person today utilizes left-brain processes more effectively, is it possible that other, more right-brained forms of intelligence are underdeveloped? Compare the rich verbal expressions of 19th-century writers and average citizens to works of present-day people. I'm reminded of Picasso's remark as he emerged from viewing the cave paintings in Lascaux, France: "We have invented nothing!"

JEANNE ROBERTSON
St. Louis, Mo.

The only IQ test I ever took was a Stanford-Binet, back in high school in 1947. It had a maximum score of 140. Just a decade later my daughter was given her first IQ test, in elementary school. Her tests were apparently open-ended, and she consistently scored some 25 points higher than I. She is now in her forties, and my personal, lifetime assessment is that her actual IQ is little, if any, higher than mine. Early on, nobody expected such high scores, so tests were not designed to be open-ended in their scoring system. Is this an explanation for the gradual rise in IQ tests over the years?

OWEN W. DYKEMA
West Hills, Calif.

Dawkins's DNA Denied?

Richard Dawkins's article "God's Utility Function" [SCIENTIFIC AMERICAN, November 1995] was full of errors of logic. He chastises those of us who would assume that living organisms have some inherent purpose or reason for being. Yet he argues that the basis of the grand scheme of life—in other words, its sole purpose—is to protect and pass on genetic material. The argument becomes almost comical when the author states that "the true utility function of life, that which is being maximized in the natural world, is DNA survival." It is as if he assumes that substituting "utility function" for "purpose" somehow makes his argument more valid.

LAWRENCE P. REYNOLDS
North Dakota State University

Reading the popular writings of neo-Darwinians like Dawkins sometimes makes me uncomfortable because they seem in danger of being hijacked by their own metaphors. While denying supernatural design, they teeter on the brink of attributing some pretty malevolent characteristics to nature. Some phrases in his article: "Nature is not cruel, only piteously indifferent…simply callous," and my personal favorite, "Genes don't care about suffering, because they don't care about anything." As if they could!

T. MICHAEL MCNULTY
Marquette University

If the purpose of DNA is to perpetuate life and thus itself, then Dawkins's argument is circular and unsatisfying. What is the purpose of a V-8 engine? To make money for General Motors! That seems to be the level of his argument.

TOM SALES
Somerset, N.J.

From the undisputed fact that the structure of DNA explains much, one cannot logically conclude it explains all, as Dawkins proposes. A key question is to find the limits of its influence. To many natural scientists, it is evident that matters of ethics and aesthetics in a human society do not dance to the music of the double helix.

ALWYN SCOTT
University of Arizona

Open Seas

Advances in technology are not the problem that Carl Safina makes them out to be in his article "The World's Imperiled Fish" [SCIENTIFIC AMERICAN, November 1995]. In other industries, technological advances make goods and services safer, cheaper and more plentiful. In the oceans, they do the opposite, because without ownership there are no rewards for taking care of a resource. Only when fishermen have a vested interest in the health of their fishery, and the ability to exclude those who do not, will they turn to technologies such as sonar and satellite systems for protecting, not exploiting, resources.

MICHAEL DE ALESSI
Competitive Enterprise Institute
Washington, D.C.

Cold Snap

I understand from Wallace S. Broecker's article "Chaotic Climate" [SCIENTIFIC AMERICAN, November 1995] that the melting of the polar ice cap could alter the "Atlantic conveyor," causing a major drop in temperatures in northern Europe. It seems ironic that global warming could in fact lead to cooling, something European antienvironmentalists should perhaps be informed of before a temperature downturn has them up in arms.

A.T.W. BEARDON
London, England

Analyzing Aliens

Robert Sheaffer's book review ["Truth Abducted," "Reviews and Commentaries," SCIENTIFIC AMERICAN, November 1995] raises the glaring question of why so many people believe in the veracity of reports of alien abductions. These stories violate basic laws of nature. They also, however, correspond with common unconscious fantasies from early childhood, when we were all relatively helpless and when, in contrast, our parents were perceived as all-wise and all-powerful. Such fantasies are clear examples of basic psychoanalytic concepts first elucidated 100 years ago by Sigmund Freud. The increasing belief in abduction and other bits of magic is, I believe, a sign of the bewilderment of our general population with the complexities of modern life.

HENRY KAMINER
Tenafly, N.J.

Letters may be edited for length and clarity. Because of the volume of mail, we cannot answer all correspondence.

50, 100 AND 150 YEARS AGO

MARCH 1946

"The problem of crew comfort on the control-decks of long-range aircraft is basic to the full realization of air transport's potential value. This problem was once considered to be satisfactorily solved when the pilot had a comfortable cockpit in which to sit, but this now appears as a serious misconception. Many of the studies of crew comfort hinge on biomechanics—the combined study of biology and mechanics. Some developments, such as the pressurized cabin, high-altitude flying suits, and improved food for both pilots and passengers, represent a tremendous amount of research work on the part of highly specialized medical men as well as aeronautical engineers."

"A millionth of a second X-ray machine, designed originally for basic research, is moving straight into the practical end of factory operation. This device can look through an inch of steel at the fastest moving mechanisms ever built, and produce pictures which tell what each hidden machine part is doing. Some smart shop is going to obtain a worthwhile cost advantage over its competitors when it X-rays the operation of metal-cutting tools on a high-speed lathe working on one of the hard-to-machine alloy steels. The X-ray can look through floods of cutting oil to see how metal is cut under actual operating conditions and can determine machineability of any lot of steel on the first few turns of a lathe spindle."

"Selective recovery of valuable materials is now possible with synthetic organic ion exchange resins. When a solution of electrolytes is passed through one of a variety of resins, it can absorb certain ions. Gold and platinum can be recovered by converting them to complex acids which can be absorbed by the proper resin. High quality pectin can be prepared from grapefruit hulls, when resin is added to a slurry of the rind in water and is later removed by centrifuging. In another commercial process, part of the calcium is removed from milk to make it more digestible for infants."

SCIENTIFIC AMERICAN

MARCH 1896

"Development and manufacture of the typewriter has grown into an industry of large proportions within a comparatively short space of time. This most useful, we might say indispensable, invention, with its busy 'click,' which was at first regarded as nothing more than an interesting toy, now gives employment to many thousands of operatives, and entails a heavy investment in capital in numerous large and thoroughly equipped factories."

Otto Lilienthal, whose work later influenced the Wright brothers, writes for Scientific American: "Formerly men sought to construct flying machines in a complete form, but our technical knowledge and practical experiences were by far insufficient to overcome a mechanical task of such magnitude without more preliminaries. For this purpose I have employed a sailing apparatus very like the outspread pinions of a soaring bird. It consists of a wooden frame covered with shirting (cotton twill). The frame is taken hold of by the hands, the arms resting between cushions. The legs remain free for running and jumping. The principal difficulty is the launching into the air, and that will always necessitate special preparations. As long as the commotion of the air is but slight, one does not require much practice to soar quite long distances. The danger is easily avoided when one practices in a reasonable way, as I myself have made thousands of experiments within the last five years." *Editor's note: Ironically, Lilienthal died in August 1896, after his glider crashed at Stölln, Germany.*

Otto Lilienthal in flight

"The theory that the two cerebral hemispheres are capable, to some extent, of independent activity, has been evoked to account for those strange cases in which an individual appears to possess two states of consciousness, such cases as afford the basis of fact for Robert Louis Stevenson's weird romance of 'Dr. Jekyll and Mr. Hyde.' Dr. Lewis C. Bruce records a case which is strongly in favor of the double brain theory. An inmate of the Derby Borough Asylum was a Welshman by birth, and a sailor by occupation. His mental characteristics had different stages at different times. In an intermediate stage he was ambidextrous, and spoke a mixture of English and Welsh, understanding both languages; but he was right handed while in the English stage and left handed in the Welsh stage."

MARCH 1846

"A wonderful account is given of the discovery of a monstrous wild man, in the swamps about the Arkansas and Missouri line. His track is said to measure 22 inches; and his toes as long as a common man's fingers. We are of the opinion that either the 'wild man,' or the man who raised the story, is a *great monkey."*

"Lumbering is the business—almost the only business—of Bangor, Maine, and the business is immense, with mills that contain 187 saws for the cutting of coarse lumber within 12 miles. One would think that such an everlasting and universal slashing as is going on in the woods north of Bangor, would very soon exhaust all the pine timber there is; but we are told there is no danger of this for a great many years to come."

"Settlers on the Missouri River have evinced serious alarm on the discovery that beavers have built their dams several feet higher than usual. This is regarded as an omen of an unprecedented freshet on that river."

Regulating the Body Business

The future is not what it might have been

More than 30 years ago the science-fiction writer Larry Niven envisioned a world in which organ transplants were common. To ensure a continuing supply of organs for the public at large, draconian laws mandated capital punishment for a host of offenses, and shadowy "organ-leggers" plucked victims off the street to extend the lives of the rich and remorseless.

K. VENKATESH Courtesy of *India Today*

SURGICAL SCARS on these Indian villagers may be marks of organ theft. Before laws forbade kidney sales, some physicians reportedly took organs without consent.

The world has not developed quite along the lines that Niven foresaw. Rumors of organ-theft syndicates appear to be pure urban legend. Then again, China does harvest the bodies of executed prisoners, and Austria has instituted "presumed consent" rules that permit doctors to remove organs from brain-dead patients unless specifically forbidden to do so beforehand.

There is also the question of body parts for sale. Until last year, $30,000 in India could buy a new kidney from doctors who paid a living donor less than $1,000 for taking part in the operation. Finally, prompted by general resentment that India was serving as a spare-parts repository for richer nations as well as by reports of shady practitioners who cut costs by stealing organs rather than buying them, the Indian legislature passed laws banning organ sales.

In the aftermath of the crackdown, doctors say the number of transplants performed throughout the country has declined substantially. Rishi Raj Kishore, assistant director general of the Indian health service, predicts, however, that rates will probably begin increasing again soon. At the same time that it outlawed the body trade, Kishore notes, India gave legal sanction to the concept of "brain death"—the notion that someone is no longer alive once brain function has ceased, even though the heart and other organs continue to work.

The laws of most Asian countries consider a body dead, and thus eligible for harvesting of organs, only after the heart has stopped beating. But more than a few minutes of stopped circulation at body temperature renders organs useless for transplantation.

Some countries, such as Japan, are working to change laws and attitudes to make donations possible, but regional disparities in wealth present serious obstacles. In 1994 a Japanese company was forced to abandon a plan for harvesting kidneys from Philippine donors, and a Japanese government plan for an international organ-transplant network found little support in neighboring countries.

There are enough brain-dead patients in India to supply the country's organ needs, according to Kishore—it is just a matter of convincing families to permit donations and of rethinking the country's transplant infrastructure to make the best use of them. China, in contrast, seems to be much closer to a realization of Niven's vision: the country reportedly relies almost entirely on capital punishment as a source of organs. Although the official numbers are a state secret, human-rights organizations estimate that China executes somewhere between 3,000 and 20,000 prisoners every year and harvests organs from at least 2,000. Government officials have asserted that prisoners agree to the harvesting, but critics argue that meaningful consent is impossible under the conditions that precede an execution. (Indeed, the U.S. has forbidden such transplants from executed prisoners, no matter how vigorously prisoners offer.)

Robin Munro of Human Rights Watch/Asia says that demand for organs does not appear to be driving the increase in capital punishment in China (as some observers have suggested). Instead the ready supply of relatively healthy cadavers appears to be prompting market development. News accounts have noted that transplant operations in China are common immediately after major holidays, when the government often conducts large numbers of executions.

Munro and other sources say that prison officials are allegedly "botching" executions so that prisoners die slowly, giving transplant teams more time to get to their organs. Some districts have also apparently replaced the legally required bullet to the back of the head with a lethal injection to prolong circulation and thus facilitate harvesting.

Stories of this kind horrify transplant specialists elsewhere. In the U.S. and other Western nations, there are concerns that government or private involvement in organ sales anywhere taints the entire global enterprise, leading to secretiveness that itself engenders suspicion. David Rothman of Columbia Presbyterian Hospital, who organized a conference on international organ trafficking in Bellagio, Italy, in September 1995, refused to discuss the proceedings or even release the names of any of the participants. And Paul Terasaki of the University of California at Los Angeles, who keeps statistics on U.S. transplants, counseled SCIENTIFIC AMERICAN to consider the entire notion of selling body parts a myth.

Indeed, the United Network for Organ Sharing (UNOS), which coordinates all transplants in the U.S., extracted public apologies from television soap operas that ran episodes with organ-sale premises. There is now a standing Hollywood committee to guard against the recurrence of similar innuendoes.

Nicholas Halasz, chair of the organ network's ethics committee, reports that he and his colleagues have even investigated transplants involving celebrities—notably baseball star Mickey Mantle, who died last year shortly after receiving a new liver—to ensure that no undue influence accompanied the rapid fulfillment of their needs. Halasz explains that UNOS rules give top priority to the sickest patients, but the organization must trust physicians not to exaggerate their patients' conditions.

Any signs of impropriety could cause people to rethink their signatures on donor cards or families to withhold consent for the removal of organs from brain-dead relatives, Halasz says. The same reasoning, he notes, has led many transplant doctors to oppose any compensation—even funeral expenses—to the families of organ donors.

Ironically, even in countries that rely entirely on voluntary donations, transplant donors tend to be poorer than the recipients. Although physicians are making progress in harvesting organs from stroke patients and other older, brain-dead persons, the typical donor is still young and a victim of accidental or deliberate violence—which tends to strike the disadvantaged in disproportionate numbers.

Studies have shown that minorities (especially blacks) suffer at a higher rate than whites do from diseases, such as kidney failure, that new organs could alleviate. Yet minorities are less likely to be offered transplantation. In Niven's grim future, at least the state made organs available to all who needed them.

—Paul Wallich and Madhusree Mukerjee

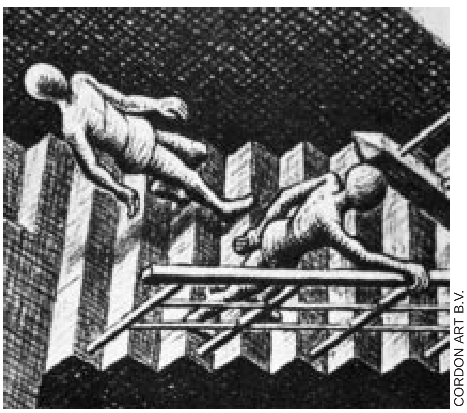
Additional reporting by Sanjay Kumar in New Delhi.

Escher for the Ear

When computer-generated tones are played repeatedly in certain sequences, they can appear to rise or descend endlessly in pitch. Other patterns of notes are heard to ascend by some people but to descend by others. Diana Deutsch, a psychologist at the University of California at San Diego, now reports that childhood plays a crucial role in how one perceives certain Escheresque melodies.

Using a computer, Deutsch constructed notes that lack a clear octave relation. For example, to make an ambiguous C note, she combines the harmonics of all C notes and manipulates their relative amplitudes (in essence, playing all six C notes on a keyboard simultaneously). As a result, a listener might be able to identify the note as C but remain unsure if it is middle C or the C an octave above or below.

Deutsch then paired each such ambiguous note with another exactly one half of an octave away (a musical distance called a tritone). For instance, subjects heard a C followed quickly by F sharp or an A sharp followed by E. The listeners were asked to judge whether a pair formed a rising sound or a descending one. Because octave information about the notes was eliminated, there was no objective answer to the question.



DETAIL FROM* RELATIVITY, *by M. C. Escher, visually parallels musical paradoxes.

The responses to this tritone paradox, as Deutsch calls it, depend on the area from which the listeners hail. Subjects from the south of England tended to hear a pair as ascending, whereas Californians heard it as descending, and vice versa. Tests of those with regional dialects in the U.S. produced similar variations.

To explain the finding, Deutsch hypothesized that people form a fixed, mental template that places musical notes in an octave in a circle, like the numbers on the face of a clock. Californians apparently fixed C in the upper half of the clock (between the nine and three o'clock positions); being half an octave away, F sharp must fall in the lower half for these listeners. Hence, they heard the C–F sharp pair descend. Britons, in contrast, place C in the lower positions of the clock and so heard the same pair rise.

Deutsch's latest results suggest that this template forms in childhood. When she tested the tritone paradox on parents and their children, the youngsters followed the perceptions of their mothers. That correlation persisted even if the mother grew up in a different locale: a California child heard what her British mother perceived, rather than hearing what other Californians heard. "This very strong correlate must reflect the fact that we are very attuned to the pitch of speech," says Deutsch, who will present her case at a meeting of the Acoustical Society of America in May.

The findings prove certain points in music theory but may have broader implications. Composers should know that "the music they are hearing may be perceived quite differently by other listeners," Deutsch observes. The paradoxes might also prove interesting in the study of neurological disorders. People suffering from certain types of brain damage often have flat speech intonations and may have unusual responses to the aural paradoxes, and manic-depressives might hear the tones in ways that reflect their moods. *—Philip Yam*

Samples of musical paradoxes appear on Scientific American'*s area on America Online. The illusions are available on an audio CD; for information, call 1-800-225-1228 or go to http://www.philomel.com on the Internet.*

Out of This World

Tracking the origin of cosmic rays

From outer space they come, striking the earth's atmosphere from all directions at nearly the speed of light. As they streak between the stars, they heat and alter the composition of giant gas clouds, subtly influencing the evolution of our entire galaxy.

They are cosmic rays, subatomic particles or atomic nuclei that carry at least a billion times the amount of energy in a photon of visible light. Ever since Austrian physicist Victor F. Hess discovered cosmic rays in 1912, astronomers have debated the origin of these enigmatic particles. Recent observations from *ASCA,* an astronomical satellite jointly operated by the U.S. and Japan, seem at last to have solved at least part of the puzzle.

Theorists had long suspected that supernova explosions—among the most violent events in our galaxy—could provide the jolt necessary to accelerate particles to cosmic-ray energies. The acceleration would come not from the explosion itself but from the resulting shock wave. Magnetic turbulence at the front of the shock creates a kind of magnetic "mirror" in which charged particles bounce back and forth, picking up energy on each pass. For the first time, *ASCA* has caught this process in the act.

A group led by Katsuji Koyama of Kyoto University trained *ASCA*'s sensitive x-ray detectors on the remains of a nearby supernova that exploded in the year 1006. Extreme heat from the explosion gives rise to an x-ray glow, which *ASCA* could easily pick up. At the edge of the supernova remnant, Koyama and his colleagues discovered another, distinctly different kind of x-ray emission—one that seems to come from high-speed electrons racing through a magnetic field. Koyama's group infers that those electrons, and presumably other, less readily detected particles, are being accelerated to enormous energies at the edge of the supernova shock. When particles like those reach the earth, they are seen as cosmic rays.

"Since we found cosmic-ray accelera-

FIELD NOTES

Plane Scary

Many viewers had to squint at the final credits to see how the eerily realistic zero-gravity scenes in the recent movie *Apollo 13* were filmed. Having no way to transport the cast to outer space, the moviemakers appealed to the next best thing: the National Aeronautics and Space Administration's Reduced-Gravity Office. Since the early days of the space program, that small arm of the agency has been lobbing a specially equipped KC-135A jetliner into the sky in such a way as to reproduce temporarily the weightlessness that astronauts feel in space. Curious to see how people with only ordinary amounts of the right stuff take to a taste of "zero g," I interviewed some recent passengers of NASA's KC-135A "Vomit Comet."

NASA

FLOATING ENGINEERS—Tibbitts, Richardson, Priebe and Christiansen* (clockwise from upper right)*—ride NASA's "Vomit Comet."

Scott F. Tibbitts, president of Starsys Research Corporation in Boulder, Colo., needed to test weightless operation of two devices his firm had developed for the upcoming *Pathfinder* and *Cassini* space probes. He decided to take advantage of a tool he had used once before—NASA's zero-gravity airplane.

Tibbitts and three co-workers (whom he describes as "people who, since they were little kids, had dreamed of being in space") dutifully passed through the preflight hoops required by NASA. They were required, for example, to sit in an altitude chamber and to discover the shock of explosive decompression and the drunken stupor that ensues when the air thins at great heights.

After they had properly dispensed with such mildly challenging formalities, the Starsys quartet arrived ready and eager in late 1995 at NASA's Ellington Field in Houston. "We were so excited, with silly little smiles on our faces," Tibbitts admits. Their tests of the two mechanisms (a paraffin-controlled valve and a spring-driven instrument cover) were to be conducted in parallel with the experiments of several other groups—investigations of weightless behavior that involved everything from supercooled liquid metals to human blood.

At one end of the plane's 60-foot-long, padded interior stood a big, red readout: the g-meter. During a typical three-hour flight, that display shifts 40 times from one g (normal flight) to two gs (as the plane noses up 45 degrees), then to a 25-second period of zero g as the jet coasts through a gently curved parabolic arc and back to two gs (when the craft pulls out of its dive).

Tibbitts explains how he reacted to the initial zero-g parabola when he first flew on the KC-135A in 1990. "First you think it's like a roller coaster, then like a big roller coaster, then you realize it's something new." Jason E. Priebe, a novice zero-g flyer in the Starsys group, was not sure what to expect, but he knew that two out of three people on board usually get sick. He and team member Mark T. Richardson reacted to their inaugural dose of zero gravity as they might to a good amusement park ride: "We started screaming."

Although Priebe says he "was shaking pretty badly after the first three or four parabolas," he and the three others from Starsys acclimated quickly enough. One of them, Scott S. Christiansen, even ventured to float to the cockpit to watch the scenery scroll upward as the plane tipped from steep climb to nosedive. Yet not one of the four men succumbed to the epidemic of motion sickness that swept through the plane, so the Starsys group conducted its experiments unimpeded. The zero-gravity trials confirmed that the valve worked flawlessly and found some unanticipated behavior in the spring-loaded cover.

Having completed the tests with time to spare, the engineers produced some other mechanical apparatus to examine in zero gravity—a slinky, a yo-yo and a paddleball. They also investigated the rotational dynamics of some not so rigid bodies. "I did a back flip at one point, and I wasn't sure when to come out of it," Priebe recounts. "I flipped over and was standing on the ceiling." So it would seem that for engineers such as those from Starsys, who want to ensure that a particular mechanism will work as expected in space, a KC-135A flight can be not only a great comfort but also a lot of good fun. —*David Schneider* tion under way in the remnant of supernova 1006, this process probably occurs in other young supernova remnants," notes Robert Petre of the Goddard Space Flight Center, one of Koyama's collaborators. But some rays are so potent that even supernova events probably cannot account for their existence. (The most extreme of these rays contain as much energy as a Nolan Ryan fastball—crammed into a single subatomic particle!)


NATIONAL OPTICAL ASTRONOMY OBSERVATORIES

EXPANDING REMNANTS of supernova explosions **(such as the Crab Nebula, above)** ***may be the birthplace of many cosmic rays.***

Many researchers have assumed that the high-energy cosmic rays must originate in even greater shocks—those surrounding active, or "exploding," galaxies. In a recent paper in *Science,* however, Günter Sigl of the University of Chicago and his colleagues suggest a more exotic possibility. Sigl's group analyzed data from the Fly's Eye detector in Utah and other, similar experiments that study the flash of light and spray of particles unleashed when cosmic rays collide with atoms in the earth's upper atmosphere. The researchers find an odd "gap" in the data: at progressively higher energies, the number of cosmic rays seems to trail off but then abruptly increases again.

No known process could produce such a gap, so why is it there? One possibility is that the highest-energy cosmic rays are the product of an entirely new, still hypothetical physical mechanism—the evaporation of cosmic strings, for instance, or the decay of proposed supermassive particles. On the other hand, the total number of high-energy cosmic rays detected is quite small, so the perceived gap "could be a statistical fluctuation," Sigl admits. "We don't have the data" to tell for sure, he laments.

Help may soon be on the way. Last November physicists from 19 countries committed themselves to building the Pierre Auger Cosmic Ray Laboratory, a $100-million detector that would far exceed the sensitivity of any existing device. Tentatively scheduled to begin operating at the beginning of the next century, the Auger Laboratory could quickly settle many current questions about cosmic rays. "It could confirm new physics, or it could rule it out," Sigl reflects. "Either way, it will be very interesting." —*Corey S. Powell*

Viral Tracers

Neuroscientists use viruses to map out pathways in the brain

Imagine trying to make sense of a railway map if none of the lines were labeled. It would be nearly impossible to know which trains ran between which towns. Neuroscientists long faced a similar problem: the chemicals they used to trace lines of communication between brain regions vanished after a single stop. "They only went from one station to the next," says Peter L. Strick of the Veterans Administration Medical Center in Syracuse, N.Y. Knowing but short stretches of certain tracks, he adds, made it exceedingly difficult to determine where any one train—or nerve signal—ultimately went.

Recently, though, Strick has turned to a new, more powerful technique, one that enlists itinerant viruses to chart brain circuits in monkeys. "The viruses move from one neuron to another, right on down the line," he notes. "Happily, there are strains of virus that do this by crossing over synaptic connections." These viruses cross in only one direction. A strain of the herpes simplex type I virus, for example, follows the flow of nerve impulses through neighboring cells: the virus particles pass down a neuron's axon, across a synapse, into another neuron, down its axon, over another synapse and so on. A different strain moves in the opposite direction.

Unlike conventional tracers, a little virus goes a long way. Because the strains are living, they replicate in every cell, thus increasing in number before each leg of the journey. "You get an on-line amplification of sorts of the tracer signal," Strick points out. "So we can see the signal more clearly than we ever could before." Already the method has revealed new facts about the cerebellum in primates. Traditionally, scientists believed that this structure integrated information from the cerebral cortex with sensory input from the muscles. It then presumably sent nerve signals back to other motor regions in the brain, enabling the body to perform skilled movements.

Strick, among others, has found that the cerebellum may also coordinate the movement of thoughts. Using viral tracers, he demonstrated that the cerebellum sent signals, via the thalamus, to regions in the cerebral cortex used solely for cognition, among them areas in the prefrontal cortex involved in short-term memory and decision making. "People proposed that the cerebellum had cognitive functions back in the 1980s," Strick says, "but I thought they were nuts. Now I'm a believer."

Most recently, he has discovered some far-reaching contacts that the basal ganglia make. These structures were also thought to preside primarily over motor functions. But viral tracers exposed output from them to sections of the temporal cortex responsible for visual tasks, such as recognizing objects. The finding, Strick suggests, could help explain why Parkinson's disease patients who take dopamine can experience visual hallucinations as a side effect. The dopamine given to humans may act on those same cells in the basal ganglia that in monkeys talk to visual areas in the temporal cortex.

Among other projects, Strick plans to determine whether the cerebellum plays a role in focusing attention. Damage to it may well provide the physical basis for the attentional deficits in autistic children. "I have spent some 30 years studying motor areas, but this technique is allowing me to look more globally at the circuits in the brain," Strick comments. In time, he adds, the viral tracing technique could elucidate some of the circuits that malfunction in a number of mental and neurological illnesses. —*Kristin Leutwyler*

Rain Forest Crunch

Amazonian forests may have been smaller in the last ice age

Going up that river was like travelling back to the earliest beginnings of the world, when vegetation rioted on the earth and the big trees were kings." Joseph Conrad's evocative portrayal of the Congo would seem to apply as well to the Amazon. That river travels across the South American continent from Peru to the Atlantic Ocean, cutting through nearly four million square kilometers of undisturbed woodlands. But is the Amazon rain forest truly a primeval jungle, a steamy, green mass that has endured for millions of years? Perhaps not, according to new results from the high Andes.

The current findings challenge a perception, which first emerged in the 1970s, that tropical climates remained virtually unchanged while the great ice sheets of North America and Europe waxed and waned through a series of Pleistocene ice ages. That view was based largely on a study of microscopic shells from the ocean floor. Analyses of the kinds of creatures that had thrived in tropical seas during glacial periods indicated that the earth's equatorial regions had kept close to their present-day temperatures.

But a growing body of evidence has been slowly eroding the notion of persistently balmy tropical climates. Early chinks appeared with studies of mountain glaciers; snow lines had been substantially lower during the ice ages, even at equatorial latitudes. Such observations created a conundrum for scientists: How could high altitudes have been colder while temperatures stayed almost fixed at sea level? Researchers began to suspect that equatorial climates might not be so simple. Still, no one anticipated what Lonnie G. Thompson and his colleagues found when they returned to Ohio State University with an ice core extracted from a Peruvian mountain glacier called Huascarán.

In core samples dating from the final grip of the Northern Hemisphere's ice sheets, Thompson and his co-workers discovered a bevy of dust particles that had settled on this peak. They reported last summer in *Science* that atmospheric dustiness was about 200 times higher during the last ice age than during modern times. Thompson maintains that the dusty ice is a relic of climate conditions that had prevailed in Amazonia (upwind of his coring site) about 15,000 years ago. "If you look out to the east from Huascarán, you're looking into the Amazon rain forest," Thompson explains. When he trains his mind's eye on the ancient Amazon basin, Thompson sees a region that was quite a bit drier and a rain forest that "was much smaller."

To buttress his interpretation of the dust, Thompson points to chemical evidence in the ice core. Dissolved nitrate (which he and his colleagues believe emanates from the rain forest) shows dramatically reduced levels for the last glacial interval. Nitrate concentrations increased only slowly after the dusty period ended, perhaps reflecting the span of years the trees needed to grow back. "When you combine that with the increase in dust, you almost have to believe that the rain forest was much more restricted than it is today."

This view clashes with some other recent evidence from the South American continent. Paul A. Colinvaux of the Smithsonian Tropical Research Institute in Panama and his colleagues study ancient pollen entrapped in lake sediments. They find that some lowland habitats of the ice-age Amazon basin were populated by plant species that now thrive only at higher, cooler elevations. But Colinvaux cautions that his pollen records do not show the rain forest drying out and turning to savanna. According to his view, any ice-age drying was "insufficient to affect the forest."

Other researchers, however, have difficulty accepting that a cool but moist regime reigned in South America during glacial times. That combination troubles some scientists who study the earth's changing climate with numerical computer models. "If it were cooler, it would undoubtedly have been much drier," remarks David H. Rind, an atmospheric scientist at the Goddard Institute for Space Studies in New York City. John E. Kutzbach, a climate modeler at the University of Wisconsin, echoes those sentiments: "My experience is that colder continents are also drier continents—substantially drier."

What would a drier climate have meant for the extent of the rain forest? Debate on that point goes back over a quarter of a century. In 1969 Jurgen Haffer suggested that arid ice-age conditions had broken the Amazonian rain forest into several separate blocks. Haffer, a professional geologist and amateur bird-watcher, proposed that the ice-age drying had eliminated the rain forest from the Amazonian lowlands, leaving only isolated forest "refugia" on higher ground. Haffer's theory explained the patchy distribution of certain forest-dwelling birds and insects: these animals had not had time to migrate out of their ice-age forest refuges.

Some biologists assert that Haffer's hypothesis helps to account for the Amazon's enormous biodiversity. Isolation would have allowed regional differences to develop and, eventually, to spawn new species. Changing climate may have been a "species pump."

Whether the massive Amazonian rain forest was truly reduced to a scatter of fragments 15,000 years ago is an open question. Nevertheless, the overall extent of these ancient woodlands may soon be revealed. —*David Schneider*

LONNIE G. THOMPSON *Byrd Polar Research Center, Ohio State University*

ICE CORING on a Peruvian mountain uncovered evidence that dry, dusty conditions prevailed in parts of the Amazon some 15,000 years ago.

Getting the Goats

A national park struggles to rid itself of a charismatic ungulate

What's black and white and red all over? That's no joke on Washington State's Olympic Peninsula, where the U.S. Park Service has gruesome plans for some 300 mountain goats that inhabit the craggy peaks of Olympic National Park.

Park biologists have long contended that the goats, descendants of a small herd brought to the peninsula in the 1920s, are damaging the sensitive and unique alpine ecosystems the park was established to protect. Indeed, the service has already spent hundreds of thousands of dollars on sterilization and live-capture programs in an effort to eliminate the alien ungulates. But those programs have proved to be dangerous and only marginally effective. As a consequence, the service has hatched an alternative proposal: come this summer, it will shoot the remaining goats from helicopters.

"The goats are innocent bystanders," admits park superintendent David K. Morris. "But the mandate of the park is to keep the ecosystem in as natural a state as possible."

Park naturalists claim that mountain goats do not belong in the Olympic range, which is separated by Puget Sound and lowland forests from thriving native goat populations in the neighboring Cascade Range. The geographic isolation of the peninsula has probably prevented Cascade goats from migrating into the Olympics, the naturalists say.

Similarly, that isolation allowed the evolution of eight species of plants found nowhere else in the world. Although none of those species has been federally recognized as endangered or threatened, at least one plant, the Olympic Mountain milk vetch, is thought to be very rare, with fewer than 4,000 individuals remaining on the peninsula.

The milk vetch (variants of which reportedly boost milk production of female goats) is avidly consumed by the animals—as are three other rare Olympian plants. In fact, the goats eat everything from moss to tree branches, and biologists worry that their numbers will not respond to the declining fortunes of any one plant.

"They could eat the last milk vetch on the face of the earth without any impact on their own population," says

Rescuing an Endangered Tree

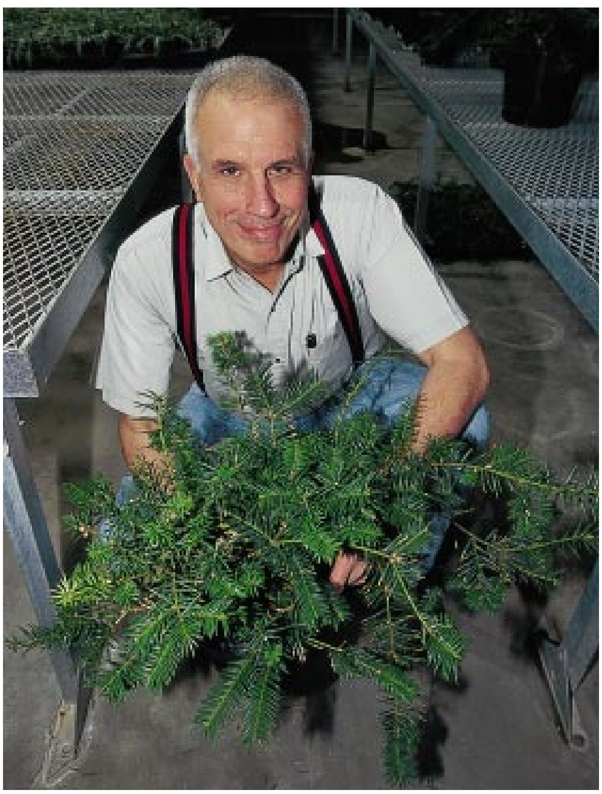
The number of stinking yew trees, named for the pungent odor of their needles, has been dropping since the 1950s. Today the tree (*Torreya taxifolia*) is considered one of the rarest in North America, with only about 1,500 specimens still alive. Efforts to preserve the tree have a particular urgency: the stinking yew is related to the Pacific yew, known for the anticancer drug taxol found in its bark. But only recently have botanists identified what is killing the trees.

Gary Strobel of Montana State University, Jon Clardy of Cornell University and their colleagues report in *Chemistry & Biology* that the dying trees, also known as Florida torreya, are infected with the fungus *Pestalotiopsis microspora,* which belongs to a group of microorganisms known as endophytic fungi. Although not all endophytic fungi harm their hosts, according to Strobel the type living inside torreya trees appears to be on the edge between symbiotic and pathogenic.

The fungus seems to cause disease in the torreya only in arid environments; when moisture levels are high, the microorganism does not appear to harm the tree. The region of northern Florida where the trees grow was once a humid pine forest, but logging wiped out the pines, leaving the land very sandy and dry. "When most of the forest was destroyed, the microorganisms were affected," Strobel says. Starting this summer, Mark Schwartz of the University of California at Davis will begin studying various fungicides that should help protect the remaining plants.

Because there are so few trees still alive, no one has been able to gather enough plant material to test conclusively for taxol or any related substances in the torreya. But Strobel feels the tree is promising as a possible source both of taxol-related compounds and of other pharmaceuticals. Ironically, some of these substances may come from the very fungi that are killing the Florida yew.

It turns out that endophytic fungi, found in most trees and shrubs, produce numerous chemicals, many of which have never been studied. Strobel speculates that fungi found in the Florida yew and many other plants have enormous potential as a source of drugs. Tapping the fungi for new compounds has an important advantage over other harvesting methods. To cultivate the organisms, workers need to remove only a small branch. The rest of the plant continues growing normally. In this way, Strobel states, "you can get plant products without endangering the plants." —*Sasha Nemecek*



DANIEL J. COX

STINKING YEWS are one of the rarest trees in North America. Gary Strobel of Montana State University nurses young trees transplanted from Florida.

Edward Schreiner, a research biologist with the National Biological Service, who has studied the goats' effects on plant communities for 15 years. In addition to grazing, Schreiner says, goats injure rare plants and disturb their habitats by trampling and wallowing (pawing up dirt trenches to lie in during the hot summer months).

Conservation groups, including the Audubon Society and the Sierra Club, have endorsed the Park Service's plan, saying it reflects sound management of so-called exotics that threaten so-called endemics. But animal welfare groups have attacked the plan as arbitrary and inhumane. The New York City–based Fund for Animals, which in the 1970s rescued from the Grand Canyon more than 500 feral burros that had been marked for extermination, claims that the park's argument for eliminating the goats is flawed on several counts. Historical records suggest that mountain goats might be native to the Olympics after all, says fund attorney Roger Anunsen; in any case, goat populations have plummeted as a result of the control programs instituted in the 1980s. And, he adds, the animals do not inflict nearly as much damage as has been claimed.

Proponents of the plan concede that two decades of field studies have failed to determine the risk posed to native ecosystems by mountain goats. "But because we can't measure the impact the goats are having, we must err on the side of caution to protect native species," says Lynn Cornelius, a biologist for the Nature Conservancy in Seattle.

An unnatural end for the Olympic mountain goats would be the ultimate irony in an already twisted tale. The dozen or so goats brought from British Columbia and Alaska 70 years ago were meant to flourish for the benefit of hunters. And flourish they did: by 1983 they numbered 1,200 strong, the majority of which lived in a national park where hunting was no longer permitted. Visitors to the park's alpine meadows had come to cherish the showy wildflowers there as well as the shaggy creatures that ate them, stepped on them and dug them up. Park officials, siding with the wildflowers, began field sterilizations and airlifts to curtail goat populations and managed to reduce the herds considerably before safety concerns halted the programs.

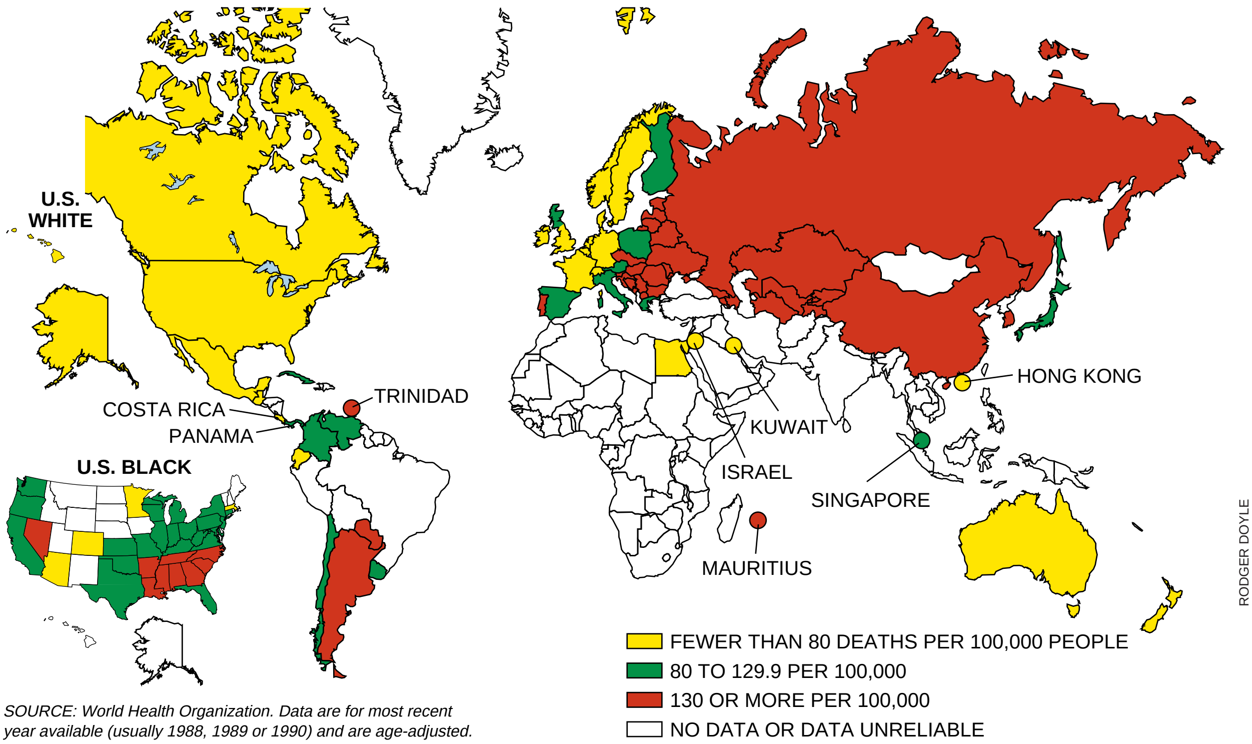
FRANCIS E. CALDWELL AND DONNA L. CALDWELL

WALLOWS created by mountain goats* (inset) *are damaging native flora in the Olympic range.

Now the public seems to be siding with the critters. In a survey conducted for the Fund for Animals last year, almost three fourths of the Washington residents polled opposed the Park Service's plan. Those figures agree roughly with the balance of comments received by the Park Service following publication of its draft environmental-impact statement last year. Park sources say public opposition could sway the service's final decision, which is due in April. And they expect a legal challenge no matter what is decided.

But even if the scheme is implemented this summer, it probably will not be curtains for the Olympic mountain goats, according to Morris. For better or worse, their complete eradication on the Olympic Peninsula is highly unlikely, because goats also occupy the national forest that surrounds much of the park. "We've already caught the village idiots," the park superintendent says; only the most elusive animals are still at large. Morris adds, ruefully, "We'll never get the last goat."—*Karen Wright*

Stroke Mortality in Men Ages 35 to 74



SOURCE: World Health Organization. Data are for most recent year available (usually 1988, 1989 or 1990) and are age-adjusted.

In 1994 half a million Americans suffered a stroke; of these, 154,000 died, more than the number who died from any other cause except coronary heart disease. With more than three million stroke survivors currently incapacitated, it is the leading cause of disability in the U.S. Worldwide, probably more than six million people died from stroke in 1994. The disease, which occurs because of blocking or hemorrhaging of blood vessels in the brain, may result in paralysis of limbs, loss of speech, and other infirmities. The stroke mortality rate of women in most countries is 60 to 90 percent that of men, but because the rates rise steeply with age, and because women live longer than men do, more women actually die of the disease.

Differences in stroke mortality among countries are wide; for example, the former Soviet Union has a rate more than five times that of the U.S. Part of the difference, at least when comparing western countries with eastern European countries, is the result of inferior medical treatment in eastern Europe; however, risk factors, including hypertension, the dominant precursor of stroke, are higher in the East, and cigarette smoking and excessive drinking, which also contribute to the disease, are more widespread there as well. Other, more hypothetical risk factors may add to the eastern European rates, such as a scarcity of citrus fruit, a prime source of vitamin C. Vitamin C and other antioxidants block formation of free oxygen radicals, thought to play a role in the development of atherosclerosis, the underlying condition leading to blockage of arteries.

In the U.S., as in virtually every other country for which data are available, stroke mortality rates have declined dramatically in recent decades. One reason is better detection of milder, more treatable strokes through computed tomography scanning. Since the 1970s, public health programs designed to reduce hypertension through drugs, diet and exercise have been in place, and some countries, such as the U.S., have benefited from declining consumption of cigarettes and alcohol.

Black Americans die from stroke at more than twice the rate of white Americans, primarily because of higher blood pressure levels, apparently resulting, at least in part, from greater sensitivity to dietary salt. Blacks may also be at risk because of poor fetal and infant nutrition, which may contribute to hypertension in later life. In the northern U.S., blacks have a lower stroke mortality rate, perhaps because they are more affluent and hence less apt to suffer nutritional deprivation as children. —*Rodger Doyle*

No Light Matter

Precious helium is blowing in the wind

Most people may associate helium with parties and parade balloons, but the lightest inert gas also has its serious side. Helium is used in a wide variety of scientific and technical applications, from cryogenics to arc welding. It is also one of the earth's most limited resources, found in usable concentrations in just a handful of natural gas wells in the U.S. and Canada. And yet Congress and the Clinton administration are acting to squander, rather than conserve, the helium supply, leading the American Physical Society to issue a warning against their "economically and technologically shortsighted" policies. Edward Gerjuoy of the University of Pittsburgh offers a more personal reaction: "It is morally wrong for this generation to waste a resource that might be precious to a future one."

Helium forms underground from alpha particles—essentially the nuclei of helium atoms—emitted by radioactive elements in the earth's interior. Over millions of years the gas builds up and finds its way into the underground reservoirs where natural gas also collects. Every year drilling companies collect about 3.3 billion cubic feet of helium. A similar amount simply mixes with the atmosphere when natural gas is burned. That helium is for all practical purposes lost forever (the gas can be extracted directly from the air, but only through an expensive and extremely energy-hungry process).

Back in 1960 the federal government, concerned about the strategic value of helium, ordered the Bureau of Mines to establish a reserve in the Cliffside gas field near Amarillo, Tex. In the current budget battle, however, "somehow this thing became a metaphor for a boondoggle," says Robert L. Park of the American Physical Society. The Bureau of Mines is being eradicated, and Congress is seeking to sell off almost all the helium now in storage. Never one to mince words, Park blames "ignorant freshman Republicans" for this situation, although he notes that President Bill Clinton, too, has referred to the reserve as an "anachronism."

Gordon Dunn of the University of Col-

ANTI GRAVITY

It's All Happening at the Zoo(logy Meeting)

In a striking example of convergent evolution, small children sitting in the backseats of cars and male fiddler crabs exhibit common behaviors. That kids in cars and crabs constantly move sideways is well established. Recent research also shows that if you wave at fiddler crabs, they wave back. This was just one of the findings reported at the annual meeting of the American Society of Zoologists (ASZ) in Washington, D.C. The ASZ conference was one of the few things that remained open in our nation's capital last December while Congress fiddled with the budget.

Thought to be an advertisement of territoriality, "waving" had been documented for large groups of fiddler crabs. Denise Pope of Duke University isolated the behavior, showing that an individual crab responds when it sees another individual wave its gigantic claw. She did this by cleverly incorporating a three-inch Sony Watchman screen into one wall of a tank and showing captive fiddlers videos of other fiddlers waving. Although the responding gestures were probably a direct response to a perceived threat, an alternative explanation is that the subject fiddler crabs desperately wanted the remote control in order to change channels (most likely to *Baywatch*).

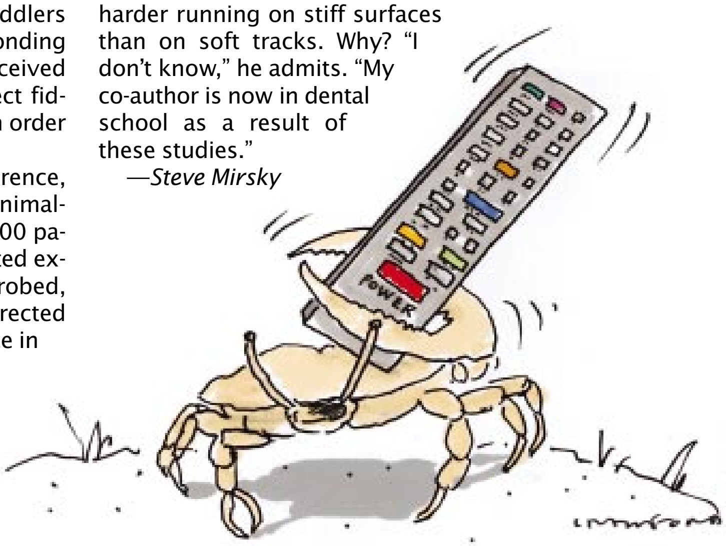
Also waving, outside the hotel hosting the conference, were picket signs, held aloft by a small band of animal-rights activists. While zoologists presented some 600 papers or posters, most of which in some way recounted experiments in which some animal was poked, probed, sliced, frappéed or drugged, the demonstrators directed their displeasure at a fur-warehouse sale taking place in the hotel basement.

A theme of many of the zoology lectures to which the demonstrators were oblivious concerned the energetic costs and kinematics of locomotion. A standard technique in this research is to put animals on a treadmill, again emulating Congress. Lectures and posters described treadmill studies using alligators, horses, lobsters, dogs, wild turkeys, goats, numerous species of lizards, and crabs that weren't waving.

One study, entitled "The Energetics and Kinematics of Running Upside-Down," used as subjects American cockroaches. "We tried using rabbits first, but it didn't work very well," comments roach wrangler Alexa Tullis of the University of Puget Sound. Running a treadmill upside down, she found, requires about twice the energy output of a right-side-up roach run. Previous studies, however, showed that scrambling up a 90-degree incline requires three times as much energy output as running over a level, horizontal surface. The scary conclusion: after a roach scurries up your wall, loping across your ceiling is a breather.

Robert Full of the University of California at Berkeley also ran roaches. "We don't study them because we like them," Full says. "Many of them are actually disgusting. But they tell us secrets of nature that we cannot find out from studying one species, like humans. General principles, once discovered, can be applied to robot design as well as animal locomotion." Roaches can be surprising as well as disgusting—Full found that they worked harder running on stiff surfaces than on soft tracks. Why? "I don't know," he admits. "My co-author is now in dental school as a result of these studies."

—*Steve Mirsky*



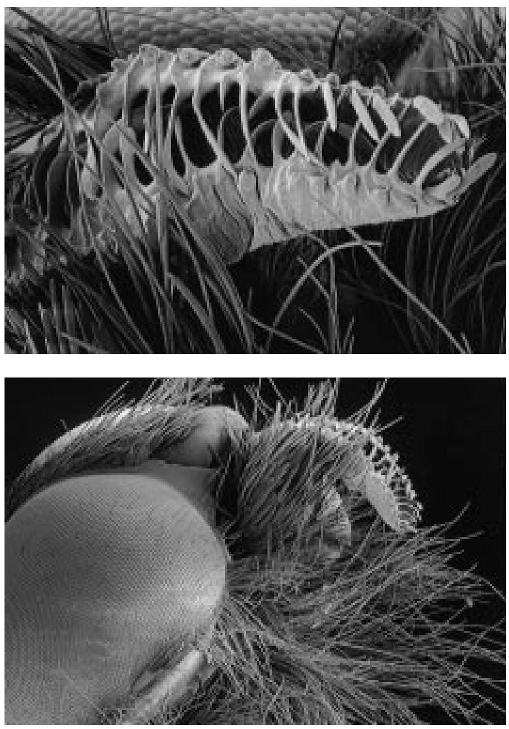
MICHAEL CRAWFORD

orado, a physicist who once served in Congress, attributes much of the current assault on the helium reserve to a lack of understanding of its value on the part of both government officials and the media. About one quarter of all helium is liquefied and used to achieve the ultracold conditions currently needed for some medical imaging devices, as well as for a variety of physics and astronomy experiments. No other element can reach the low temperature of liquid helium; most electrical superconductors require such intense cold to function. Other applications (ballooning, welding, high-purity fabrication techniques) also rely, albeit less critically, on helium.

At present, helium demand is rising about 10 percent a year, so Dunn plausibly argues that the U.S. should be adding to its reserve, not abandoning it. He anticipates that "the helium supply may be largely depleted by 2015," the date by which Congress proposes to have phased out the reserve. Gerjuoy notes that it would cost a substantial $150 million a year for the government to buy up the helium now wasted by natural gas providers, but he suggests that the money would be "a terrific investment, not an expenditure," because the cost of helium is sure to rise as supplies grow tight. So why aren't private companies buying up helium? "Industry does not operate on a 25-year timescale," he sighs.

Even more modest proposals—a natural-gas consumption fee that would finance a helium storage fund, for instance—face an uphill battle in the present political environment. Park is optimistic that a future Congress will at least restore the helium reserve. After all, maintaining it costs the U.S. only $2 million a year, much of which is offset by money earned from short-term storage of helium for private providers. And when that reserve runs out? "We'll have to find new technologies," Park reflects. "Nature did not provide any substitute." —*Corey S. Powell*

Pass the Salt, Please



A REAL SQUIRT: puddling moth ejects wastewater while using its proboscis* (micrographs) *to drink for sodium.

Sodium, essential for nearly every metabolic reaction in the body, is sometimes a rare commodity, so it's no wonder that many animals seek out salt licks, sweat, termite mounds and Big Macs. Some moths, though, are able to meet their salt needs with an ability that would put any barfly to shame. They guzzle almost 40 milliliters of water in a few hours—the human equivalent of more than 40,000 liters, at four liters per second—to absorb the sodium they need.

Naturalists, after long observing butterflies and moths drinking water from puddles, came to suspect that the quest for salt drove the behavior. In the 1970s researchers found that when given a choice of progressively saltier solutions, butterflies usually drank from the saltiest mixture available. Now chemical ecologists Scott R. Smedley and Thomas Eisner of Cornell University have confirmed the role that "puddling" plays in sodium procurement, after studying nature's champion puddlers, male *Gluphisia septentrionis* moths.

These tiny moths, which are only 1.5 centimeters long, drank more than 38 milliliters in three and a half hours, an extraordinary amount that is more than 600 times their body mass. As they drank, the moths powerfully expelled the wastewater up to half a meter away, thus avoiding dilution of their salty puddle. The Cornell ecologists found that the excretions contained less salt than the drink did, confirming that sodium is indeed what *Gluphisia* moths thirst for. The male *Gluphisia* is also well adapted for its quest: it has a proboscis with projections that act as a sieve and a longer intestine to absorb sodium better.

Adult male moths do not live beyond a week, so why did the species develop such specialized physiology and behavior? In part, for the sake of its progeny: the favorite food of *Gluphisia* larvae—poplar leaves—is sodium-poor, so responsible parents must have some other way to ensure that their young will have an ample supply of the critical ion while growing. The burden falls on the male: female *Gluphisia* moths do not engage in puddling activity.

Smedley and Eisner confirmed this theory by showing that sodium is particularly concentrated in the reproductive system of the male. Through its sperm pack, the male passes on excess salt during copulation; the female will not only be impressed by the resourcefulness shown by her mate's gift but can also transfer the sodium to her eggs. With the head start, the larvae defer their salt worries, at least until they reach drinking age. —*Kai Wu*

THOMAS EISNER AND MARIA EISNER *Cornell University*

Having It All

In the economist's jargon, lifetime utility is a function of money earned plus nonmonetary benefits from family and other interests. During the past 10 years or so, attention has focused on how—or whether—women can maximize their returns from both family and career. Many advocates of women's rights contend that the two ought not be mutually exclusive and decry the conditions that make it so difficult to achieve success in both.

It is hard enough to figure out how to juggle work, marriage and children in reality but perhaps even more difficult to conceive an economic analysis that would reliably indicate how well women (or men) are succeeding. Claudia Goldin, a labor economist and economic historian at Harvard University, has been tracking data collected on working women over the past century, starting with the era when work and marriage were almost polar opposites for many women. How well these statistics augur today—or even whether the right numbers are available—is unclear.

There is no way to tell if men combine successful careers with families.

Goldin focused on college-educated women, the prototypes of today's mythical yuppie supermoms. During the first decades of the 20th century, higher education was a ticket to the single life for women. Half of those few who completed college bore no children, and almost a third never married. (In contrast, more than 90 percent of college men at the time married.) Going to college during the 1960s or 1970s had a much smaller impact on the chances of a "successful" family life, at least in part because college had become far more common.

As far as careers go, Goldin found herself with a conundrum: How do you define success? She settled on a simple earnings test: women who were paid more than the bottom 25 percent of men the same age for three years in a row were "successful." About 45 percent of her sample met this criterion (as compared with about 65 percent of men who stayed out of the bottom quartile three years running).

Putting the two sides of lifetime utility together, however, was possible for only about a sixth of the women Goldin studied. At least half of those who managed a career had no children (as opposed to about one in five of those who did not make successful careers). Maximizing returns from family required harsh career sacrifices, and maximizing a career often meant forgoing family entirely (in addition to increased childlessness, "successful" women were more likely to be divorced than their noncareer counterparts).

To anyone who has been reading the newspapers, none of these facts come as a great surprise. In some ways, though, what the numbers do not reveal is almost as interesting as what they do. For example, there is no way of telling how many men combine successful careers with families—as opposed to marriages—because the U.S. does not collect any information on how many offspring men father. Indeed, as Goldin notes, the National Longitudinal Survey, an ongoing study of how Americans' lives change as they grow older, tracks only women. The male side of the survey was cut in 1983, in part because of problems with follow-up. "Men disappear," Goldin says.

What about women (and men) of the 1980s and 1990s? The returns will not be in until well after the turn of the millennium, but some of the conditions that have traditionally made career and family so hard to combine are beginning to ease. It has been more than a generation since employers could automatically fire women who married or became pregnant, day care has become more widely available, and the differences between male and female career choices are narrowing.

Audrey L. Light of Ohio State University has documented the convergence of college majors pursued by men and women, which even 20 years ago were still rigorously separated by sex. Nevertheless, as long as economists and statisticians don't even care enough to survey men about their family lives, the implied bias is clear. Perhaps by the time the lifetime utility is measured the same way for both sexes, their situations will have equalized as well. —*Paul Wallich*

Hot Pork

Energy officials resume plans for reprocessing plutonium

Over roughly four decades, immense complexes near Hanford, Wash., and Aiken, S.C., produced some 100 metric tons of plutonium, the wherewithal of the cold war. To put the metal in a pure form suitable for making weapons, it was extracted from irradiated nuclear fuel in an industrial-chemical sequence known as reprocessing. In 1992, after the cold war ended, the Department of Energy, which oversees the weapons complex, halted all its reprocessing operations.

Now the DOE has announced that it will resume reprocessing on a small scale—and that it will continue work on a new reprocessing technology despite its limited need for reprocessing in the foreseeable future. The development has provoked charges, including some from within the DOE itself, that the plan is fraught with pork-barrel politics. At the same time, at least one public-interest group is challenging the need for reprocessing.

James R. Giusti, a DOE spokesman, says the work will achieve environmental and safety objectives rather than military ones. It is necessary to "stabilize" irradiated fuels that have corroded and therefore represent a safety risk at the DOE's Savannah River site in South Carolina. The reprocessing will put the waste into forms more suitable for storage and eventual disposal, Giusti adds.

Critics of the plan argue that the reprocessing will unnecessarily risk the health and safety of workers and people living near the plant, add more plutonium to the country's large accumulation and undermine U.S. nuclear-nonproliferation efforts to curb reprocessing and plutonium production by other countries. The groups, including the Institute for Energy and Environmental Research and the Energy Research Foundation, also question the need for the new reprocessing technology at a time when the U.S.'s accumulation of plutonium has become a serious liability. The technology is being developed at the DOE's Idaho National Engineering Laboratory. The decision, opponents suggest, was motivated partly by the desire to preserve jobs in areas that are somewhat depressed economically.

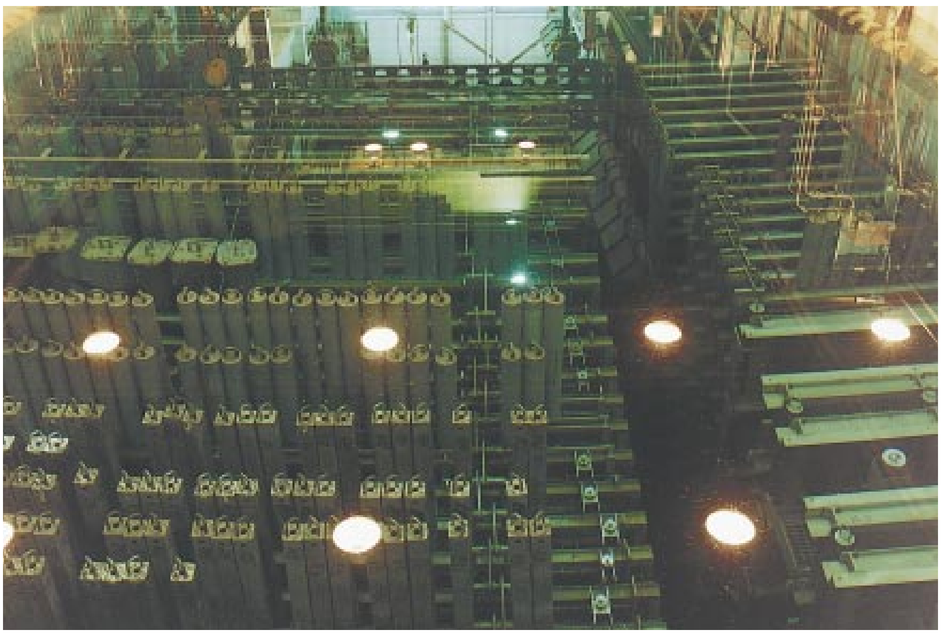
The fuels being considered for reprocessing, approximately 5 percent of the DOE's total, are being stored underwater in basins at Savannah River. Protective claddings have become corroded, or protective containers have failed, releasing radioactive materials into the water and making maintenance more difficult, costly and potentially hazardous, according to DOE officials. The fuels include 140 metric tons of weapons-reactor fuel, 81 canisters of fuel from a Taiwanese research reactor and one from a U.S. breeder reactor no longer in operation. Two other types of weapons-reactor fuel, amounting to seven metric tons, will probably also be reprocessed. Giusti says the plutonium recovered from the reprocessing will be stored at Savannah River until "the department finalizes its plutonium disposition plans."

According to Noah Sachs of the Institute for Energy and Environmental Research, reprocessing of the fuels would take about 10 years and produce 11,600 cubic meters of high-level waste, 31,600 cubic meters of low-level waste and 720 cubic meters of transuranic waste (which contains elements with atomic numbers greater than 92). The quantities would increase Savannah River's accumulations of these three types of waste by 9, 4.8 and 8.1 percent, respectively. The institute issued a report in January calling on the DOE to handle the corroding wastes by shoring up the storage tanks where the fuels were put and then transferring the fuels to a better wet-storage environment or to dry storage, in air or an inert gas.

Gordon M. Nichols, Jr., director of the chemical-separation division at the DOE's Savannah River site, responds that "we are improving our wet-storage options because it is the prudent thing to do, but there's no way we see wet storage as a long-term option." He adds that "we don't believe we know enough at this time about dry storage to place aluminum-clad fuels in dry storage safely for the amount of time that would be required."

The DOE has already begun transforming for storage various solutions containing isotopes of plutonium, uranium, americium, curium and neptunium. The solutions, totaling some 300 cubic meters, are by-products of reprocessing left in the facilities when they were abruptly shut down almost four years ago. Some of the materials will be converted into dry oxides; others will be placed in small glass logs.

Most of the fuel reprocessing planned so far will be done in a plant, or "canyon," designated F. One controversial issue, which was expected to be resolved in February or March, was whether the



WESTINGHOUSE SAVANNAH RIVER COMPANY

CONTAINERS OF SPENT FUEL rods are cooled and shielded in a storage pool at the Savannah River site in South Carolina. The Energy Department's plans to reprocess these and other fuels, ostensibly to make them safer for long-term storage, have generated controversy.

DOE would also use a second canyon, called H. Powerful interests have supported the use of H-canyon, including the Defense Nuclear Facilities Safety Board and Senator Strom Thurmond of South Carolina. It is also likely that the DOE's main contractor at the site, Westinghouse Savannah River Company, will make more money if it operates two canyons rather than one.

A recent DOE study concluded, however, that the department could save up to $200 million over 10 years by phasing out H-canyon. In January a DOE employee knowledgeable about Savannah River operations told SCIENTIFIC AMERICAN: "The only reason I could see [to continue using H-canyon] is if the nation completely reversed itself and the decision was made for the U.S. to begin reprocessing civilian [power reactor] nuclear fuel. But I don't think we're anywhere near that kind of a policy shift."

The work at the Idaho laboratory on the new reprocessing technology has also been characterized as unnecessary. An expert familiar with the congressional deliberations over the project called it "pure pork. It was forced down [the DOE's] throat by the Senate Armed Services Committee"—ultimately, as part of the Energy and Water Appropriations Bill. The bill made available $25 million for continued development of the technology, called electrometallurgical processing or pyroprocessing. Another $25 million was set aside for related work on spent breeder-reactor fuel. "It's just obscene," the source says.

"It was a political deal the administration cut to get rid of the Integral Fast [breeder] Reactor," the expert explains. "The administration didn't want to fund another breeder, and the way to get the Idaho and Illinois [congressional] delegations to agree to get rid of the breeder was to agree to fund the reprocessing technology. But they [the delegations] are still pushing for the reactor, so I don't know what was achieved."

Senator Dirk Kempthorne of Idaho has been one of the supporters of the technology. Through a spokesman he said: "This country has turned its back on solving a problem [nuclear waste] that must be solved. We could shut down every nuclear facility tomorrow, and we'd still have waste that needs to be dealt with.... [Electrometallurgical processing] is a promising technology that allows the U.S. to use its best minds and facilities to prepare spent nuclear fuels for final disposition."

Over the past year, officials at the Idaho lab have portrayed electrometallurgical processing as a waste management tool, but as such it has won few supporters outside of Idaho. —*Glenn Zorpette*

Smart Shopping

The Pentagon tries to teach itself new tricks

The Pentagon has never been a paragon of adept buying practices, but a new Defense Department initiative may help bring advanced technologies to soldiers more quickly and for less money. So far the program has delivered an unmanned reconnaissance aircraft to troops in Bosnia in record time, and its proponents believe they can simplify the Pentagon acquisition system by allowing the military "users"—the ones who take the stuff to war—to influence the design of weapons and other technologies from the beginning.

GENERAL ATOMICS

PREDATOR, an unmanned reconnaissance drone developed in less than two years under an innovative new Pentagon procurement plan, is now being tested by U.S. soldiers deployed in the Balkans.

It may sound simple, but this is the Pentagon, where a system thought up in one decade might not be used by a soldier for another two. During the cold war, when the U.S. worried almost exclusively about the Soviet Union, this inertia worked because both nations knew enough about the other's weapons programs to keep pace, even with 30-year acquisition cycles. "There was a well-understood relationship between us and our military capability and them and theirs," says Jack Bachkosky, deputy undersecretary of defense for advanced technology.

With the explosion in information and computer technology, however, the old way won't do. Now the Pentagon needs to deliver new technologies to its soldiers within years, not decades, because many may grow obsolete by the time they make it to the field. For years, the Pentagon has made noise about simplifying its buying system, calling the attempts "acquisition reform," with mixed results. Under the current administration, many of the arcane rules and regulations governing military procurement practices have been swept away, a crucial first step for the new initiative. It doesn't have a name, but its products are called advanced concept technology demonstrations (ACTDs). They are so different from traditional efforts that Pentagon officials actually use the word "informal" to describe them.

In three years ACTDs have grown to command about $1 billion of the defense budget, and Pentagon acquisition executive Paul Kaminski calls the program "one of the fundamental core elements in improving our acquisition system." About 20 ACTDs are in development, and 10 or more are set to be initiated each year.

An ACTD is created when the Pentagon recognizes a military requirement and matches it with a mature technology that has not yet been adapted for military purposes. Users and developers work hand in hand to design the technology, and no commitment to producing it in large numbers is made until soldiers who have tested prototypes in the field testify to its performance. If the technology does not work, it is scrapped; if it does, it can be fed into the traditional system and built on a large scale, or it can be further refined. In any case, a small set of operational hardware is produced that can be used by soldiers—even in wartime.

Not everyone is a fan. Within the military, along with the typical resistance to new ideas, there is a more specific dislike of civilians in the Pentagon getting deeply involved in weapons systems development. Nevertheless, Bill McCorkle, who heads the Army Missile Command's research and development center, credits the ACTD process for "saving" a system he helped to get off the ground, called the enhanced fiber-optic guided missile. McCorkle believes the army shied away from a promising technology for years because it was new, it was "not invented here," and it did not fit well into existing service structures. Now, he hopes, the ACTD will allow for the missile system to be demonstrated in realistic settings, which may convince the Pentagon to buy it in large numbers.

The Defense Department still needs to figure out what to do with an ACTD once its technology is ready for production. And the Pentagon's track record on ideas for saving money is not good. Still, one of the first ACTDs has already shown promise. In 1993 the military needed an inexpensive reconnaissance airplane, and it needed it quickly. Simultaneously, the Defense Department was kicking around the ACTD concept. The two ideas were matched, and a contract was awarded within months to General Atomics in San Diego.

About a year later, in 1995, an unmanned prototype aircraft known as Predator was flying missions over the former Yugoslavia. This year it will be back over Bosnia serving as an eye in the sky for U.S. soldiers, and those same soldiers will have a say in deciding if the Pentagon should commit millions to building more. Now that's acquisition reform. *—Daniel G. Dupont*

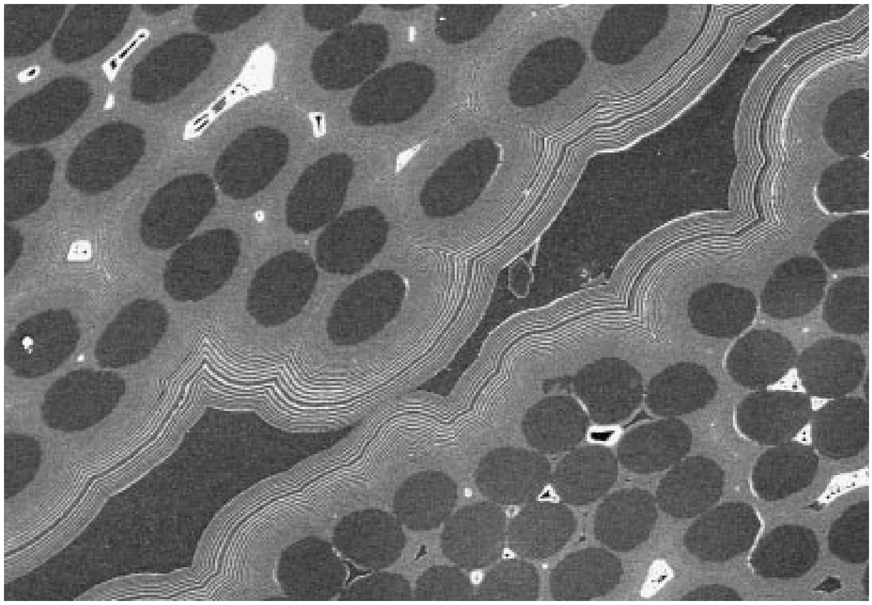
Tough Stuff

Ceramic composites may get stronger—and cheaper

Ceramic composites, the darlings of the material world, have been too precious and fragile to realize their full potential. Certainly in toughness, advanced ceramics surpass metals that weigh much more. But composites' brittleness and high cost have made them impractical for many applications, such as airplane engines and heat exchangers, the performance and efficiency of which are limited by inadequate materials.

Those limits may soon be overcome—and the cost of strong composites cut—thanks to a new class of materials recently invented at the Georgia Technology Research Institute (GTRI). The advance simply combines two strategies long used to strengthen materials. The first is reinforcement. Much as metal rods can internally buttress concrete bridges, carbon fibers toughen ceramics grown around them. In metals, platelets of silicon carbide serve the same function. A second well-known way to make strong stuff stronger is to stack thin layers of two metals or ceramics into a laminate. The product is, like an oyster's shell, much harder than the mere sum of its parts.

In the past, the main obstacle to laminating reinforcements, according to W. Jack Lackey, who led the GTRI team, has been that growing layers of ceramic around a mesh of fibers is quite tricky. Laminating round platelets—which can cost 10 to 100 times less than carbon fibers—is even harder. Lackey's group succeeded on both counts, however, using a process called chemical vapor infiltration.



GEORGIA TECHNOLOGY RESEARCH INSTITUTE

SILICON CARBIDE PLATELETS are laminated with ceramic layers through a process called chemical vapor infiltration.

Lackey has not yet performed the mechanical tests that will show whether his new "laminated matrix composites" are indeed stronger and more resistant to heat than any other composites yet produced. But past experience suggests that the materials ought to be tougher than the competition, so the GTRI has filed a patent on the invention. If the technology pans out, cheap ceramic composites may become a lot more commonplace. *—W. Wayt Gibbs*

Bad Timing

A loophole is found in a popular encryption scheme

A new cryptanalytic attack has shaken confidence in the security of some very popular encryption schemes—and computer experts are stunned at how easy it can be to unravel "securely" coded messages. A public-key encryption scheme uses a public key to encrypt messages, but the decryption key is kept private. Now Paul C. Kocher, a cryptography consultant, has found a back door. Kocher proved that a wily snooper can figure out what the secret key is—by keeping track of how long a computer takes to decipher messages.

Public-key cryptography relies on certain mathematical functions that are very easy to do but very hard to undo. For instance, it is easy to multiply two numbers together to get a larger number yet hard to factor a large number into its component primes.

Encryption schemes take advantage of this fact; because the operations are easy to do, it takes very little time for people to encrypt messages using the public key. Only the authorized user knows the easy way to decrypt them, however. Typically, for a would-be attacker to crack the code, he would have to perform very difficult operations, such as factoring.

Kocher's attack makes an end run around the mathematics. Just as a burglar might guess the combination to a safe by seeing how long it took for somebody to turn the dial from number to number, a computer hacker can figure out the cryptographic key by timing the computer as it decrypts messages. The burglar has no need to crack the safe; the hacker has no need to factor a large number.

The attack depends on the public nature of the key. Like everybody else, an attacker can use the public key to encrypt a message and send it to a computer. If he times how long it takes for the computer to respond, however, he can get a rough idea of how much time the decryption process took. Because he knows the public key, the timing measurements can reveal a lot of information about where the bottlenecks in the process are. After a number of observations (with accurate timing, it takes only several hundred to a few thousand tries), the hacker can analyze those bottlenecks to learn what secret number the computer is using to decrypt the messages.

Is the timing attack a real threat to security? "Oh, God, yes!" exclaims Bruce Schneier, author of *Applied Cryptography,* published in 1995. "You can't belittle the realness of it. It's not only a theoretical attack—you can do this!" Worse news for security buffs: the attack is self-correcting. It guesses the bits of the secret key, one by one; if a code breaker makes a mistake in guessing a bit, he will quickly discover that the attack stops making progress. The hacker goes back, corrects the bad bit, then resumes the attack anew.

Fortunately, the attack does not affect the security of systems that a snooper cannot time accurately. Popular e-mail encryption schemes such as Pretty Good Privacy (PGP) are not compromised. The vulnerable systems are those that respond quickly to outside requests and depend on digital "certificates" to verify the user's identity. Network servers, for example, pass certificates back and forth to ensure that only authorized users get access to a particular part of a network.

"A certificate is essentially who you are—it is your digital identity," Kocher says. Certificates are embedded in "smart" cards, credit-card-size devices containing a processor and memory, to protect them from outside attack. But if a computer cracker can run timing tests on the smart card (either directly or by timing a network into which the card is plugged), its secret code could be broken.

One way of locking the back door is to make the computer wait to spit out answers rather than responding as quickly as it can. Another method uses a process called blinding, in which the computer multiplies the message by a random number before exponentiating it. This process prevents the attacker from knowing what numbers are being knocked around inside the computer. According to Kocher, Netscape, makers of popular software for browsing on the World Wide Web, will now use blinding to prevent timing attacks on encrypted transmissions—protecting consumers who use their credit cards on the Web.

Although Kocher's key-extraction technique can be foiled, it shows that cryptographers can never be complacent; the real world has traps that mathematicians may not foresee. "In theory, there are other attacks," Schneier remarks. "You can measure power consumption or heat dissipation of a chip; timing is just one way. The moral is that there's always something else out there." —*Charles Seife*

PROFILE: ALBERT LIBCHABER

Seeing the World in a Snowflake

The walls of his bedroom are lined with books, original editions of works by the greats—Newton, Descartes, Leibniz, Galileo, Poincaré, anyone I can think of. Albert Libchaber pulls out the volumes one by one, running his fingers along favorite passages and translating for my benefit. Kepler's musings, in 1611, on a snowflake: a "nothing" that reveals in its symmetry the atomic structure of matter. Hooke's drawings of a fly's eye, revealed by one of the earliest microscopes. Lyapunov's treatise on the stability of motion, presaging chaos theory. His heroes, Libchaber explains, are Huygens and Kepler: "They are more passionate, more human, more romantic, therefore less well known than Galileo or Newton." He is disappointed that I cannot name any heroes of my own; he would have liked to thrill me by pulling out their works.

Staring at Huygens's exquisite diagrams of the pendulums he crafted and studied, I suddenly see the wellspring that inspires Libchaber. "I have a feeling, if an experiment is aesthetic it will tell me something," he had said earlier in his soft, almost inaudible voice. "I will not do an experiment if it is not beautiful." Libchaber emulates his heroes, whose genius touches him through the pages of these books. He asks direct, simple questions—doing, as someone said, 19th-century physics with 21st-century equipment. In 1979 his precise techniques led him to see, in a tiny cell of liquid helium at the École Normale Supérieure in Paris, how a fluid's flow becomes disordered—the first close look at chaos in nature.

When I first met Libchaber, some 10 years ago at the University of Chicago, I did not know he was a famous man. The aroma of his pipe would announce his presence in the research buildings, allowing me, then a graduate student, to waylay him with queries about vortices and smoke rings. He would answer in detail, with illustrative waves of his hands and pipe. Downstairs, his laboratories were filled with endlessly entertaining ventures, featuring colorful liquid crystals, pulsating magnetic bubbles or long fingers of oil pushing through water. At the Rockefeller University's new Center for Physics and Biology, the peripatetic professor is now turning his laboratory to the study of life-forms. "How did life start?" is the question that occupies him.

His restlessness, Libchaber explains, has early roots. Born in 1934 in Paris to Jewish immigrants from Poland, he was six when Germany invaded France. His parents, who had strong—and revealing—accents, decided that Albert and his older brother, Marcel, were more likely to survive on their own. Posing as orphaned Catholics from Alsace, the brothers lived out the war in the south of France, moving from family to family for safety.

"We are alive because a number of French people helped," Libchaber acknowledges. Still, the boys lived in perpetual fear of being found out—of being overheard discussing their plight or being revealed as circumcised. Albert was confused about why he had to lie: "When you are small and you know people want to take you someplace, you don't understand why." Albert enjoyed going to church and had to be reminded sometimes by his brother that he was Jewish. "I didn't know what that meant," Libchaber recalls. In the streets, homesick German soldiers would stop to hug and kiss the little boy.

After four years in hiding, word arrived one day that the Americans were coming. Albert and Marcel ran ahead from their village near Marseilles to greet the troops. "They asked us where the Germans were," Libchaber relates. "We told them they had left." The relieved soldiers gave the children chewing gum and played games with them. "They were very young men, joyful. Left a wonderful impression of America," Libchaber smiles.

The brothers were reunited with their parents, who had somehow survived the war. Marcel was happy. But Albert cried, refusing to go to them: he did not know them anymore. As it was, the couple had enough on their minds. "My mother lost everybody, my father 90 percent," Libchaber says. "When people say the camps did not exist, I can hold up a long list of aunts, cousins—"

The ever growing tally of murdered relatives finally brought home to Albert what it meant to be Jewish. But the nomadic childhood also left a mark. "It has affected the fact that I move so well," remarks Libchaber, with a slight, sad twist of his mouth. "America, France, I don't feel in touch. I'm a wandering Jew."

His father, Chil Libchaber, was in fact the son of a rabbi. Endowed with a deep love of books, he studied into the night after each day of work in post-war Paris. "My father would say, 'You learn or you go to work, there is nothing else,' " Libchaber laughs. Although captivated by city life, the adolescent Albert absorbed this passion for scholarship, inventing his own special blend. "I am Jewish and French," Libchaber declares. "French [means] rational, mathematical. Jewish—a mystical view of study and learning." What scientists do, he contends, is little different from what Talmudists do: "There is a coded message, [you] find the code."

The Frenchman obtained a bachelor's degree in mathematics, a year after marrying, at the age of 20, his sweetheart, Irene Gellman. Although the marriage was happy, paper and pencil were not enough to satisfy his intellect. "The essential constraint that separates physics from the mystical is experiment. I felt the need to do tests," Libchaber recounts. He became enamored of the exciting new electronic devices—transistors, amplifiers, diodes, lasers—that were rapidly coming out. A Fulbright fellowship soon took him to the University of Illinois at Urbana-Champaign to study with John Bardeen, the wizard of solid-state physics. Libchaber, long used to the hierarchical habits of French academe, was stunned by Bardeen. "He had already one Nobel Prize, but his office door was always open. I could see as much of him as I wanted to," he says.

ALBERT LIBCHABER'S tabletop experiments unscrambled the codes of chaos.

ARNOLD NEWMAN

The apprenticeship, which left a deep impression on Libchaber, was not, however, to last. A year and a half into his graduate studies, the student was drafted and sent to the Sahara to fight a new French war—against Algerians demanding independence. Assigned to the atomic weapons squad, Libchaber had to be at "point zero" the day after an explosion to measure the radioactivity and deduce the energy released. The extreme desert environment made his task doubly challenging: "I learned the hard way, experimental science."

After the war, Libchaber returned to Paris—and a newborn son—to finish his doctorate, with Pierre Aigrain, at the École Normale. There he did an influential experiment. Radiation does not normally pass through a metal; Libchaber showed how an extra magnetic field could create helical waves that nonetheless penetrate metals with ease. Dividing his time between New Jersey and Paris, he began to work with C. C. Grimes of Bell Laboratories on diverse problems in metals and superconductors.

In the mid-1970s Libchaber, who now headed his own research group at the École Normale, turned his attention to the flow of superfluid helium. The quantum liquid glides along smoothly until vortices form; these then trip over one another, causing turbulence. Libchaber decided to first study an isolated vortex or two. With an engineer, Jean Maurer, he crafted a tiny metal cell, three millimeters wide and 1.25 millimeters tall. Gently heated from below, helium in the cell flowed upward near the center and back down at the sides, forming two parallel rolls. Probes at the top measured the fluid's temperature.

As the heat increased, waves began to ripple up and down the length of the two vortices. The probes told a surprising tale—each new wave had exactly twice the wavelength of the preceding one. Ultimately, all the new waves jumbled up to make the flow chaotic. Libchaber's jagged graph, with spikes marking the different waves, found its way into the hands of Mitchell J. Feigenbaum, now at Rockefeller. Within a few months the theorist wrote to the experimenter. The helium cell had revealed the first route to chaos, now designated the "period-doubling cascade."

Until then, chaos had been a mere curiosity, a plaything for mathematicians. "The moment Albert did his experiment, and chaos showed up in a real thing, not to mention a fluid, it completely changed the reaction of the [physics] world," Feigenbaum declares. Physicists had always believed that as a system becomes disordered, waves of many arbitrary lengths develop. Instead what happens is extremely precise and ordered—only subharmonics of a fundamental wave appeared in Libchaber's cell. The elegance of the helium experiment was, in fact, vital to its success: a comparable study in water would have required a tub, too big to control.

Despite his achievement, Libchaber was restless at the École Normale. When Leo P. Kadanoff, a brilliant, gruff theorist from the University of Chicago, visited to recruit students, Libchaber offered, only half-joking, "I can come, too." In truth, America enticed the Frenchman. "It's an adolescent country," he explains. "Has vitality. Makes no plans, makes mistakes, can recover from anything. France is late middle-aged." Besides, the country appealed to his nomadic instincts, offering the freedom "to go, to move, to do." The transfer, in 1983, started a fruitful "strong interaction" with Kadanoff. "We would be exchanging ideas," the American theorist recalls. "From the ideas would flow inspiration for experiments, and from the experiments, new ideas."

Life advances by trial and error. There is no planning, only evolution.

At Chicago, Libchaber demonstrated a second route to chaos. Thomas Halsey, then an assistant professor, describes how Libchaber came into his office at 11:00 P.M. one Friday to announce that—after months of effort—his project was finally making sense. "That experiment was so beautiful that it killed the field," Halsey states. It demonstrated in exquisite and exhaustive detail the quasiperiodic route to chaos, in which the new waves are not subharmonic but have wavelengths related by the number 1.618—the Golden Mean.

Libchaber went on to study full-blown turbulence. Chaos is just a first step toward disorder, involving only a few kinds of motion. "You freeze space, play with time" is how the experimenter puts it. He wanted to play with space as well, and with his students, he found curious mushroom-shaped plumes and other long-lived structures in turbulent water. But there was no theory to explain these findings: the problem of completely disordered motion remains unsolved. "I did not see why continue just getting data," Libchaber comments. Convinced that the interesting and doable problems in condensed matter had all been done, Libchaber made another move in 1991—to biology, and to Princeton University and the NEC Institute.

"Chicago is an unstable fixed point," he explains. Scientists are attracted to it, but many leave for the true fixed points on the East and West coasts. But Princeton could not hold him either. "People work at home, you don't see much of them. I didn't interact very well," he sums up. In 1994 Libchaber moved on to Rockefeller to surround himself with biologists. One unlikely reason for the move: New York City. "It's ugly, dirty, bankrupt," Libchaber says. "But it's alive. In New York, everything is possible." He hikes for hours down the city streets, reveling in the diversity of the faces. He does not miss nature; there is enough in the lab.

Much in biology appears oddly familiar to Libchaber. As in his turbulent plumes, the fluids in cells are highly uncooperative. Their large viscosity impedes all motion, while their molecules constantly bombard the minute cellular bodies, making them bounce around. "How can the laws of physics apply to such a difficult environment and create high technology?" Libchaber asks. A host of microscopic machines, he observes, keeps things moving: "There are pumps, motors, channels, highways. So much so that you have the feeling that we never invented anything, life did it before. We are just rediscovering."

But Libchaber does note a fundamental distinction between physics and biology. Physicists construct a simplified world whose behavior they can predict; biologists study the world as they find it, reflecting the nature of life. "If you are an engineer and you want to make a rocket engine, you design it completely different from usual engines. [Life] will start from a classical engine, add some more, make it much more complex. There is no planning, only evolution. Some things are useless, [but] you don't throw them out." Even so, this untidy process of "trial and error" produces machines and computers of utmost efficiency, able to detect a single photon or molecule.

The first studies by Libchaber's new group are teasing out the forces exerted by the tiny cellular machines and the means by which microtubules—the "highways"—grow. The physicists next plan to use DNA in making test-tube computers. But these are only warm-up problems. One day, Libchaber dreams, he will build from scratch an elementary living world—as small and complete as a snowflake. —*Madhusree Mukerjee*

Urban Planning in Curitiba

A Brazilian city challenges conventional wisdom and relies on low technology to improve the quality of urban life

by Jonas Rabinovitch and Josef Leitman

As late as the end of the 19th century, even a visionary like Jules Verne could not imagine a city with more than a million inhabitants. Yet by the year 2010 over 500 such concentrations will dot the globe, 26 of them with more than 10 million people. Indeed, for the first time in history more people now live in cities than in rural areas.

Most modern cities have developed to meet the demands of the automobile. Private transport has affected the physical layout of cities, the location of housing, commerce and industries, and the patterns of human interaction. Urban planners design around highways, parking structures and rush-hour traffic patterns. And urban engineers attempt to control nature within the confines of the city limits, often at the expense of environmental concerns. Cities traditionally deploy technological solutions to solve a variety of challenges, such as drainage or pollution.

Curitiba, the capital of Paraná state

LAKESIDE PARKS (*right*) serve multiple functions in Curitiba, the capital of the state of Paraná in southeastern Brazil (*above*). In addition to providing green space for citizens and forming part of the metropolitan bicycle-path network, they help to control the floods that once plagued the city. The artificial lakes, created during the 1970s, are designed to facilitate drainage and to hold excess rainwater and keep it from inundating low-lying areas.

in southeastern Brazil, has taken a different path. One of the fastest-growing cities in a nation of urban booms, its metropolitan area mushroomed from 300,000 citizens in 1950 to 2.1 million in 1990. Curitiba's economic base has changed radically during this period: once a center for processing agricultural products, it has become an industrial and commercial powerhouse. The consequences of such rapid change are familiar to students of Third World development: unemployment, squatter settlements, congestion, environmental decay. But Curitiba did not end up like many of its sister cities. Instead, although its poverty and income profile is typical of the region, it has significantly less pollution, a slightly lower crime rate and a higher educational level among its citizens.

Designing with Nature

Why did Curitiba succeed where others have faltered? Progressive city administrations turned Curitiba into a living laboratory for a style of urban development based on a preference for public transportation over the private automobile, working with the environment instead of against it, appropriate rather than high-technology solutions, and innovation with citizen participation in place of master planning. This philosophy was gradually institutionalized during the late 1960s and officially adopted in 1971 by a visionary mayor, Jaime Lerner, who was also an architect and planner. The past 25 years have shown that it was the right choice; Rafael Greca, the current mayor, has con-

JONAS RABINOVITCH *Courtesy of the city of Curitiba*

JONAS RABINOVITCH

JONAS RABINOVITCH

24-HOUR STREET, an arcade of shops and restaurants that never closes, helps to keep Curitiba's downtown area vital. The city has also regulated the locations of banks, insurance companies and other nine-to-five businesses to prevent the district from becoming a ghost town after working hours.

tinued the policies of past administrations and built on them.

One of Curitiba's first successes was in controlling the persistent flooding that plagued the city center during the 1950s and early 1960s. Construction of houses and other structures along the banks of streams and rivers had exacerbated the problem. Civil engineers had covered many streams, converting them into underground canals that made drainage even more difficult—additional drainage canals had to be excavated at enormous cost. At the same time, developers were building new neighborhoods and industrial districts on the periphery of the city without proper attention to drainage.

Beginning in 1966 the city set aside strips of land for drainage and put certain low-lying areas off-limits for building. In 1975 stringent legislation was enacted to protect the remaining natural drainage system. To make use of these areas, Curitiba turned many riverbanks into parks, building artificial lakes to contain floodwaters. The parks have been extensively planted with trees, and disused factories and other streamside buildings have been recycled into sports and leisure facilities. Buses and bicycle paths integrate the parks with the city's transportation system.

This "design with nature" strategy has solved several problems at the same time. It has made the costly flooding a thing of the past even while it allowed the city to forgo substantial new investments in flood control. Perhaps even more important, the use of otherwise treacherous floodplains for parkland has enabled Curitiba to increase the amount of green space per capita from half a square meter in 1970 to 50 today—during a period of rapid population growth.

Priority to Public Transport

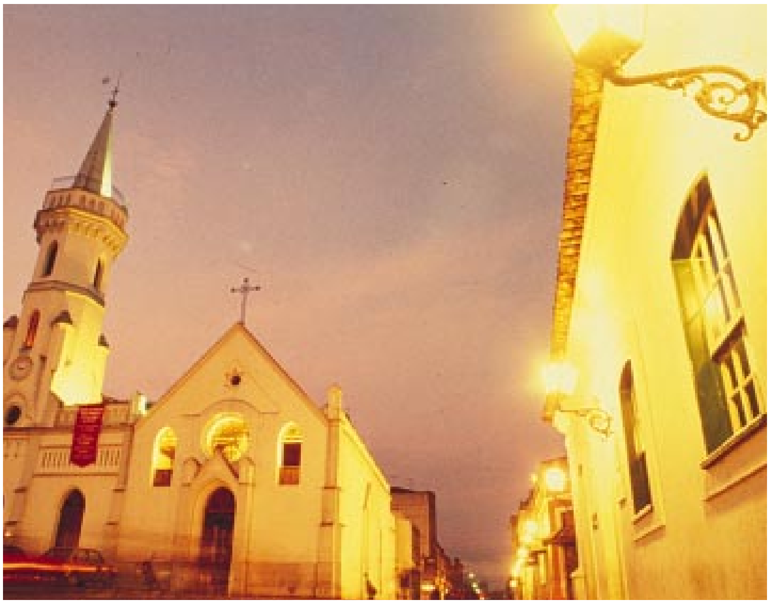
Perhaps the most obvious sign that Curitiba differs from other cities is the absence of a gridlocked center fed by overcrowded highways. Most cities grow in a concentric fashion, annexing new districts around the outside while progressively increasing the density of the commercial and business districts at their core. Congestion is inevitable, especially if most commuters travel from the periphery to the center in private automobiles. During the 1970s, Curitiba authorities instead emphasized growth along prescribed structural axes, allowing the city to spread out while developing mass transit that kept shops, workplaces and homes readily accessible to one another. Curitiba's road network and public transport system are probably the most influential elements accounting for the shape of the city.

Each of the five main axes along which the city has grown consists of three parallel roadways. The central road contains two express bus lanes flanked by local roads; one block away to either side run high-capacity one-way streets heading into and out of the central city. Land-use legislation has encouraged high-density occupation, together with services and commerce, in the areas adjacent to each axis.

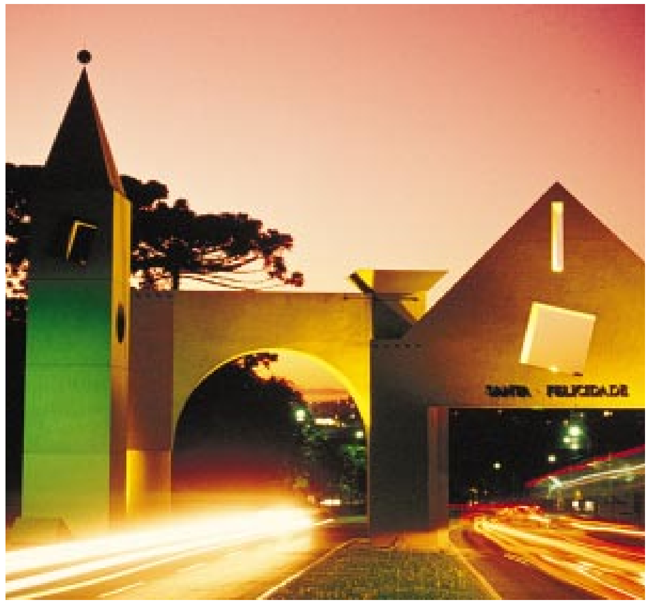
The city augmented these spatial changes with a bus-based public transportation system designed for convenience and speed. Interdistrict and feeder bus routes complement the express bus lanes along the structural axes. Large bus terminals at the far ends of the five express bus lanes permit transfers from one route to another, as do medium-size terminals located approximately every two kilometers along the express routes. A single fare allows passengers to transfer from the express routes to interdistrict and local buses.

The details of the system are designed for speed and simplicity just as much as the overall architecture. Special raised-tube bus stops, where passengers pay their fares in advance (as in a subway station), speed boarding, as do the two extra-wide doors on each bus. This combination has cut total travel time by a third. Curitiba also runs double- and triple-length articulated buses that increase the capacity of the express bus lanes.

Ironically, the reasoning behind the choice of transportation technology was not only efficiency but also simple economics: to build a subway system would have cost roughly $60 million to $70 million per kilometer; the express



JONAS RABINOVITCH



JONAS RABINOVITCH

HISTORIC CENTER of Curitiba (*left*) has received special planning protection, including incentives to build elsewhere, that preserves old buildings. Many of the district's streets have been converted to pedestrian use, reducing pollution and fostering a sense of neighborhood. Ceremonial gates mark sections of the central city that were once enclaves for particular immigrant groups (the entrance to the former Italian quarter is shown at the right).

bus highways came in at $200,000 per kilometer including the boarding tubes. Bus operation and maintenance were also familiar tasks that the private sector could carry out. Private companies, following guidance and parameters established by the city administration, are responsible for all mass transit in Curitiba. Bus companies are paid by the number of kilometers that they operate rather than by the number of passengers they transport, allowing a balanced distribution of bus routes and eliminating destructive competition.

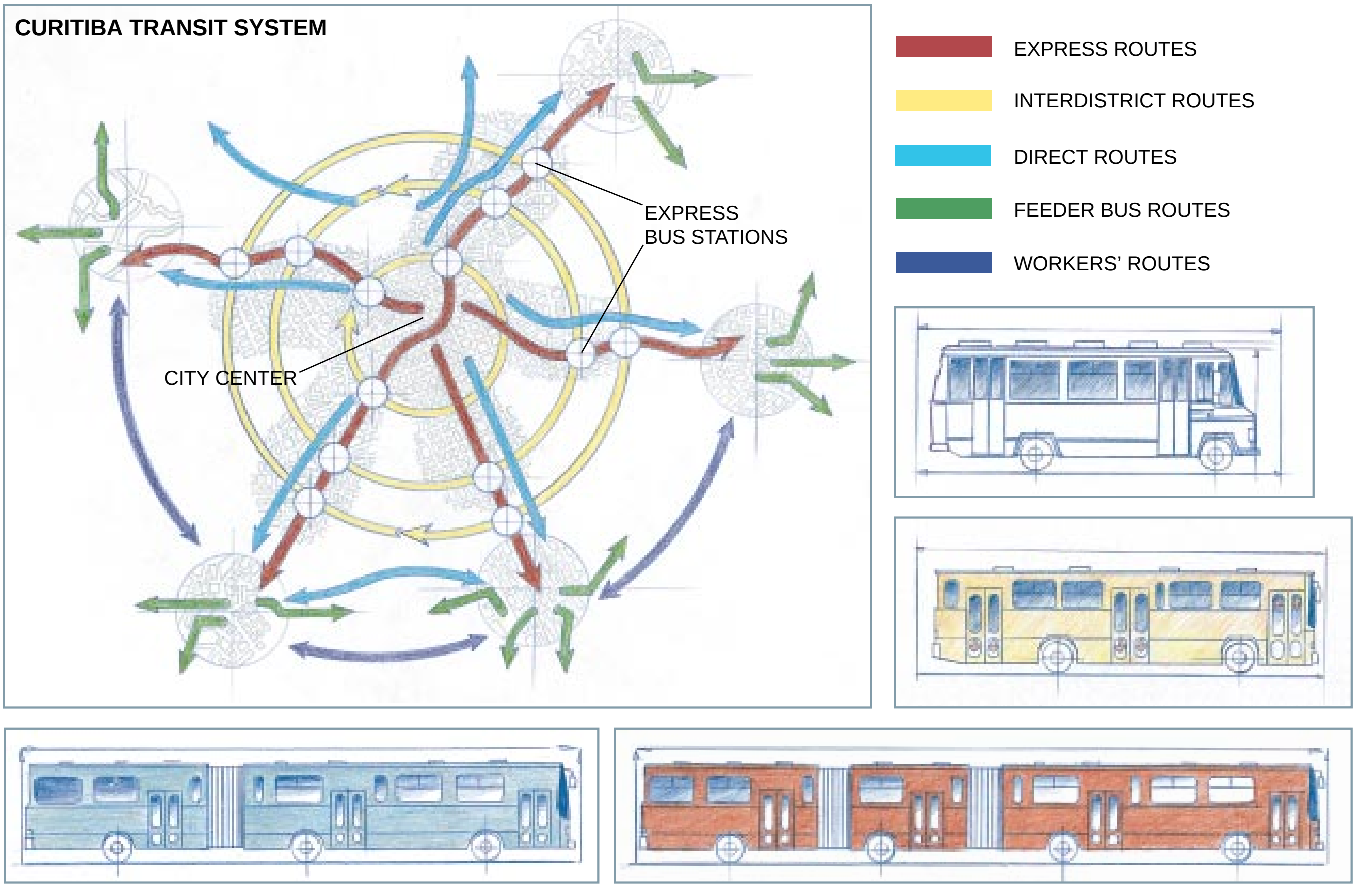
As a result of this system, average low-income residents of Curitiba spend only about 10 percent of their income on transport, which is relatively low for Brazil. Although the city has more than 500,000 private cars (more cars per capita than any Brazilian city except the capital, Brasília), three quarters of all commuters—more than 1.3 million passengers a day—take the bus. Per capita fuel consumption is 25 percent lower than in comparable Brazilian cities, and Curitiba has one of the lowest rates of ambient air pollution in the country. Although the buses run on diesel fuel, the number of car trips they eliminate more than makes up for their emissions.

In addition to these benefits, the city has a self-financing public transportation system, instead of being saddled by debt to pay for the construction and operating subsidies that a subway system entails. The savings have been invested in other areas. (Even old buses do not go to waste: they provide transportation to parks or serve as mobile schools.)

The implementation of the public transport system also allowed the development of a low-income housing program that provided some 40,000 new dwellings. Before implementing the public transport system, the city purchased and set aside land for low-income housing near the Curitiba Industrial City, a manufacturing district founded in 1972, located about eight kilometers west of the city center. Because the value of land is largely determined by its proximity to transportation and other facilities, these "land stocks" made it possible for the poor to have homes with ready access to jobs in an area where housing prices would otherwise have been unaffordable. The Curitiba Industrial City now supports 415 com-

JONAS RABINOVITCH

MAIN BOULEVARD of Curitiba, now devoted to pedestrian traffic, is the site of a weekly celebration of children gathering to paint. The ceremony began more pragmatically in 1972: when motorists threatened to ignore the traffic ban and drive on the street as usual, city workers blocked them by unrolling enormous sheets of paper and inviting children to paint watercolors.



KARL GUDE

BUS ROUTES have grown with the city. Express bus routes define Curitiba's spoke-shaped structural axes; interdistrict and local lines fill in the space between spokes. Each route is serviced by a bus of appropriate scale, from minibuses that carry 40 people on local trips to giant 270-passenger biarticulated vehicles used for express travel.

panies that directly and indirectly generate one fifth of all jobs in the city; polluting industries are not allowed.

Participation through Incentives

The city managers of Curitiba have learned that good systems and incentives are as important as good plans. The city's master plan helped to forge a vision and strategic principles to guide future developments. The vision was transformed into reality, however, by reliance on the right systems and incentives, not on slavish implementation of a static document.

One such innovative system is the provision of public information about land. City Hall can immediately deliver information to any citizen about the building potential of any plot in the city. Anyone wishing to obtain or renew a business permit must provide information to project impacts on traffic, infrastructure needs, parking requirements and municipal concerns. Ready access to this information helps to avoid land speculation; it has also been essential for budgetary purposes, because property taxes are the city's main source of revenue.

Incentives have been important in reinforcing positive behavior. Owners of land in the city's historic district can transfer the building potential of their plots to another area of the city—a rule that works to preserve historic buildings while fairly compensating their owners. In addition, businesses in specified areas throughout the city can "buy" permission to build up to two extra floors beyond the legal limit. Payment can be made in the form of cash or land that the city then uses to fund low-income housing.

Incentives and systems for encouraging beneficial behavior also work at the individual level. Curitiba's Free University for the Environment offers practical short courses at no cost for homemakers, building superintendents, shopkeepers and others to teach the environmental implications of the daily routines of even the most commonplace jobs. The courses, taught by people who have completed an appropriate training program, are a prerequisite for licenses to work at some jobs, such as taxi driving, but many other people take them voluntarily.

The city also funds a number of important programs for children, putting money behind the often empty pro-

JONAS RABINOVITCH

TRANSPORT NETWORK includes bicycle paths integrated with streets and the bus network for most efficient travel. The bicycle paths also connect the city's main parks.

Integrated Design Makes Busways Work

JONAS RABINOVITCH


JONAS RABINOVITCH

Curitiba's express bus system is designed as a single entity, rather than as disparate components of buses, stops and roads. As a result, the busways borrow many features from the subway system that the city might otherwise have built, had it a few billion dollars to spare. Most urban bus systems require passengers to pay as they board, slowing loading. Curitiba's raised-tube bus stops (*above*) eliminate this step: passengers pay as they enter the tube, and so the bus spends more of its time actually moving people from place to place.

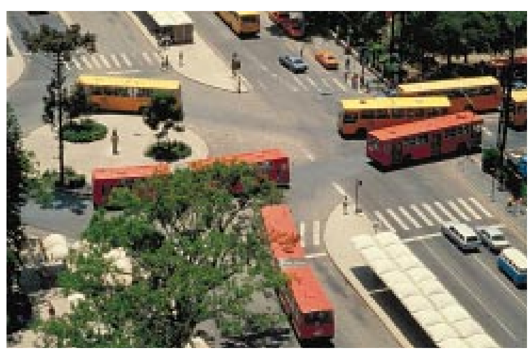
Similarly, the city installed wheelchair lifts at bus stops rather than onboard buses (*top right*), easing weight restrictions and simplifying maintenance—buses with built-in wheelchair lifts are notoriously trouble-prone, as are those that "kneel" to put their boarding steps within reach of the elderly. The tube-stop lifts also speed boarding by bringing disabled passengers to the proper height before the bus arrives.

Like subways, the buses have a track dedicated entirely to their use (*right*). This right-of-way significantly reduces travel time compared with buses that must fight automotive traffic to reach their destinations. By putting concrete and asphalt above the ground instead of excavating to place steel rails underneath it, however, the city managed to achieve most of the goals that subways strive for at less than 5 percent of the initial cost.

Some of the savings have enabled Curitiba to keep its fleet of 2,000 buses—owned by 10 private companies under contract to the city—among the newest in the world. The average bus is only three years old. The city pays bus owners 1 percent of the value of a bus each month; after 10 years it takes possession of retired vehicles and refurbishes them as free park buses or mobile schools [*see middle left illustration on next page*].

Companies are paid according to the length of the routes they serve rather than the number of passengers they carry, giving the city a strong incentive to provide service that increases ridership (*bottom right*). Indeed, more than a quarter of Curitiba's automobile owners take the bus to work. In response to increased demand, the city has augmented the capacity of its busways by using extra-long buses—the equivalent of multicar subway trains. The biarticulated bus, in service since 1992, has three sections connected by hinges that allow it to turn corners. At full capacity, these vehicles can carry 270 passengers, more than three times as many as an ordinary bus. —*J.R.*


JONAS RABINOVITCH


JONAS RABINOVITCH


JONAS RABINOVITCH


JONAS RABINOVITCH


JONAS RABINOVITCH


JONAS RABINOVITCH

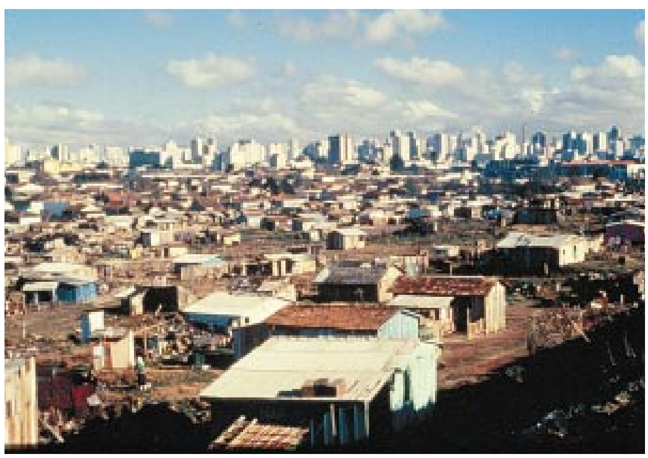
RECYCLING in Curitiba takes many forms. As in many other cities, families sort their garbage to ease recovery of glass, metal and plastic (*top left*). In addition, old buses find second and third careers as free transportation to city parks (*middle left*) or as mobile offices and classrooms (*bottom left*). Even the city's old electrical utility poles find new life as parts of park buildings and public offices, including the Free University for the Environment (*above*).

nouncements municipalities make about the importance of the next generation. The Paperboy/Papergirl Program gives part-time jobs to schoolchildren from low-income families; municipal day care centers serve four meals a day for some 12,000 children; and SOS Children provides a special telephone number for urgent communications about children under any kind of threat.

Curitiba has repeatedly rejected conventional wisdom that emphasizes technologically sophisticated solutions to urban woes. Many planners have contended, for example, that cities with over a million people must have a subway system to avoid traffic congestion. Prevailing dogma also claims that cities that generate more than 1,000 tons of solid waste a day need expensive mechanical garbage-separation plants. Yet Curitiba has neither.

The city has attacked the solid-waste issue from both the generation and collection sides. Citizens recycle paper equivalent to nearly 1,200 trees each day. The Garbage That Is Not Garbage initiative has drawn more than 70 percent of households to sort recyclable materials for collection. The Garbage Purchase program, designed specifically for low-income areas, helps to clean up sites that are difficult for the conventional waste-management system to serve. Poor families can exchange filled garbage bags for bus tokens, parcels of


JONAS RABINOVITCH


JONAS RABINOVITCH

POOR DISTRICTS will probably always be part of fast-growing cities; Curitiba's garbage-purchase program, which pays 40,000 families in bus tokens or food in exchange for waste from areas that conventional sanitation services cannot reach, has at least mitigated some of the unsanitary conditions that usually prevail (picture at the left was taken before the garbage purchases began). A school-based garbage-exchange plan also supplies poor students with notebooks.

surplus food and children's school notebooks. More than 34,000 families in 62 poor neighborhoods have exchanged over 11,000 tons of garbage for nearly a million bus tokens and 1,200 tons of surplus food. During the past three years, students in more than 100 schools have traded nearly 200 tons of garbage for close to 1.9 million notebooks. Another initiative, All Clean, temporarily hires retired and unemployed people to clean up specific areas of the city where litter has accumulated.

These innovations, which rely on public participation and labor-intensive approaches rather than on mechanization and massive capital investment, have reduced the cost and increased the effectiveness of the city's solid-waste management system. They have also conserved resources, beautified the city and provided employment.

JONAS RABINOVITCH

BOTANICAL GARDENS were once a city dump. In addition to providing space for recreation, the gardens serve as a research center for studies of plant compounds.

Lessons for an Urbanizing World

No other city has precisely the combination of geographic, economic and political conditions that mark Curitiba. Nevertheless, its successes can serve as lessons for urban planners in both the industrial and the developing worlds.

Perhaps the most important lesson is that top priority should be given to public transport rather than to private cars and to pedestrians rather than to motorized vehicles. Bicycle paths and pedestrian areas should be an integrated part of the road network and public transportation system. Whereas intensive road-building programs elsewhere have led paradoxically to even more congestion, Curitiba's slighting of the needs of private motorized traffic has generated less use of cars and has reduced pollution.

Curitiba's planners have also learned that solutions to urban problems are not specific and isolated but rather interconnected. Any plan should involve partnerships among private-sector entrepreneurs, nongovernmental organizations, municipal agencies, utilities, neighborhood associations, community groups and individuals. Creative and labor-intensive ideas—especially where unemployment is already a problem—can often substitute for conventional capital-intensive technologies.

We have found that cities can turn traditional sources of problems into resources. For example, public transport, urban solid waste, and unemployment are traditionally considered problems, but they have the potential to become generators of new resources, as they have in Curitiba.

Other cities are beginning to learn some of these lessons. In Brazil and elsewhere in Latin America, the pedestrian streets that Curitiba pioneered have become popular urban fixtures. Cape Town has recently developed a new vision for its metropolitan area that is explicitly based on Curitiba's system of structural axes. Officials and planners from places as diverse as New York City, Toronto, Montreal, Paris, Lyons, Moscow, Prague, Santiago, Buenos Aires and Lagos have visited the city and praised it.

As these planners carry Curitiba's examples back to their homes, they also come away with a crucial principle: there is no time like the present. Rather than trying to revitalize urban centers that have begun falling into decay, planners in already large cities and those that have just started to grow can begin solving problems without waiting for top-down master plans or near fiscal collapse.

The Authors

JONAS RABINOVITCH and JOSEF LEITMAN are urban planners, Rabinovitch at the United Nations and Leitman at the World Bank. Rabinovitch earned his bachelor's degree in architecture and urban planning from the Federal University of Rio de Janeiro and received a master's degree in the economics of urban development from University College London, with a specialization in urban traffic and transport planning. Before coming to the U.N. three years ago, he was an adviser to the mayor and director of international relations for Curitiba, having joined the city's research and urban planning institute in 1981. Leitman is a senior urban planner at the World Bank. He received his doctorate in city and regional planning from the University of California, Berkeley, in 1992. He earned bachelor's and master's degrees from Harvard University. This article reflects the opinions of the authors, not necessarily those of the city of Curitiba, the United Nations Development Program or the World Bank.

Further Reading

CURITIBA: TOWARDS SUSTAINABLE URBAN DEVELOPMENT. Jonas Rabinovitch in *Environment and Urbanization,* Vol. 4, No. 2, pages 62–73; October 1992.

ENVIRONMENTAL INNOVATION AND MANAGEMENT IN CURITIBA, BRAZIL. Jonas Rabinovitch with Josef Leitman. United Nations Development Program, Habitat and World Bank. Urban Management Program, Working Paper Series No. 1; June 1993.

JANE'S URBAN TRANSPORT SYSTEMS. Edited by Chris Bushell. Jane's Information Group, 1995.

A SUSTAINABLE URBAN TRANSPORTATION SYSTEM: THE "SURFACE METRO" IN CURITIBA, BRAZIL. J. Rabinovitch and J. Hoehn. Environmental and Natural Resources Policy and Training Project, Working Paper No. 19; May 1995 (ISSN 1072-9496). For information, contact J. Rabinovitch via fax at (212) 906-6793.

Collisions with Comets and Asteroids

The chances of a celestial body colliding with the earth are small, but the consequences would be catastrophic

by Tom Gehrels

Are we going to be hit by an asteroid? Planetary scientists are divided on how worrisome the danger is. Some refuse to take it seriously; others believe the risk of dying from such an impact might even be greater than the risk of dying in an airplane crash. After years of studying the problem, I have become convinced that the danger is real. Although a major impact is unlikely, the energies released could be so horrendous that our fragile society would be obliterated.

Early in our planet's history, asteroids and comets made life possible by accreting into the earth and then by bringing water to the newborn planet. And they have already destroyed, at least once, an advanced form of that life. The dinosaurs were killed by such an impact, making way for the age of the mammals. Now for the first time, creatures have evolved to a point where they can wrest control of their fate from the heavenly bodies, but humans must come to grips with the danger.

Some four and a half billion years ago, the solar system formed out of a swirling cloud of gas and dust. Initially the planetesimals—coarse collections of rocky materials—coagulated, merging with one another to create planets. Because of the energy released by the colliding rocks, the earth began as a molten globe, so hot that the volatile substances—water, carbon dioxide, ammonia, methane and other gases—boiled off. As the material of the inner solar nebula was mopped up by the growing planets, the bombardment of the earth slowed. The glowing planet cooled, and a crust solidified. Only then did water—the life-giving fluid that covers three quarters of the earth's surface—return, borne on cold comets arriving from the solar system's distant reaches. Fossil records show that simple life-forms started evolving almost right away.

Comets and asteroids are, in fact, leftover planetesimals. Most asteroids inhabit the vast belt between the orbits of Mars and Jupiter. Being quite close to the sun, they were formed hot; as on the early earth, the high temperatures vaporized the lighter substances, such as water, leaving mostly silica, carbon and metals. (Only recently have astronomers found some rare asteroids that contain crystalline water embedded in rocks.)

Comets, on the other hand, hover at the outer edges of the solar system. As the solar system was formed, a good deal of matter was thrown outward, beyond the orbits of Uranus and Neptune. Coalescing far from the sun, the comets were born cold, at temperatures as low as –260 degrees Celsius. They retained their volatile materials, the gas, ice and snow. Sometimes called dirty snowballs, these objects are usually tenuous aggregates of carbon and other light elements.

Fiery Visitors

In 1950 Jan H. Oort, professor of astronomy at Leiden University in the Netherlands, was teaching a class that I was allowed to attend as an undergraduate. While reviewing astronomical calculations for his students, Oort noted that a number of known comets reach their farthest point from the sun—called the aphelion—at a great distance. He went on to formulate the idea that a cloud of comets exists as a diffuse spherical shell at about 50,000 or more astronomical units. (One astronomical unit is the distance from the earth to the sun.) This distant cloud, containing perhaps some 10^{13} objects, envelops the solar system.

The Oort cloud reaches a fifth of the distance to the nearest star, Alpha Centauri. Inhabitants of this shell are thus loosely bound to the sun and readily disturbed by events beyond the solar system. If the sun passes by another star or a massive molecular cloud, some of these cometary orbits are jarred. The planetesimal might then swing into a narrow elliptical orbit that brings it toward the inner solar system. As it nears the sun, the heat vaporizes its volatile materials, which spew forth as if from a geyser. In ancient cultures, this celestial spectacle was sometimes an ominous event.

Some visitors from the Oort cloud are never seen again; others have periods that get shorter with each successive pass. The best known of these comets are those that return regularly, such as Halley's, with a period of 76 years. The chance that such a comet will collide with the earth is exceedingly small, because it comes by so infrequently. But the patterns of their orbits suggest that in the next millennia, comet Halley or Swift-Tuttle (with a period of 130 years) will sometimes swing by too close for comfort.

In 1951 Gerard P. Kuiper, then at Yerkes Observatory of the University of Chicago, surmised that another belt of comets exists, just beyond Neptune's orbit, much nearer than the Oort cloud. Working at the University of Hawaii, David C. Jewitt and Jane Luu discovered the first of these objects in 1992 after a persistent search; by now some 31 bodies belonging to the Kuiper belt have been found. In fact, Pluto, with its unusually elliptical orbit, is now considered to be the largest of these objects; Clyde Tombaugh, who discovered Pluto in 1930, calls it the "King of the Kuiper belt."

Comets belonging to the Kuiper belt are not directly disturbed by rival stars. Instead they can stray close to Neptune, which may either help stabilize them or, conversely, throw them out of orbit. (An as yet unknown 10th planet may also be stirring the comets' path, but the evidence for its existence is inconclusive.) The comets may then come very close to the sun. Although those from the Kuiper belt tend to have shorter periods than those from the Oort cloud, both types of comets can be captured in tight orbits around the sun. It is therefore impossible to tell where a particular comet—such as Tempel-Tuttle, which sweeps by at 72 kilometers per second every 33 years—originated from.

Some comets are bound into small orbits and have short periods, on the order of 10 years. These comets pose more of a concern than the ones that come by only every century or so. A collision with such a short-period comet might occur once in some three million years.

However infrequent a cometary collision might be, the consequences would be calamitous. The orbits of comets are often steeply inclined to the earth's; oc-

F. GOHIER *Explorer*

METEOR CRATER in northern Arizona, a depression 1.2 kilometers in diameter, was carved out by an asteroid that struck the earth 50,000 years ago. The asteroid was only 30 meters wide but, being metallic, was strong enough to penetrate the atmosphere without disintegrating. The earth collides with an object of this size or larger once in a century.

casionally, a comet is even going in the opposite direction. Thus, comets typically pass the earth with a high relative velocity. For example, Swift-Tuttle, which is about 25 kilometers across, flies by at 60 kilometers per second. It would impact with cataclysmic effect.

Unless it runs into something, a comet probably remains active, emitting gases and dust for some 500 passages by the sun. Eventually, the volatile materials are used up, and the comet fades away as a dead object, indistinguishable from an asteroid. Up to half of the nearest asteroids might in fact be dead, short-period comets.

Falling Rocks

Indeed, most of the danger to the earth comes from asteroids. Like comets, asteroids have solar orbits that are normally circular and stable. But there are so many of them in the asteroid belt that they can collide with one another.

The debris from such collisions can end up in unstable orbits that resonate with the orbit of Jupiter. By virtue of its immense mass, Jupiter competes with the sun for control of the motions of these fragments, especially if an asteroid's orbit "beats," or resonates, with that of the giant planet. So, for instance, if the asteroid goes around the sun thrice in the same time that Jupiter orbits once, the planet's gravitational influence on the rock is greatly enhanced. Just as a child on a swing flies ever higher if someone pushes her each time the swing returns, Jupiter's rhythmic nudges ultimately cause the asteroid to veer out of its original orbit into an increasingly eccentric one.

The asteroid may either leave the solar system or move in toward the terrestrial, rocky planets. Eventually, such vagrants collide with Mars, the earth-moon system, Venus, Mercury or even the sun. A major fragment enters the inner solar system once in roughly 10 million years and survives for about as long.

To estimate the chances of such a rock hitting the earth, the asteroids have first to be sorted according to size. The smallest ones we can observe, which are less than a few tens of meters across, rarely make it through the earth's atmosphere; friction with air generates enough heat to vaporize them. The asteroids that are roughly 100 meters and larger in diameter do pose a threat. There are 100,000 or so of these that penetrate the inner solar system deeper than the orbit of Mars. They are called near-earth asteroids.

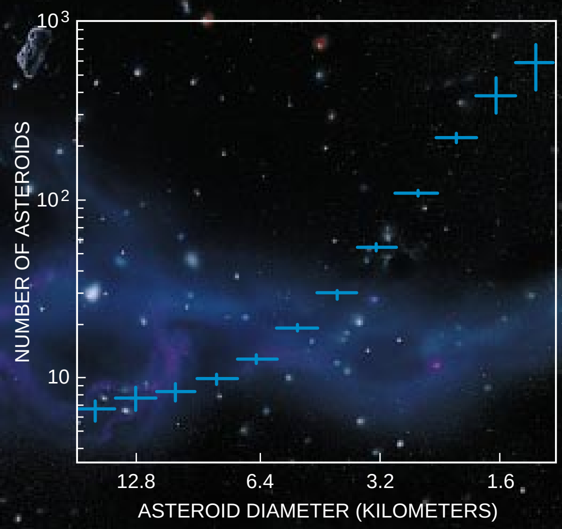
In 1908 one such object, a loose conglomerate of silicates about 60 meters wide, entered the atmosphere and burst apart above the Tunguska Valley in Siberia. The explosion was heard as far away as London. Although the fragments did not leave a crater, the area below the explosion is still marked by burnt trees laid out in a region roughly 50 kilometers across. The identity of the Tunguska object inspired a lot of nonsensical speculation for decades, and some highly imaginative suggestions were made, including that it was a mini–black hole or an alien spacecraft. Scientists, however, have always understood that it was a comet or asteroid.

Events such as the Tunguska explosion may occur once a century, and it is most likely that they would occur over the oceans or remote land areas. But they would be devastating if they happened near a populated area. If one exploded over London, for instance, not only the city but also its suburbs would be laid waste.

Of the smaller asteroids, the few metallic ones are tough enough to penetrate the atmosphere and carve out a crater. The 1.2-kilometer-wide Meteor Crater in northern Arizona is an example; it came from a metallic asteroid about 30 meters in diameter that fell some 50,000 years ago.

The Threat from Asteroids

The asteroid belt, home of most asteroids, lies between the orbits of Mars and Jupiter. The chart (*below*), put together from Spacewatch observations, shows that smaller asteroids, produced by fragmentation of the larger ones, are more numerous. The rocks normally remain in circular, stable orbits, but collisions, along with the gravitational influence of Jupiter, can throw them into narrow, unstable orbits. Then the asteroids may enter the inner solar system, where they pose a threat to the earth.



ROBERT JEDICKE *University of Arizona*

ALFRED T. KAMAJIAN (*illustration*); JOHNNY JOHNSON (*graph*)

An even greater peril is posed by the 1,000 or 2,000 medium near-earth asteroids that are roughly one kilometer and larger in size. One of these asteroids is thought to collide with the earth once in about 300,000 years. Note that this estimate is only a statistical average. Such a collision can happen at any time—a year from now, in 20 years or not in a million years.

Frightful Darkness

The energies liberated by an impact with such an object would be tremendous. The kinetic energy can be calculated from $1/2\ mv^2$, where m is the mass of the object, and v is the incoming velocity. Assuming a density of about three grams per cubic centimeter, as known from meteorites, and an average velocity of 20 kilometers per second, a one-kilometer-wide object would strike with a shock equivalent to tens of billions of tons of TNT—millions of times the energy released at Hiroshima in 1945.

Granted, asteroids do not emit the nuclear radiation that caused the particular horrors of Hiroshima. Still, an explosion of millions of Hiroshimas would do more than destroy a few cities or some countries. The earth's atmosphere would be globally disrupted, creating the equivalent of a nuclear winter. Large clouds of dust would explode into the atmosphere to obscure the sun, leading to prolonged darkness, subzero temperatures and violent windstorms.

Even more dangerous are the largest near-earth asteroids, which are about 10 kilometers in diameter. Fortunately, there are only a few such threatening objects, perhaps just 10. (Even more fortunately, they happen to be mere fragments of the objects in the asteroid belt, which can be as large as 1,000 kilometers across.) An asteroid of this size collides with the earth only once in 100 million years or so.

One such event is evident in the fossil record. The impact of a celestial object marks the end of the Cretaceous geologic period and the beginning of the Tertiary, 65 million years ago. After years of searching, the crater from that event—a depression about 170 kilometers in diameter—has been identified in the Yucatán Peninsula of Mexico. Although the crater cannot be directly seen, it has fortuitously been identified by drillings for oil and in images taken from the space shuttle *Endeavour*. The depression resulted from the explosive impact of an object perhaps 10 to 20 kilometers in diameter.

Studies of the effects of that explosion paint a frightening picture. An enormous fireball ejected rocks and steam into the atmosphere, jarred the earth's crust and triggered earthquakes and tsunamis around the globe. Vast clouds of dust, from the earth and the asteroid, erupted into the stratosphere and beyond. There ensued total darkness, which lasted for months.

Acid rain began to fall, and slowly the dust settled, creating a layer of sediment a few centimeters thick over the earth's surface. Below this thin sheet we see evidence of dinosaurs. Above it they are missing, as are three fourths of the other species. The darkness following the explosion must have initially plunged the atmosphere into a freeze.

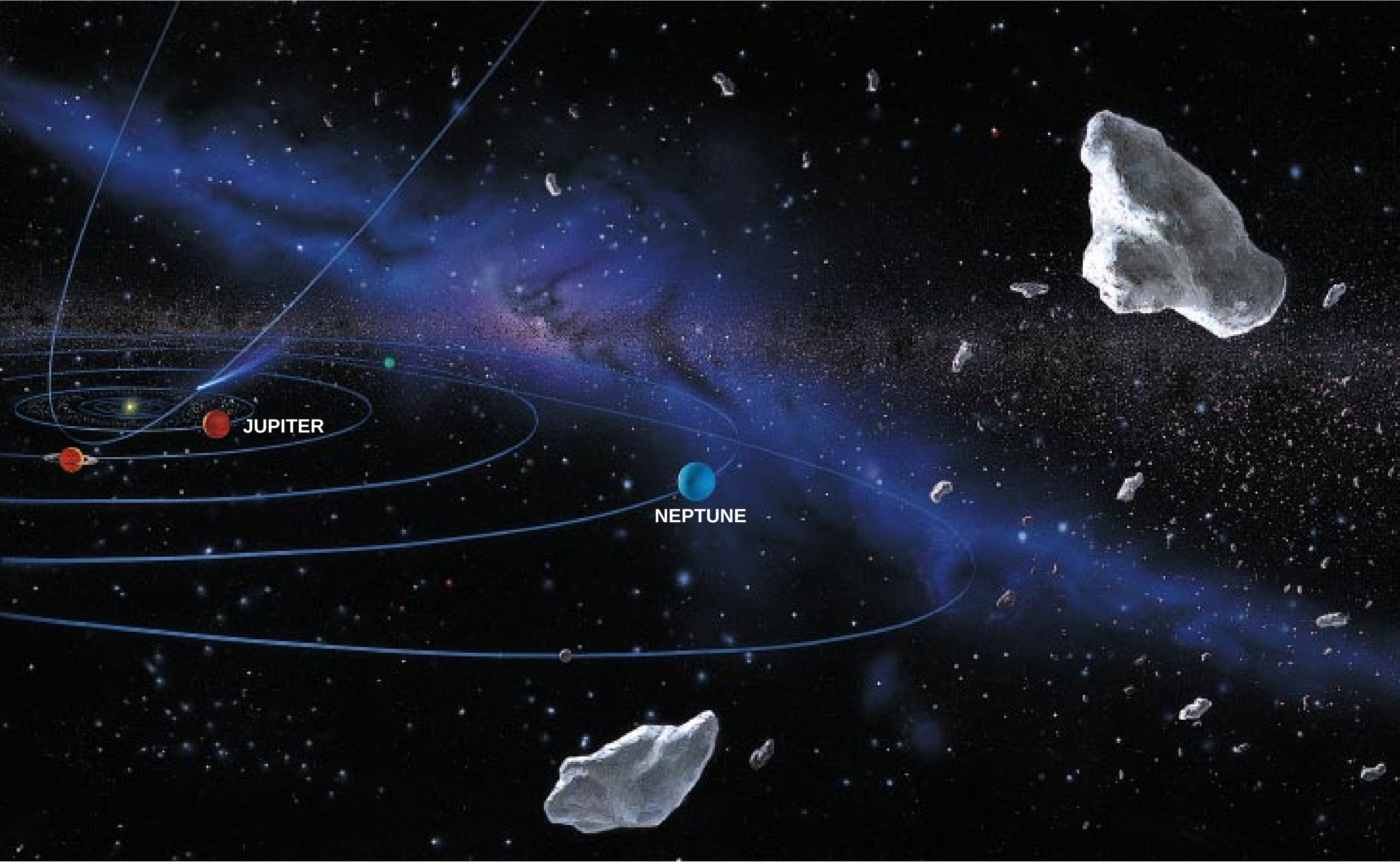
Over many centuries, the reverse effect—a slow greenhouse warming, by as much as 15 degrees Celsius—had an equally devastating outcome. The asteroid had struck the earth in a vulnerable place, slicing into a rare region with a deep layer of limestone. (Less than 2 percent of the earth's crust has so much limestone; Australia's Great Barrier Reef is an example.) The explosion ejected the carbon dioxide from the limestone into the atmosphere, where, along with other gases, it helped to trap the earth's heat. Jan Smit of the Free University, Amsterdam, has proposed that the severe warming, rather than the initial freeze, killed the dinosaurs—there is some evidence that they died off slowly.

Spacewatch

So—are we going to be hit? To begin with, the answer lies in the domain of planetary astronomy. The dangerous objects have to be located, as soon as possible, to diminish the chances of our unexpected demise. Furthermore, they have to be tracked on the succeeding nights, weeks, months and even years so that their orbits can be accurately extrapolated into the future.

In the early 1970s a 0.46-meter photographic camera at the Palomar Observatory in southern California was dedicated to the search for near-earth objects. Eleanor Helin of the Jet Propulsion Laboratory in Pasadena, Calif., led one of the teams of astronomers, and Eugene M. and Carolyn S. Shoemaker of the U.S. Geological Survey led the other. The scientists photographed the same large areas of the sky at half-hour intervals. As asteroids orbit the sun, they move with respect to the background stars. If near to the earth, the asteroid is seen to travel relatively fast; the motion is easily recognized from the multiple exposures.

Since the pioneering efforts at Palo-

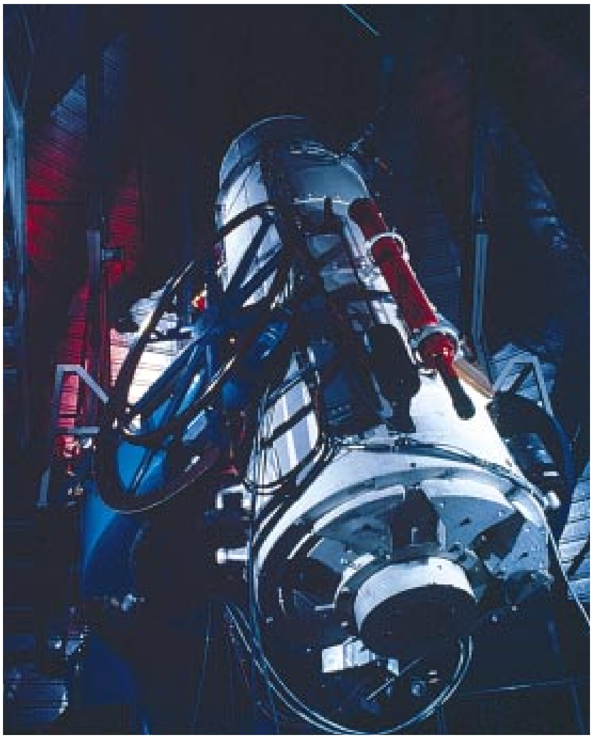


mar, other observers have become interested in near-earth asteroids. At Siding Spring in the mountains of eastern Australia, a dedicated group of scientists uses a 1.2-meter photographic camera to hunt for these rocks. In 1994 observers in California and Australia, with their photographic methods, jointly discovered 16 near-earth asteroids. (At the end of that year, the Palomar project closed as more modern techniques were developed elsewhere.)

About 15 years ago Robert S. McMillan, also at the University of Arizona, and I began to realize that at this rate, it would take more than a century to map the 1,000 or more asteroids that are larger than one kilometer across. By taking advantage of electronic detection devices and fast computers, the rate of finding asteroids could be greatly increased. Spacewatch, a project dedicated to the study of comets and asteroids, was born in Tucson. A 0.9-meter telescope at the University of Arizona's Steward Observatory on Kitt Peak, 70 kilometers west of Tucson, is now dedicated to Spacewatch. Robert Jedicke, James V. Scotti, several students and I, all from Tucson, use this facility regularly for finding comets and asteroids. McMillan, Marcus L. Perry, Toni L. Moore and others, also from Tucson, use it for finding planets around other stars.

Instead of photographic plates, our electronic light detectors are charge-coupled devices, or CCDs. These are finely divided arrays of semiconductor picture elements, or pixels. When light hits a pixel, its energy causes positive and negative electrical charges to separate. The electrons from all the pixels provide an image of the light pattern at the focal plane of the telescope. A computer then compares images of the same patch of sky scanned at different times, marking the objects that have moved.

In this manner, Spacewatch observers may find as many as 600 asteroids a night. Most of these are in the asteroid belt; only occasionally does an object move against the star field so fast that it must be close to the earth. (Similarly, an airplane high above in the sky seems to move slower than one coming in low for a landing.) In 1994 Scotti found an asteroid that


UNIVERSITY OF ARIZONA

SPACEWATCH telescope on top of Kitt Peak in southern Arizona is dedicated to searching for comets and asteroids.

The Threat from Comets

Comets reside beyond the orbit of Neptune in the Kuiper belt and the Oort cloud and, like asteroids, come near the earth only when dislodged from their circular paths. The Kuiper belt probably merges into the Oort cloud, which extends a fifth of the distance to the nearest star, Alpha Centauri. Comet Halley (*below*) is a visitor from the Oort cloud that has swung into a steeply elliptical orbit around the sun, having a period of 76 years.

ALFRED T. KAMAJIAN (*illustration*); ROYAL GREENWICH OBSERVATORY *Science Photo Library/Photo Researchers, Inc.* (*photograph*)

passed within 105,000 kilometers of the earth. Also in that year, Spacewatch reported 77,000 precise measurements of comet and asteroid positions. One gratifying aspect of Spacewatch is that it has private and corporate supporters (currently 235) in addition to the U.S. Air Force Office of Scientific Research, the National Aeronautics and Space Administration, the Clementine space program, the National Science Foundation and other governmental organizations.

Spacewatch has discovered an abundance of small asteroids, those in the range of tens of meters. The numbers of these objects exceed predictions by a factor of 40, but we do not as yet understand their origins. These asteroids we call the Arjunas, after the legendary Indian prince who was enjoined to persist on his charted course. Military reconnaissance satellites have since also observed the Arjunas. The data, once routinely discarded but now stored and declassified, show the continuous showering of the planet by small asteroids. Because of the atmosphere, these rocks burn up with little consequence, even though similar ones scar the airless moon.

The next step for Spacewatch is to install our new telescope, which was built with an existing 1.8-meter mirror, so that we can find fainter and more distant objects. This state-of-the-art instrument, the largest in the field of asteroid observation, should serve generations of explorers to come. Meanwhile, at Côte d'Azur Observatory in southern France, Alain Maury is about to bring a telescope into operation with an electronic detection system. Duncan Steel and his colleagues in Australia are switching to electronics as well, although this project has funding problems perhaps more severe than ours. Next to join the electronic age might be Lowell Observatory near Flagstaff, Ariz., under the supervision of Edward Bowell. The U.S. Air Force is also planning to use one of its one-meter telescopes to this end; Helin and her associates already use the one on Maui in Hawaii. And amateur astronomers are coming on-line with electronic detectors on their telescopes.

If there is an asteroid out there with our name on it, we should know by about the year 2008.

Deflecting an Asteroid

And what if we find a large object headed our way? If we have only five years' notice, we can say good-bye to one another and regret that we did not start surveying earlier. If we have 10 years or so, our chances are still slim. If we have 50 years' notice or more, a spacecraft could deploy a rocket that would explode near the asteroid. Perhaps the most powerful intercontinental ballistic missiles could blast a small object out of the way. (That, incidentally, would also be a good means of getting rid of these relics from the cold war.)

It seems likely, however, that we will have more than 100 years to prepare. Given that much time, a modest chemical explosion near an asteroid might be enough to deflect it. The explosion will need to change the asteroid's trajectory by only a small amount so that by the time the asteroid reaches the earth's vicinity, it will have deviated from its original course enough to bypass the planet.

Present technology for aiming and guiding rockets is close to miraculous. I once overheard two scientists arguing about why *Pioneer 11* had arrived 20 seconds late at Saturn—after a journey of six years. But the detonation will have to be carefully designed. If the asteroid is made of loosely aggregated material, it might disintegrate when shaken by an explosion. The pieces could rain down on the earth, causing even greater damage than the intact asteroid, as hunters who use buckshot know. A "standoff" explosion, at some distance from the surface, may be the most effective in that case. Earth-based radar, telescopes and possibly space missions will be needed to determine the composition of an asteroid and how it might break up.

Further into the future, laser or microwave devices might become suitable. Gentler alternatives, such as solar sails and reflectors planted on the asteroid's surface—to harness the sun's radiation in pushing the asteroid off course—have also been suggested. A few scientists are studying the feasibility of nuclear devices to deflect very massive asteroids that show up at short notice.

Comets and asteroids remind me of Shiva, the Hindu deity who destroys and re-creates. These celestial bodies allowed life to be born, but they also killed our predecessors, the dinosaurs. Now for the first time, the earth's inhabitants have acquired the ability to envision their own extinction—and the power to stop this cycle of destruction and creation.

The Author

TOM GEHRELS was inspired to study celestial objects upon attending a class given by Jan Oort in the Netherlands, who surmised the existence of a distant shell of comets now called the Oort cloud. Gehrels is professor of planetary sciences at the University of Arizona at Tucson, a Sarabhai Professor at the Physical Research Laboratory in India and principal investigator of the Spacewatch program at Kitt Peak, Ariz., where he hunts for comets and asteroids.

Further Reading

THE ORIGINS OF THE ASTEROIDS. Richard P. Binzel, M. Antonietta Barucci and Marcello Fulchignoni in *Scientific American,* Vol. 265, No. 4, pages 88–94; October 1991.

HAZARDS DUE TO COMETS AND ASTEROIDS. Edited by T. Gehrels. University of Arizona Press, 1994.

ROGUE ASTEROIDS AND DOOMSDAY COMETS. Duncan Steel. John Wiley & Sons, 1995.

The African AIDS Epidemic

In parts of sub-Saharan Africa, nearly 25 percent of the population is HIV-positive as a result of heterosexual transmission of the virus. Could lack of circumcision make men in this region particularly susceptible?

by John C. Caldwell and Pat Caldwell

AIDS has swept across sub-Saharan Africa on an extraordinary scale. Two thirds of the roughly 16 million people in the world infected with the human immunodeficiency virus (HIV), which causes AIDS, live there. Half of the world's cases are found in what we call the AIDS belt—a chain of countries in eastern and southern Africa that is home to only 2 percent of the global population.

Heterosexual intercourse serves as the main vehicle for spreading HIV throughout sub-Saharan Africa. This is in stark contrast to the developed world, where the virus is most frequently transmitted during homosexual intercourse or when intravenous drug users share contaminated syringes. Attempts to halt the flood of AIDS cases in Africa will not succeed until researchers can determine what factors contribute to the remarkable prevalence of the disease among heterosexuals. Such a diagnosis will also help us figure out how likely it is that heterosexual epidemics might extend into Asia or the West.

One frequently mentioned explanation for the severe epidemic in the AIDS belt is that the virus originated here and continues to move outward from an epicenter of disease. But AIDS cases appeared in hospitals in Uganda and Rwanda at the same time they did in the West, and no stored human-tissue samples taken from Africans during the 1970s are HIV-positive. Furthermore, the AIDS belt is not circular but elongated, clearly not the pattern of expansion from an epicenter. (A related virus, HIV-2, most likely did originate in Africa, but it infects fewer people and kills much more slowly; for these reasons we do not deal with it in this article.)

To determine what factors might be spurring the rapid progress of HIV in sub-Saharan Africa, we decided to reexamine everything we knew, and thought we knew, about the epidemic. Were we sure that HIV was transmitted primarily by heterosexual intercourse? Were there differences between behavior in the AIDS belt and in the rest of the region that might account for the severity of the epidemic in certain countries? Could susceptibility to AIDS be linked to other health problems that were common in the heavily infected areas?

We were in a good position to study the African AIDS epidemic, because for more than three decades we have examined family dynamics, fertility and fertility control in sub-Saharan Africa. In the late 1970s we also worked on sexuality and sexually transmitted diseases there. And since 1989 we have studied the epidemiological, social and behavioral aspects of the African AIDS epidemic. Our collaborators in Africa have been, since 1989, I. O. Orubuloye of Ondo State University in Nigeria and the West African Research Group on Sexual Networking and, since 1991, James P. Ntozi and Jackson Mukiza-Gapere, both at Makerere University in Uganda, as well as John K. Anarfi of the University of Ghana and Kofi Awusabo-Asare of the University of Cape Coast in Ghana.

Heterosexual Transmission

The first assumption we had to scrutinize was the notion that AIDS in sub-Saharan Africa spreads primarily through heterosexual intercourse. We were skeptical because elsewhere the risk of acquiring the virus during heterosexual sex is extremely low. If a man and a woman are otherwise healthy except for the fact that one is HIV-positive, then in a single act of unprotected vaginal intercourse, the chance of transmission from the man to the woman is one in 300 and from the woman to the man, possibly as low as one in 1,000. This level of risk contrasts sharply with the much greater likelihood of infection during unprotected anal intercourse; when sharing needles during drug use; or from a transfusion of infected blood. These means of transfer are sufficiently risky to sustain an epidemic among small segments of a population—among homosexual men and intravenous drug users, for example. But these methods cannot sustain a society-wide epidemic in the manner that heterosexual transmission would allow.

Despite our initial skepticism, evidence for a heterosexual epidemic in Africa is convincing. The most careful studies have shown that the infection rate among females is probably 1.2 times higher than the infection rate among males and that most HIV-positive women caught the virus from their spouse. In the West, however, the number of infected men (who are more likely than women to contract HIV from intravenous drug use or homosexual intercourse) is five to 10 times greater than the number of women.

Additional studies have ruled out other typical methods of transmission in most of Africa. For instance, we have found that because anal intercourse is considered abhorrent for a variety of reasons, including its connection with witchcraft, the practice is almost completely suppressed in much of sub-Saharan Africa. Furthermore, intravenous drugs are seldom used there: marijuana is widely consumed, but injected opiates are too expensive for these impoverished societies.

Many researchers in the West have assumed that the heterosexual AIDS epidemic reflects unusual sexual practices that facilitate transmission. But by global standards, sexual activity in sub-Saharan Africa is fairly simple: even in commercial sex, there is little foreplay or violence that can cause bleeding. We did worry, however, that perhaps other traditions might make even straightforward intercourse unusually dangerous. For example, in some parts of Africa, women apply astringents such as alum (long used in the West to dry blood from shaving cuts), cloth or leaves to dry their vaginas, according to local traditional male demands. Vaginal drying is also used to remove discharge caused by infections, which occur frequently in tropical situations where hygiene is difficult and medication rare. Such methods of

GEORGE MULALA

ENTIRE FAMILIES in sub-Saharan Africa, including this one in Nairobi, Kenya, feel the devastating impact of AIDS. The wife (*right*) was probably infected with HIV by her husband (who refused to be photographed). All three children have the disease; a fourth daughter has already died of AIDS. Kenya lies in the part of Africa the authors call the AIDS belt, where a heterosexual epidemic may be sustained in part because of lack of circumcision among men in the region.

drying can scratch or alter the vaginal wall, making it more susceptible to bleeding during intercourse and possibly rendering HIV infection more likely. Yet we have not found any evidence linking these practices with an increased risk for contracting HIV.

Role of Circumcision Suggested

Although we were initially stymied in our search for what might be fueling the heterosexual epidemic, we found a new lead in 1989. A joint Canadian-Kenyan medical research team working at Kenyatta Medical School in Nairobi, where the epidemic is intense, had reported a year earlier that AIDS rates were higher among Luo migrants from western Kenya than among the Kikuyu from central Kenya. At first, the researchers assumed that the Luo ethnic group was at greater risk of the disease because they migrated from an area that is close to Uganda, which the researchers believed might be the epicenter of the HIV epidemic.

But as the epicenter idea became less tenable, the researchers proposed an alternative explanation: they surmised that the Luo were at greater risk because, unlike the Kikuyu, they were not circumcised. Apparently, uncircumcised Luo men were more likely to have chancroid—a sexually transmitted disease characterized by soft sores on the genitals—or syphilis. These men also had an unexpectedly elevated risk of contracting HIV [*see box on page 66*]: in the capital city of Nairobi, a Luo man with chancroid who had sex once with an HIV-positive prostitute had a 50 percent chance of contracting the virus.

Drawing on these findings, in 1989 an

Continued on page 66

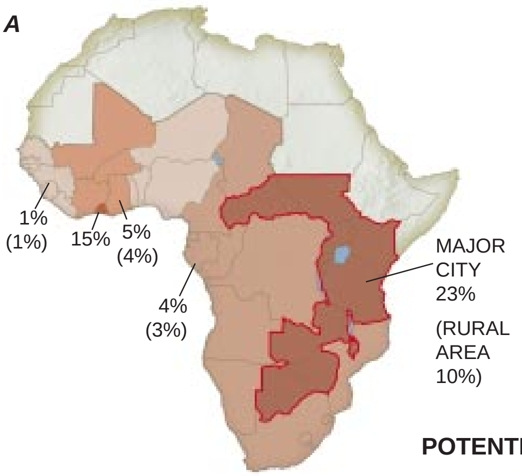
AIDS in Sub-Saharan Africa

The impact of AIDS on the continent of Africa has been anything but uniform. In the stretch of countries known as the AIDS belt (*large map on this page*), nearly 25 percent of the urban population is infected with HIV, the AIDS-causing virus. Yet in other regions, the rate of infection is as low as 1 percent (*map A*), comparable to numbers seen in some Western countries. To account for the extraordinary epidemic in the AIDS belt, we examined a variety of factors to determine what customs were both common throughout this region and unique to it and thus might explain the severity of the disease there.

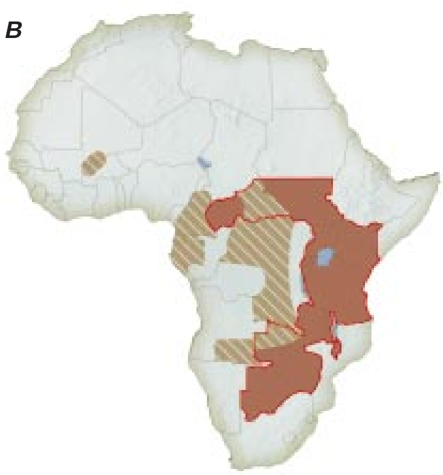
For example, high levels of sexually

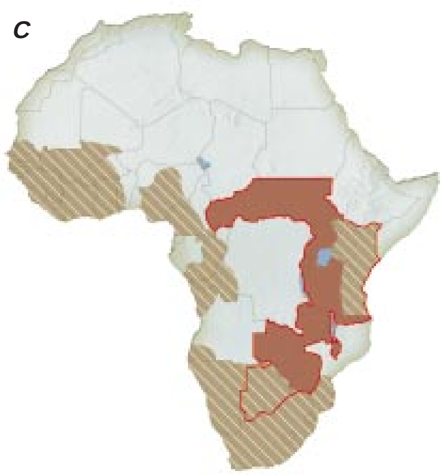
THE AIDS BELT

PERCENTAGE OF POPULATION WITH HIV

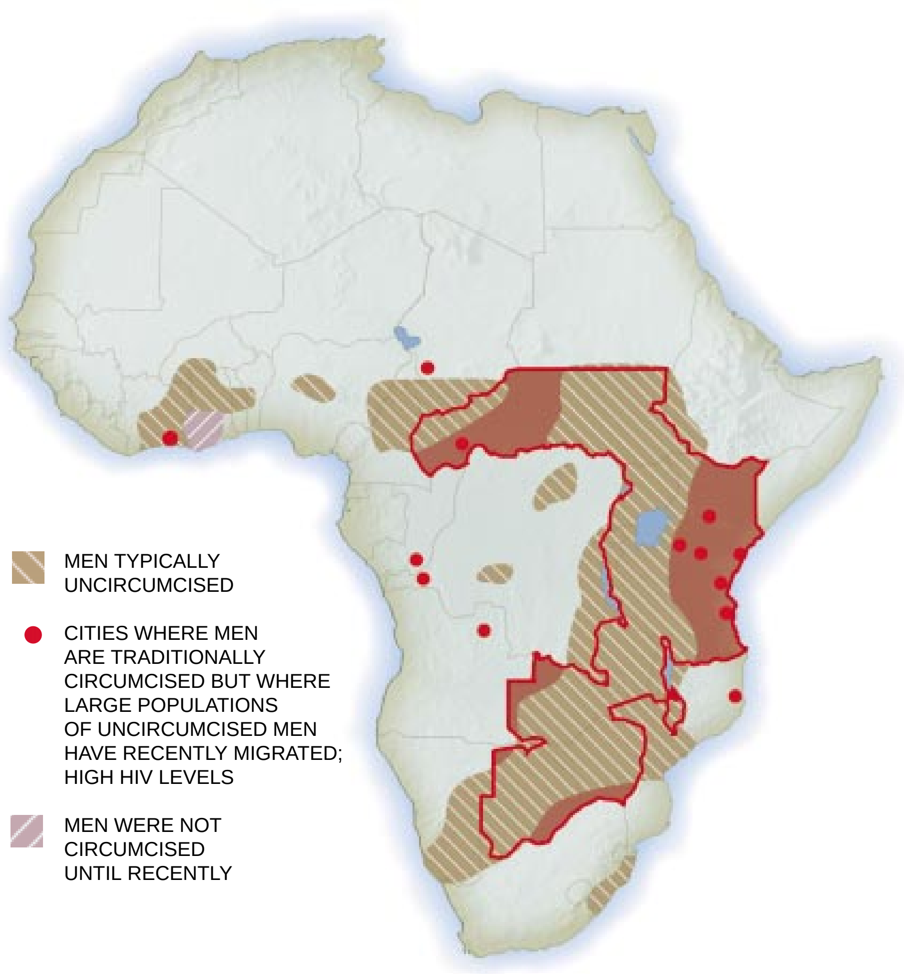


POTENTIAL EXPLANATIONS FOR AIDS EPIDEMIC



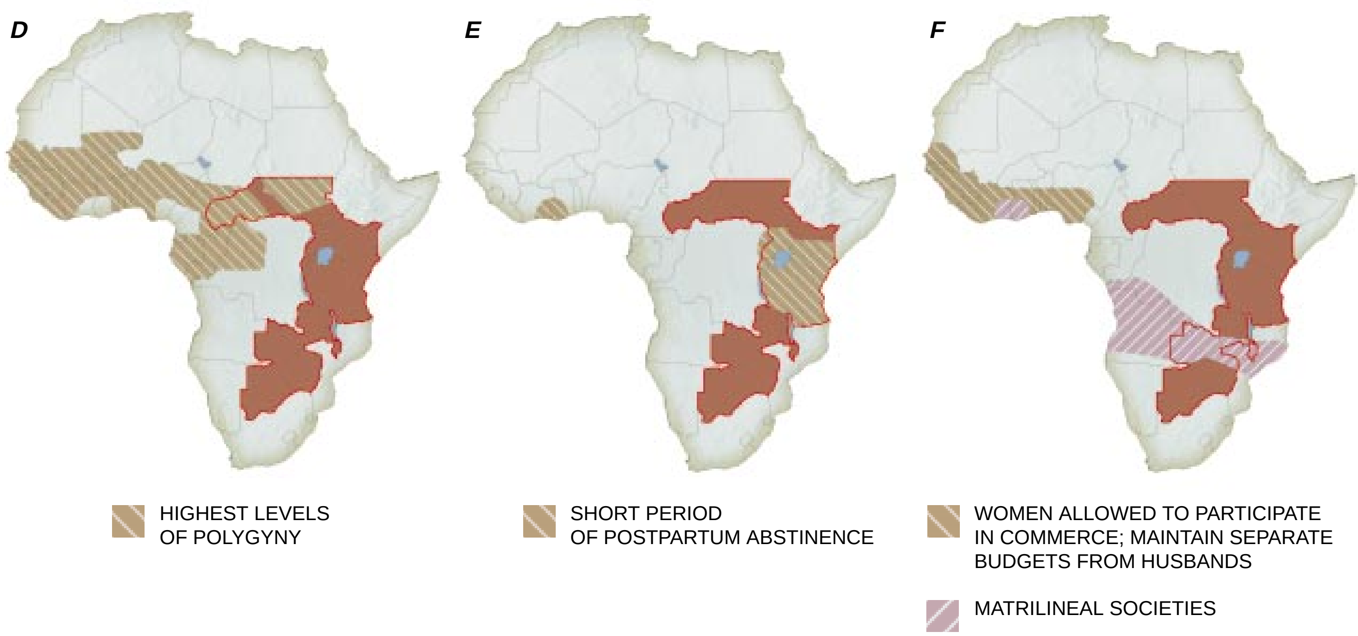


REGIONS WHERE MOST MEN ARE UNCIRCUMCISED



transmitted diseases (*map B*) often signify regions where sexual behavior is particularly risky (people have multiple partners or engage in sex with prostitutes). Also, premarital and extramarital sex—generally more conducive to spreading such diseases than sex within marriage—tend to be more common where people marry late (*map C*) or where men marry more than one wife (*map D*), thus forcing some men to stay single for many years. Sex outside marriage can be less common where women's sexual behavior is not restricted after giving birth (*map E*). Matrilineal societies or communities that allow women to participate freely in commerce foster independence in women, which may in turn promote premarital and extramarital sex (*map F*).

Most of the ideas we investigated failed to explain the extraordinarily high rate of infection in the AIDS belt. One factor did stand out, however: lack of male circumcision (*large map on this page*). In the area where men are typically uncircumcised, HIV rates are among the highest in the AIDS belt. Furthermore, the migration of uncircumcised men from this area to cities outside the region of rampant infection has led to exceedingly high numbers of AIDS cases in many urban locales as well. —*J.C.C. and P.C.*



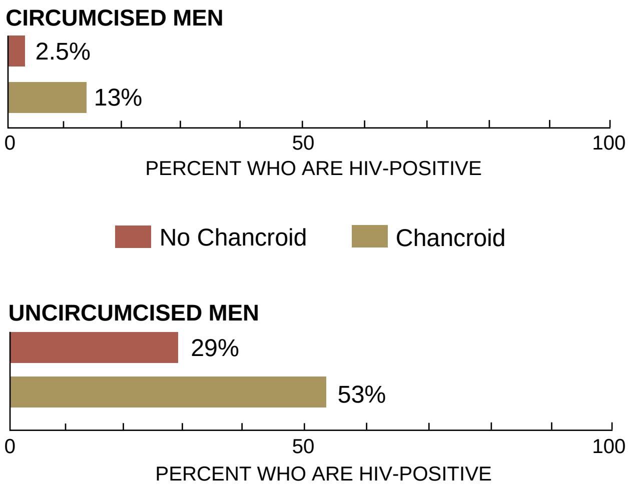
LAURIE GRACE AND PETE SAMEK

Circumcision, Chancroid and AIDS

So far no one has been able to determine precisely how lack of circumcision may make some men more susceptible to HIV infection. But current thinking emphasizes both the nature of the foreskin and the likely importance of penile hygiene. Certain sexually transmitted diseases, particularly chancroid, which causes large, soft sores on the genitals, tend to occur more frequently among uncircumcised men in poor areas where maintaining personal cleanliness is difficult. Chancroid disappeared in the West around the beginning of this century, apparently as the society became more affluent, making hygiene easier to maintain.

Studies of American and Australian soldiers during the Vietnam and Korean wars support the connection between lack of circumcision, poor hygiene and high levels of chancroid. In a 1952 study of U.S. soldiers in Korea, for instance, Joseph Asin of the U.S. Army wrote that the men were more susceptible to chancroid infection as a result of inadequate cleaning, "especially washing the genitals with soap and water following exposure." For uncircumcised men, thorough cleaning of the genitals can be particularly challenging. Only about 33 percent of U.S. soldiers in Korea were uncircumcised, but 96 percent of those infected with chancroid were uncircumcised.

Recent research in Kenya has demonstrated that uncircumcised men with chancroid are at greater risk for HIV [*see graphs at left*]. Still, these surveys cannot determine whether lack of circumcision itself makes men more susceptible to AIDS or whether lack of circumcision promotes chancroid infection, which then catalyzes transmission of HIV. We suspect that both scenarios account for high levels of HIV infection in the AIDS belt. Once men are infected with chancroid, they are at greater risk of contracting HIV because the presence of genital sores makes transmission during intercourse more likely. In contrast to healthy foreskin, infected foreskin can be penetrated more easily by HIV. *—J.C.C. and P.C.*



SOURCE: F. A. Plummer et al. in AIDS and Women's Reproductive Health, *edited by L. C. Chen et al. Plenum Press, 1991.*

JESSIE NATHANS

Continued from page 63

American team led by John Bongaarts of the Population Council published a paper showing that the regions across sub-Saharan Africa with high levels of HIV infection among local peoples corresponded remarkably well with the areas where men were typically uncircumcised. The next year Stephen Moses and his colleagues from the Canadian-Kenyan group completed their own analysis of the AIDS epidemic and arrived at the same conclusion.

Surprisingly, the publication of these two papers did not excite much interest. The World Health Organization's Global Program on AIDS emphasized that Moses and his team could not point to any physiological mechanism to explain how lack of circumcision was implicated in an increased risk of HIV infection; the research simply demonstrated a statistical correlation between circumcision, chancroid or syphilis, and HIV infection. Consequently, other researchers gave the finding little weight.

Instead many AIDS experts seemed to assume that African conditions and behavior were so "bad," in some unspecified way, that the disease would spread inexorably throughout the continent. But we felt that this explanation was wrong. We believed that we could establish what made HIV transmission so common in Africa, particularly in the AIDS belt. Was it lack of circumcision, or something else? Our team redoubled its efforts to examine the culture of sub-Saharan Africa, looking for any traditions that might account for the rampant spread of heterosexually transmitted AIDS cases.

Background Research in Nigeria

We decided to look first at the country of Nigeria, outside the AIDS belt, where we have been working over a 30-year period, studying the very traditions and sexual behavior that might affect AIDS transmission. According to established thinking about sexually transmitted diseases, much of the sexual behavior in Nigeria—where most men and many women have multiple sex partners, for example—seemed to make the country a prime candidate for a heterosexual epidemic of AIDS. Yet the percentage of HIV-positive people (0.5 percent of the population) is the same as in the U.S. and only slightly above that in other Western countries. Perhaps our work in Nigeria could help us isolate unique conditions found in the AIDS belt that contribute to the extraordinary epidemic there.

For reasons embodied in traditional culture and religion, and probably associated with certain types of farming, land tenure and inheritance, sub-Saharan African culture has always placed greater emphasis on fertility than on the repression of women's premarital and extramarital sexuality [see "High Fertility in Sub-Saharan Africa," by John C. Caldwell and Pat Caldwell; SCIENTIFIC AMERICAN, May 1990]. Infidelity might occasionally spark fights, punishment and, more rarely, marital dissolution, but it was never equated with sin and excoriated in the way that it was in traditional Western and Asian societies. Much good flowed from this permissive attitude: women were not suppressed and hidden, and girls had survival chances as great as their brothers'. But eventually these cultural traditions and attitudes did make the societies susceptible to attack by sexually related diseases.

Our work in Nigeria showed that in the southern part of the country, most married men with one wife and more than 25 percent of married men with multiple wives had experienced extramarital sexual relations within the previous year; 75 percent of single men had had relations with at least two women in that time. In addition, around 30 percent of married women had had intercourse with at least two men in the previous year, as had 50 percent of all single women. Translated into terms of the number of lifetime partners, these findings are similar to the results of recent social surveys of the U.S., Britain and France, although in these countries, a greater proportion of partnerships occurred before marriage. Apparently, even though sexuality is becoming a bit more liberated in Western societies, the culture is still more successful in discouraging extramarital sex. In Western countries, heterosexual HIV transmission remains low, however, because of better prevention measures.

Sexually transmitted diseases and AIDS spread rapidly among people with many sexual partners, poor facilities to combat infection and lack of access to preventive means, such as condoms. Not surprisingly, we found such illnesses as gonorrhea and, to some extent, syphilis to be common in Nigeria; chancroid, however, is almost unknown there. Somehow the country was not suffering the kind of AIDS epidemic raging elsewhere. Throughout the parts of sub-Saharan Africa that lay outside the AIDS belt, men and women had multiple sex partners, untreated sexually transmitted diseases were widespread, and a significant proportion of males visited prostitutes, among whom levels of HIV infection were sufficient to start an epidemic. But the disease had reached a crisis level only in the AIDS belt. What was the additional factor that characterized this area?

Rejecting Other Options

Along with other groups, we explored a variety of hypotheses. Perhaps AIDS might spread along routes of unusually high levels of sexually transmitted diseases, which tend to be most pervasive where risky sex practices, such as sex with prostitutes or multiple partners, are rampant. A test was at hand. Immediately to the west of Uganda and Rwanda, but outside the AIDS belt, lies an area known as the world's major infertility belt. In this region, an epidemic of sexually transmitted diseases, persisting for more than a century, has devastated the population, leaving many people sterile. Significantly, though, despite the extreme levels of certain sexually transmitted diseases, few people in this region suffer from chancroid, and the prevalence of AIDS here is, at most, moderate.

We looked next at the custom of scarification. In this practice, found widely throughout Africa, incisions made on the limbs or face are believed to prevent or cure illness. In some areas, all members of a household where sickness has occurred are scarified one after another with the same instruments, so the possibility of contaminating one person with another's blood is high. But this practice is not popular throughout the AIDS belt, nor is it unique to it. Thus, scarification could not explain the epidemic.

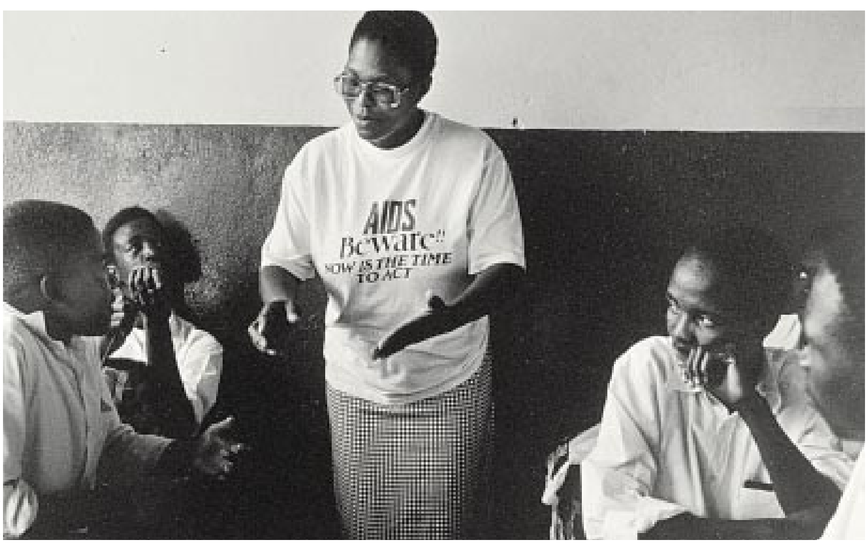
We then tried to determine what might prompt men to resort to sex—and particularly commercial sex—outside marriage. For example, in societies where men marry more than one wife, there are typically large numbers of single men seeking sex; for many years, these men are simply unable to find any available women to wed. The practice of polygyny, however, is more common outside the AIDS belt. We also looked at average male age at the time of marriage—theorizing that the longer a man remained unmarried, the longer, and the more frequently, he might resort to visiting prostitutes—but this idea, too, failed to provide an explanation.

We examined the duration of abstinence among women after they give birth: in some regions, men have no sexual access to their wives for up to 60 percent of their married life because of traditions regarding childbirth and sexual relations. But in certain parts of the AIDS belt, postpartum abstinence is rather brief compared with the practice elsewhere in Africa, reducing the time when husbands can turn only to extramarital—and hence risky—sex.

We followed up a host of other ideas: The tradition of young men participating in warrior societies, which tend to be critical of premarital pregnancies, might encourage men to visit prostitutes, among whom accidental pregnancies would have little impact on the men involved. The custom of demanding large bride-prices makes parents more worried about a marriage arrangement falling apart if the bride becomes pregnant before marriage; such concern might also prompt men to visit prostitutes before marriage. Female autonomy, found in matrilineal societies or reflected by women's participation in commerce and trade, might foster women's sexual independence from their husbands, possibly resulting in risky extramarital sex on the part of both husbands and wives. Populations with more men than women (most frequently found in urban areas as a result of immigration) might also promote prostitution. But none of these situations were both common throughout the AIDS belt and unique to it—we still had not isolated the factor that was promoting the rapid heterosexual transmission of HIV.

Back to Circumcision

In frustration, we turned once again to the earlier research that suggested a connection between lack of male circumcision and the spread of AIDS. We updated our maps with more recent data on levels of HIV infection and noted that the areas of Africa with large numbers of uncircumcised men were almost exactly the same as the regions suffering from the severe AIDS epidemic [*see box on pages 64 and 65*]. But other researchers had discounted these reports, claiming that the original source of the circumcision data was out-of-date (the information on men's circumcision status was first presented in *Eth-*



GIDEON MENDEL *Network Matrix*

EDUCATION PROGRAMS throughout Africa and the world can help slow the deadly spread of AIDS. In Hlabisa, South Africa, a volunteer informs high school students about the hazards of engaging in unprotected sex.

nographic Atlas, by George P. Murdock of the University of Pittsburgh, published in 1967) or that the analysis of the proposed relationship was faulty. And some critics felt that the findings were merely coincidental. Further complicating the debate was the controversy in the Western medical profession about whether circumcision was meaningless mutilation; many did not wish to revive that discussion. In addition, some health administrators feared that linking circumcision status, susceptibility to AIDS and ethnicity might lead to greater ethnic hostility rather than improved AIDS-prevention strategies.

Over the past three years, however, we have reexamined the methodology of the papers as well as the anthropological sources and determined that the findings are sound. Also, in continuing investigations we have found very little support for the charge that the circumcision data are no longer relevant. The link between lack of circumcision and elevated levels of HIV infection appears robust. In some parts of the AIDS belt, nearly all men are uncircumcised—a situation unlike almost anywhere else in Africa. Outside the AIDS belt, in the city of Abidjan, the capital of Ivory Coast, levels of HIV infection are as high as they are in some cities of the AIDS belt; we believe the epidemic in Abidjan is very likely sustained by immigrants who come from a surrounding area where a majority of men are uncircumcised.

Thus, we concluded that in the AIDS belt, lack of male circumcision in combination with risky sexual behavior, such as having multiple sex partners, engaging in sex with prostitutes and leaving chancroid untreated, has led to rampant HIV transmission. Unsafe sexual practices have certainly contributed to the spread of AIDS across Africa and indeed around the world. But risky behavior alone cannot sustain the epidemic seen in the AIDS belt.

Caution is warranted regarding the connection between male circumcision and AIDS. Lack of circumcision seems primarily to amplify the dangers of unsafe sexual practices: men in Asia, for example, are frequently uncircumcised, but the region has not seen the number of AIDS cases that Africa has, probably because of different sexual customs. And circumcision is not an absolute protection against AIDS; circumcised men are certainly still at risk. Everywhere in sub-Saharan Africa, even where all men are circumcised, levels of HIV infection are high among prostitutes and their customers. The epidemic in this group could potentially serve as a starting point for the virus to move into other communities.

Combating AIDS Globally

Because the virus progresses through different segments of the population at markedly distinct rates, the best chances for combating AIDS everywhere lie in targeting education and prevention programs at high-risk groups—homosexuals, prostitutes and their clients, intravenous drug users, and promiscuous men and women. In sub-Saharan Africa, circumcision could be offered as a reinforcement of other protective measures.

Universal male circumcision, which is likely to prove unacceptable, would nevertheless probably reduce the level of infection in the AIDS belt to the much lower numbers seen in the rest of sub-Saharan Africa. But eradication of the disease will require a multipronged assault: people must be encouraged to have fewer partners and practice safer sex. Improved treatment for sexually transmitted diseases, particularly chancroid, is also crucial.

There are signs of change in East Africa. Average numbers of sexual partners appear to be falling. Women in Uganda, for example, are offering alternatives to vaginal sex or refusing to have unsafe sex. And condom use has become much more accepted. No doubt as a result, the disease appears to be slowing in Uganda.

In rural southwest Tanzania, and surely elsewhere in the AIDS belt, uncircumcised men have not waited for agreement among researchers about the connection between circumcision and AIDS. Based on observations of their community and neighboring ones, they have concluded that they are at greater risk for AIDS than circumcised men. These men are appearing at hospitals in sharply increasing numbers, requesting circumcision for themselves and often for their sons. Clinics that offer adult male circumcision as a protection against AIDS now advertise in Tanzanian newspapers.

Thankfully, ethnic rivalries do not appear to have worsened as a result of this issue. If health authorities state that lack of circumcision appears to make men more susceptible to HIV infection, it now seems probable that some men will seek circumcision. At the same time, however, medical experts will have to stress to everyone that circumcision is not in itself protection against AIDS. Although the epidemic in sub-Saharan Africa may subside somewhat, because of greater use of condoms and probably increased incidence of circumcision, Africans in the AIDS belt remain at extremely high risk of HIV infection.

For AIDS programs worldwide, the sub-Saharan epidemic demonstrates an important lesson. Because a number of unusual conditions must combine to sustain a society-wide heterosexual epidemic, it seems unlikely that in the West, AIDS will spread significantly beyond high-risk groups such as homosexual men and intravenous drug users. Consequently, prevention programs—particularly those in Asia, where the epidemic may worsen—should be addressed specifically to people who are most likely to contract HIV. Such efforts should slow the ongoing epidemic in Africa and prevent similar conditions from developing elsewhere.

The Authors

JOHN C. CALDWELL and PAT CALDWELL work at the Health Transition Center of the National Center for Epidemiology and Population Health at the Australian National University in Canberra. Both John Caldwell, a demographer, and Pat Caldwell, an anthropologist, are principal investigators in the Swedish Agency for Research Cooperation project on sexual networking, sexually transmitted disease and HIV in sub-Saharan Africa. Between 1970 and 1988 John Caldwell served as head of the Australian National University's department of demography. He is also adjunct professor at Harvard University's department of population sciences and international health.

Further Reading

THE RELATIONSHIP BETWEEN MALE CIRCUMCISION AND HIV INFECTION IN AFRICAN POPULATIONS. John Bongaarts, Priscilla Reining, Peter Way and Francis Conant in *AIDS,* Vol. 3, No. 6, pages 373–377; 1989.

GEOGRAPHICAL PATTERNS OF MALE CIRCUMCISION PRACTICES IN AFRICA: ASSOCIATION WITH HIV SEROPREVALENCE. Stephen Moses, Janet E. Bradley, Nico J. D. Nagelkerke, Allan R. Ronald, J. O. Ndinya-Achola and Francis A. Plummer in *International Journal of Epidemiology,* Vol. 19, No. 3, pages 693–697; September 1990.

SEXUAL NETWORKING AND AIDS IN SUB-SAHARAN AFRICA: BEHAVIOURAL RESEARCH AND THE SOCIAL CONTEXT. Edited by I. O. Orubuloye, John C. Caldwell, Pat Caldwell and Gigi Santow. Australian National University, 1994.

FORUM: THE EAST AFRICAN AIDS EPIDEMIC AND THE ABSENCE OF MALE CIRCUMCISION: WHAT IS THE LINK? In *Health Transition Review,* Vol. 5, No. 1, pages 97–117; April 1995.

Budding Vesicles in Living Cells

A transatlantic collaboration has uncovered the machinery responsible for forming the tiny but essential containers, or vesicles, that store proteins and shuttle them to and fro in cells

by James E. Rothman and Lelio Orci

All nucleated cells—whether in colonies of yeast or in plants and people—have a complex internal organization resembling that of a well-run city. Perhaps most notably, cell and city both rely on the coordinated activities of specialized departments. In cells these departments are walled off by membranes and called organelles.

A brief tour of some of the more important departments in a cellular "city" might start at the outer membrane, itself an organelle. This structure is akin to the gated walls that once enclosed ancient cities, in that it controls the entry of food and other materials and the export of products manufactured within. Another critical department lies deep inside cells and serves as a manufacturing center. It is here, in the endoplasmic reticulum, that many proteins—the main working parts of cells—are produced. These newly made proteins are then transported to yet another department, the Golgi apparatus, where they are modified (often by the addition of sugars) and ultimately shipped to other destinations within or outside the cell. The Golgi, then, is a major distribution hub for our microscopic cities.

Cells even include their own recycling centers, known as lysosomes. These centers break down old and expired proteins and certain other molecules so that their constituents can be refashioned and reused. Lysosomes also accept material that has been hauled into the cell from the outside.

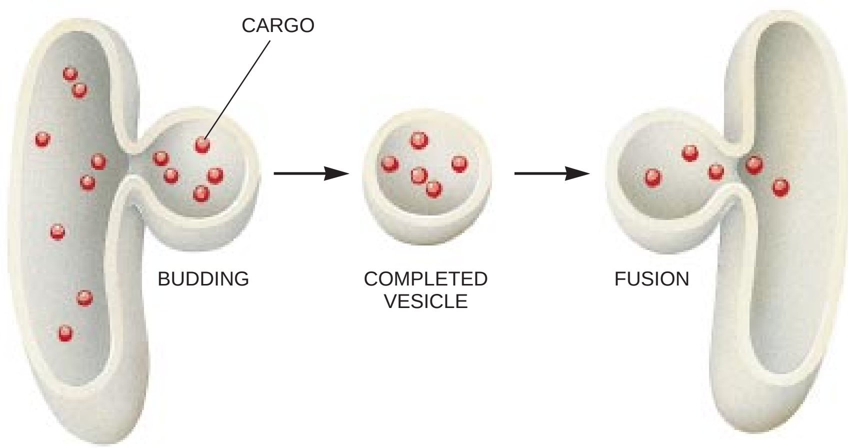
Not surprisingly, cells have evolved a complex transportation system for conveying proteins between the various organelles. It is organized like the container-based shipping systems at airports. All goods having a shared destination are packed into a common cargo container, called a transport vesicle, which is opened only when it arrives at its intended destination. Every cell produces many kinds of transport vesicles, and each type is specialized for a distinct intracellular shipping route or kind of cargo. Cells also produce transport vesicles that store substances critical to communication between cells (such as neurotransmitters and insulin); these carriers release their cargoes from cells in response to precise signals.

Clearly, transport vesicles are vital for many aspects of a cell's, and a body's, functioning. Yet for years no one knew how they formed. Recently, however, the two of us and our colleagues have sketched out many of the molecular details. Aside from having intrinsic interest, the results could eventually be of value to medicine. For instance, tumor cells can multiply only if their transport vesicles function properly. Hence, substances that impeded vesicle formation could conceivably augment anticancer drugs now on the market.

Our story is in many ways characteristic of much modern cell biology in that it has its origins in microscopy (which first revealed the existence of transport vesicles) and its resolution in biochemistry (which has detailed many molecular reactions involved in vesicle formation). This tale also illustrates the true nature of scientific research. Nonscientists often think of scientific discovery as an impersonal process in which pure intellect leads by way of inexorable logic to a clear solution to a problem. This view underestimates the role played by false steps, occasional good luck and dogged persistence. And textbook descriptions never capture the fun of the chase! For us, much of the fun has been working closely with each other despite our separation by the Atlantic Ocean, which has allowed only annual visits.

"Who *Are* You?"

Before we began collaborating, in the mid-1980s, microscopists who examined slices of whole cells had already outlined certain basic steps of protein transport by vesicles. Largely on the basis of Nobel Prize–winning research conducted in the 1960s by George E. Palade, then at the Rockefeller University, investigators knew that vesicles probably arose by budding from the membrane of an organelle, thereby becoming tiny, membrane-bound sacs.



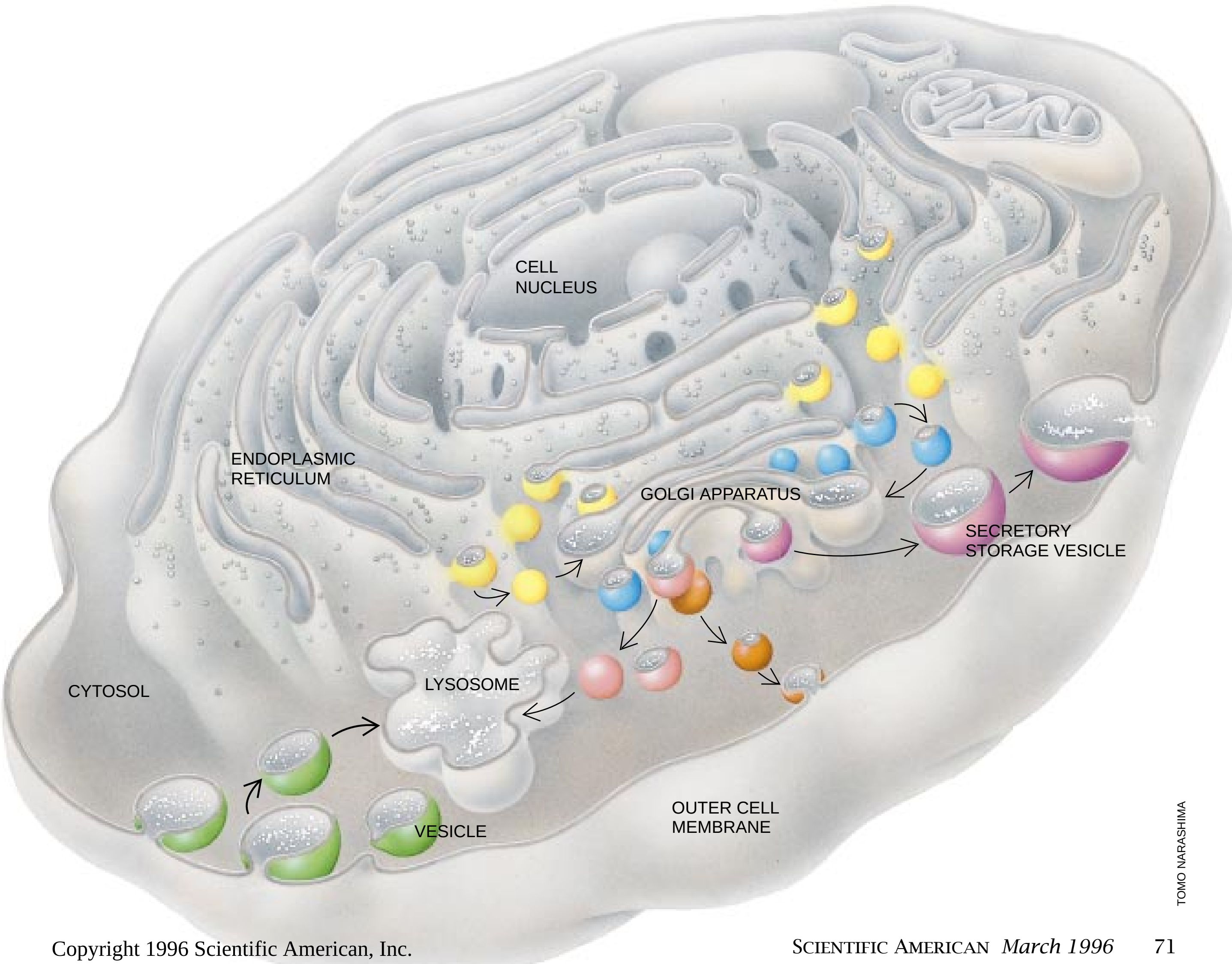
Each sac subsequently migrated through the cytosol (the protein-rich sap in which organelles float) and stuck to the outer membrane of a specific target organelle. Then the vesicle's membrane merged, or fused, with that of the target organelle, in the process pouring the stowed proteins into the target's interior. In other words, the "cargo container" was opened at its destination and its cargo released.

To clarify how budding occurred, it would be necessary to study the process in the test tube—that is, to re-create vesicle formation outside the cell, eliminating extraneous elements. With such a system one could manipulate or otherwise analyze the components of the vesicle-producing machinery to determine which molecules acted first, and next, and so on. But no one had managed to devise such a cell-free system until 1980, when one of us (Rothman) did so at Stanford University.

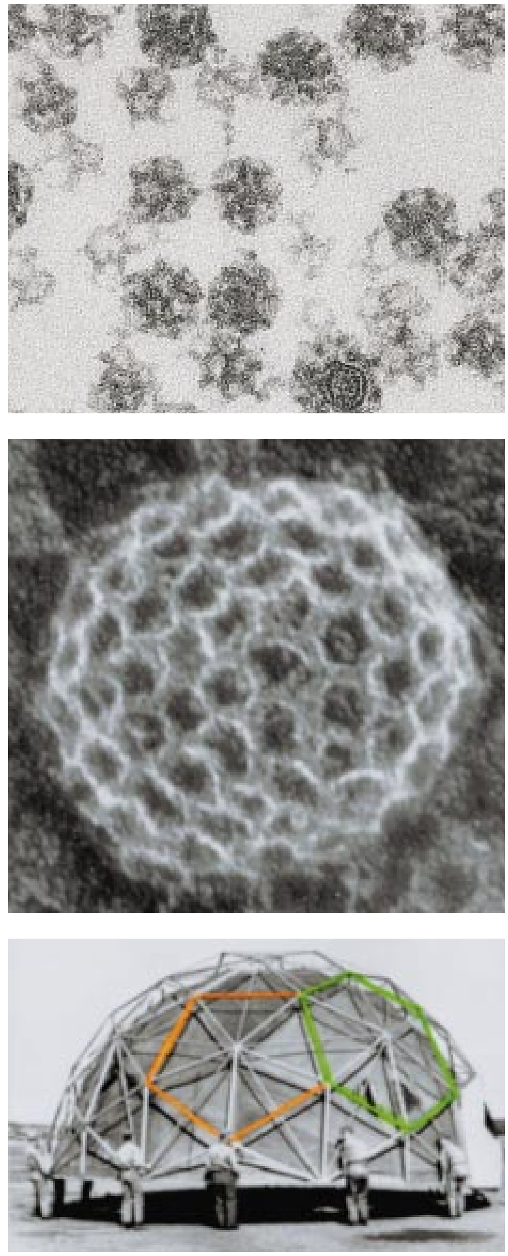
In 1984 Rothman and his colleagues published a report announcing that their cell-free system could re-create the vesicle formation known to occur in whole cells. It had enabled transport vesicles to bud from membranes of the Golgi apparatus, which has several attached compartments. Most vesicles derived from Golgi membranes convey cargo from one compartment of the Golgi to another; these were the type observed. Some Golgi vesicles, however, dispatch proteins to more distant sites [see "The Compartmental Organization of the Golgi Apparatus," by James E. Rothman; SCIENTIFIC AMERICAN, September 1985].

It was at this stage that our collaboration began, in memorable fashion. Late one afternoon at Rothman's office, less than a week after the paper's December publication, the telephone rang. Interrupting his paperwork, Rothman picked up the receiver. Before he could utter "hello," a deep voice began an intense, rapid and totally incomprehensible monologue in Italian-accented English. Rothman, not unknown himself for intense verbalization, was stunned into silence by this display. After several minutes, at a complete loss, he returned to his paperwork, allowing the monologue to continue. After another few minutes, in which the only recognizable words were "collaborate" and "we'll be brothers," Rothman finally

TRANSPORT VESICLES (*colored spheres*) abound in cells. Some (*yellow*) convey proteins made in the endoplasmic reticulum to the Golgi apparatus, where the proteins are modified. Others (*blue*) shuttle proteins from one compartment of the Golgi to another. Three kinds emerge from the Golgi. One (*orange*) exports proteins from the cell immediately. Another—the secretory storage vesicle (*dark pink*)—releases its contents when signaled to do so. A third (*light pink*) brings digestive enzymes to chambers called lysosomes, which break down various molecules, including those brought into the cell by still other vesicles (*green*). A vesicle forms (*opposite page*) by budding from an organelle; it delivers cargo (*red*) by fusing its membrane with that of the target organelle.



CLATHRIN COAT was the first type of coating identified on transport vesicles; it is found mainly on vesicles that carry proteins from outside the cell to lysosomes. Electron microscopy showing sectional (*top*) and surface (*middle*) views of such vesicles reveals the cagelike structure of the coat. Its hexagonal-pentagonal organization resembles that of a geodesic dome (*bottom*).



BLACK-AND-WHITE IMAGE COURTESY OF THE BUCKMINSTER FULLER INSTITUTE, LOS ANGELES; MICROGRAPHS BY LELIO ORCI

cried out in frustration, "Who *are* you?"

The voice paused and then very quietly and distinctly, even modestly, declared, "I am Lelio Orci. I am a professor at the University of Geneva, and I work on—"

Of course, Rothman immediately recognized the name. Orci was a well-known electron microscopist. Rothman instantly apologized, Orci spoke more slowly, and a long and productive interchange took place, even though it was already well past midnight in Geneva.

Wrong Idea, Useful Result

Within a few weeks Rothman was in Geneva, and the two of us excitedly worked out a plan by which Orci's techniques could determine whether the vesicles budding from Golgi membranes in cell-free extracts had a particular property. Specifically, we wanted to know if the vesicles were encased in a so-called clathrin coat: a thick, cagelike layer composed primarily of the fibrous protein clathrin. Structurally, the clathrin coat looks something like a geodesic dome.

We asked this question because certain transport vesicles were known to possess such coats, the only type recognized at the time. (Clathrin-coated vesicles transport proteins from outside the cell to the lysosomal system for breakdown; they bud into the cytosol from the cell's outer membrane.) But the mechanism of budding could not be deciphered for lack of a cell-free system. If Orci could confirm that the Golgi-derived vesicles were garbed in a clathrin coat, Rothman could immediately begin biochemical studies in his cell-free system to evaluate an appealing model of how clathrin-coated vesicles form.

That model emerged in the 1960s from research initiated by Toku Kanaseki and Ken Kadota of Osaka University. In 1969 Kanaseki and Kadota purified coated vesicles and discovered that the coats had a surprisingly regular structure. The reason for this regularity became apparent six years later, when Barbara M. F. Pearse of the Medical Research Council Laboratory in Cambridge, England, discovered the coat primarily consisted of many copies of one protein, which she named clathrin. Kanaseki and Kadota envisioned that the coat itself was responsible for budding. As the clathrin cage developed, presumably from membrane components, it deformed the underlying membrane (which is a highly plastic structure) into a domelike shape, literally sucking a spherical vesicle and any attached proteins out of the larger membrane.

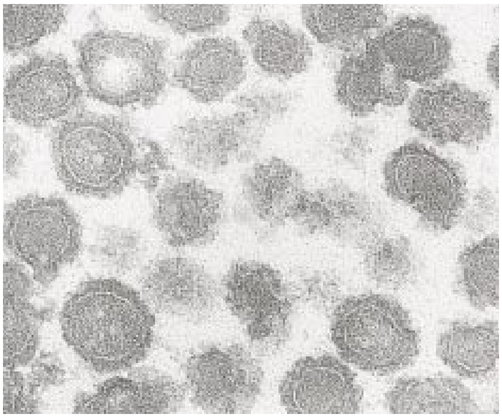
By 1984 it was clear that this proposal was a bit too simplistic. When Rothman demonstrated budding in his test-tube system, he found that the budding process could not occur from isolated membranes alone. Vesicle formation required the cytosol as well as an energy source. These results suggested that the coat must somehow be derived from substances in the cytosol. Nevertheless, Kanaseki and Kadota's basic model remained attractive and merited testing.

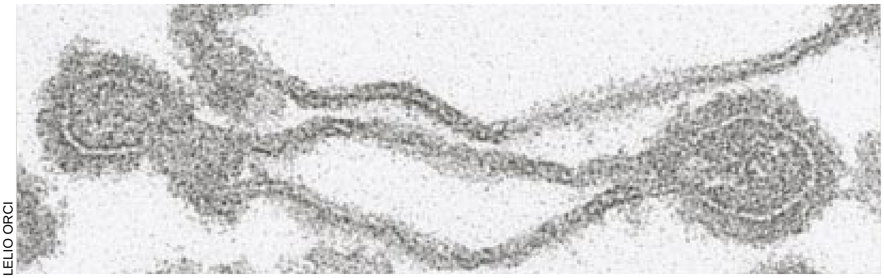
Soon after Rothman returned to Stanford from Geneva, one of his graduate students began to implement our plan. Benjamin S. Glick generated transport vesicles in the test tube by incubating the Golgi apparatus with cytosol and an energy source. He then preserved the samples and sent them to Geneva for microscopic analysis. Within days Orci had established that the Golgi-derived vesicles were indeed wrapped in a coat, but not of the expected type. Its fine structure differed from that of clathrin coats, and the material was not recognized by antibodies that bind specifically to clathrin. That surprising result meant that cells manufacture at least two kinds of transport vesicles sporting different coats. It now appears that many kinds of transport vesicles exist, each with a distinct coat.

The outcome of that study is a wonderful example of the unanticipated value of the false step. Our incorrect hypothesis—that clathrin coats mediate formation of transport vesicles in the Golgi apparatus—led us to conduct an experiment that shifted our attention away from clathrin-coated vesicles to the ones we had just identified. Our mistake thus propelled us down a fruitful new path of research that revealed much about how transport vesicles arise. (In contrast, understanding of how clathrin-coated vesicles develop has progressed more slowly. Until recently, no one had devised a cell-free system for investigating them.)

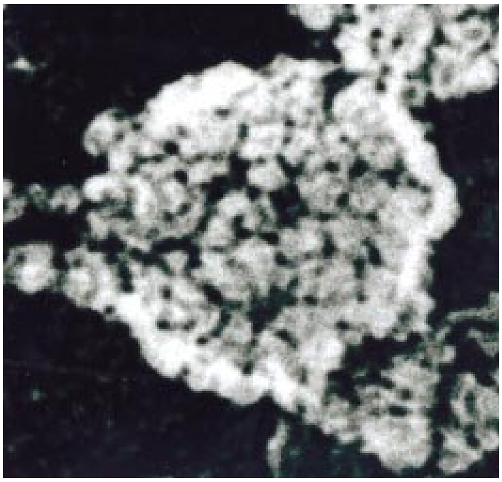
As the first step to uncovering how the Golgi-derived vesicles formed, we needed to learn the composition of their coat. This challenge was to consume several years. To begin, we had to obtain a pure supply of the tiny spheres, a daunting task at the time. After intense struggle, Vivek Malhotra and Tito A. Serafini, young investigators in Rothman's laboratory, isolated a minuscule amount in 1989. But it was enough to analyze and enabled them to demonstrate that the coat contained eight proteins. We named them COP proteins (for *co*at *p*rotein) and soon took to calling the Golgi-derived vesicles COP-coated vesicles.

To clarify the budding mechanism, we had to find out more about the COP proteins and their interactions with one another. Our earliest major insight into these interactions came in 1990, when we discovered that seven of the coat proteins assembled into a large complex before attaching as a single unit to the Golgi membrane; we called the unit a


LELIO ORCI


LELIO ORCI

COP COAT, a type identified by the authors, envelops transport vesicles that arise from the Golgi apparatus. Its difference from the clathrin coat is apparent in the electron micrographs at the left. The micrograph above captures two COP-coated vesicles in the process of budding from a Golgi membrane.


LELIO ORCI

coatomer. Only one protein bound separately. Essentially, then, the coat consisted of just two main elements: the coatomer and the eighth protein. This profoundly simplifying discovery meant we could glean the principal steps of budding without having to study each coatomer protein separately.

Here was where luck and the prepared mind came into play, for the discovery was made accidentally. Late one evening in 1990, M. Gerard Waters, a young researcher in Rothman's laboratory (by then at Princeton University), was examining data from his attempts to purify a substance needed for transport vesicles to fuse with membranes of target organelles. By coincidence, Serafini happened by and, looking over Waters's shoulder, was astonished by what he saw.

In purifying a protein, one divides the components of a cell extract into different groupings according to their physical and chemical properties. Then the groupings, or fractions, are tested for evidence of the activity of interest (such as triggering fusion of a vesicle and a membrane). Samples that show activity are further subdivided into a new set of more refined fractions and retested. Meanwhile fractions showing no activity are usually discarded. Ultimately, one obtains a sample containing only the protein responsible for the biological activity of interest—a pure protein.

To be thorough, Waters had analyzed the protein content in a fraction that had shown no evidence of triggering fusion. For months, he had, quite naturally, been discarding this fraction as useless. Serafini was struck by the similarity of proteins in Waters's throwaway fraction to those he and Malhotra had identified earlier as the COP proteins.

Quick work showed that the discarded material consisted almost exclusively of seven of the COP proteins, which were bound tightly to one another. We knew they were associated because the fraction steadfastly resisted being subdivided further. Ironically, then, while one part of Rothman's laboratory had been engaged in a soul-destroying effort to isolate coated vesicles (so as to attain trace amounts of the coat proteins), another part had been throwing away the same proteins in huge quantities. Thanks to Serafini's and Waters's alertness, the waste was halted, the existence of coatomer was recognized, and years of effort were saved. We now had at our disposal an unlimited source of the coatomer proteins needed for various experiments.

We soon turned our attention to the identity of the eighth protein. For many reasons, we suspected it was a molecule called ARF (for *A*DP *r*ibosylation *f*actor). This molecule was first described by Richard A. Kahn and Alfred G. Gilman of the University of Texas Southwestern Medical School and was known to enable the cholera toxin to cause disease. But its normal role in the body was a mystery. Fortunately, confirming our suspicion was easy. From Kahn we obtained antibodies that recognize ARF and nothing else. Then we mixed the antibodies with our protein. The antibodies bound readily, indicating that the protein was, indeed, ARF.

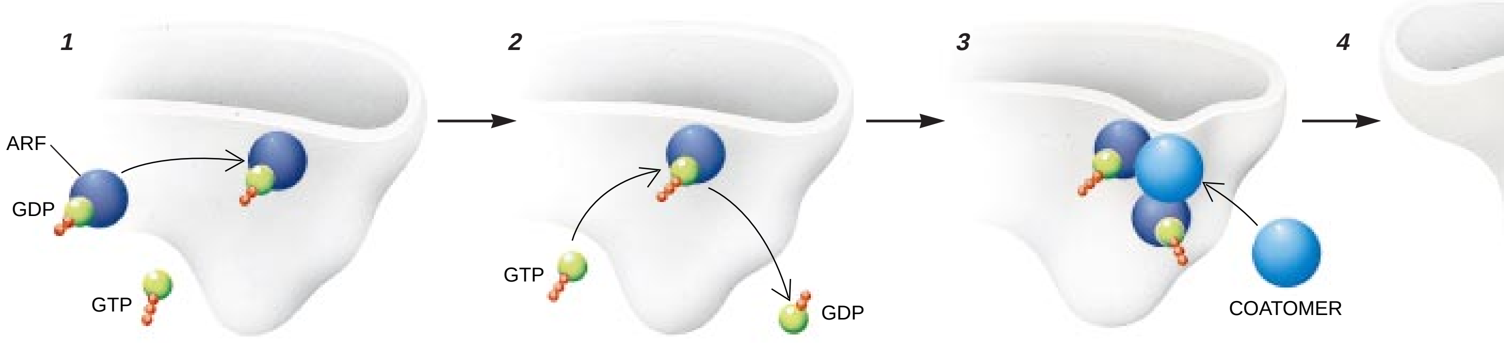
After six years of arduous labor, we had all the coat proteins in pure form and were finally ready to decipher the mechanism of budding. This task was more straightforward. We began the project in 1991, after Rothman had moved his laboratory to the Memorial Sloan-Kettering Cancer Center, and we learned a great deal in just two years.

Budding Unveiled

We first asked whether coatomer and ARF were the only proteins needed in order for COP-coated vesicles to bud from the Golgi apparatus. Did they perhaps require assistance from other proteins? The two elements proved sufficient in themselves! Although we had started with a crude cytosolic extract, containing many thousands of different proteins, the power of biochemistry had allowed us to cull the two that really mattered—coatomer and ARF—and to prove they alone do the job.

The next step was to add the proteins one at a time instead of together, as a way of deciphering the roles played by the two components. We saw that ARF bound to the Golgi membrane in the absence of coatomer and then recruited coatomer to the membrane. Coatomer, however, could not bind to the Golgi by itself (a point also demonstrated independently by Richard D. Klausner and his colleagues at the National Institutes of Health). Next, electron microscopy revealed that when only ARF binds to the Golgi membrane, the membrane remains flat. But when coatomer is added with ARF, a bud develops. These observations established that ARF and coatomer have distinct roles in producing transport vesicles. Coatomer drives curving of the membrane and budding, whereas ARF determines when budding occurs (by dictating to coatomer when and where to do its job).

Actually, the budding process is known in more detail. ARF that floats free in the cytosol is normally connected to a molecule of guanosine diphosphate



(GDP): a base linked to a sugar and two phosphate groups. When this unit encounters the Golgi apparatus, an enzyme at the membrane surface replaces the GDP with a relative obtained from the cytosol: guanosine triphosphate (GTP), which carries three phosphate groups. This modification is significant for two reasons. First, only ARF bound to GTP can hook onto the membrane surface firmly, thereby enabling coatomer complexes to come on board. Second, binding of GTP to ARF is an important step in energizing vesicle formation.

How should we picture the budding process? Individual coatomers are anchored to the membrane by one or more ARF molecules; the exact number probably is neither fixed nor critical. We think (but have not yet proved) that these complexes then link up, side by side in a regular array, to form a dome (something like that made by clathrin) on the surface of the membrane. As the array grows and curves, the part of the Golgi membrane attached to the inside of this coat is sculpted into a budding vesicle that eventually closes and is released. Hence, COP-coated vesicles arise much as was predicted by the model Kanaseki and Kadota proposed for clathrin-coated vesicles. In a modification of the model, however, our work suggests that a new coat must assemble piece by piece from cytosolic components each time a vesicle buds. Moreover, assembly is directed by ARF and requires GTP for energy.

Although the coat is essential to the budding of COP-coated vesicles, it impedes fusion of the vesicle's membrane with that of the target organelle; such fusion requires that the membranes contact one another directly. And so, for the vesicle's contents to be released, the coat must be shed. This disrobing happens by a simple reversal of the coat-assembly mechanism. The ARF molecules in a fully formed coated vesicle are still linked to GTP. But soon they clip a phosphate from the GTP, thereby converting it to GDP and releasing energy stored in the GTP. This conversion causes the ARF molecules to lose their affinity for the vesicle membrane and to fall off. Because the coatomer units require ARF for their own attachment, they dissociate as well, leaving behind an uncoated vesicle free to merge with a target organelle. In a way, then, the cell plants little GTP "time bombs" in the coat during budding. After budding, the charges are detonated to eject the coat.

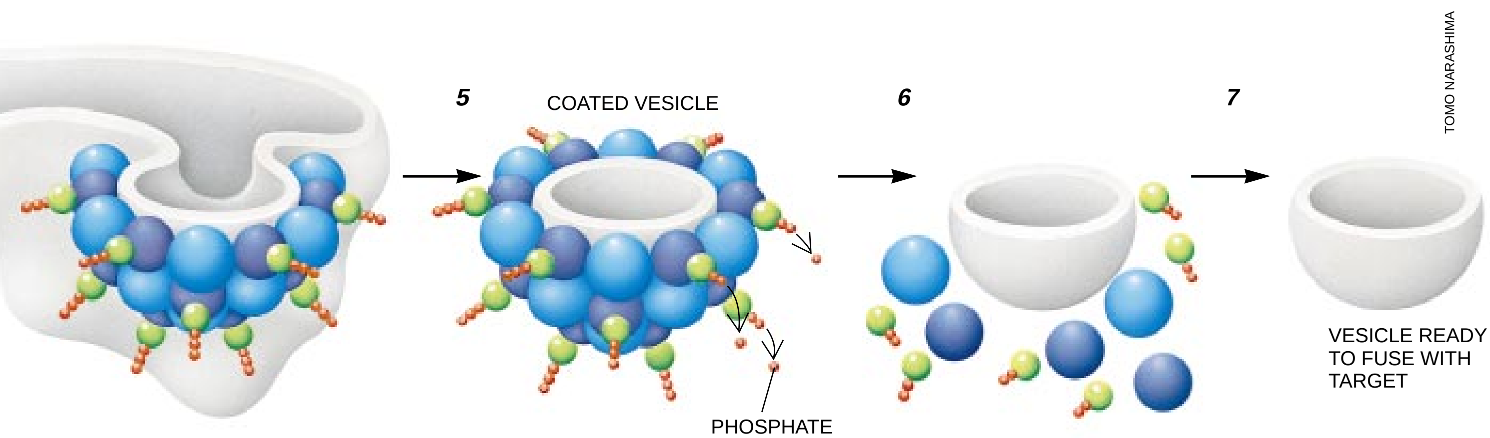
Because the coat is released before the vesicle reaches its destination, it is unlikely to be involved in directing a vesicle to its target. Rather the uncoated vesicle membrane contains a set of proteins that perform this function. These proteins also attract added proteins that initiate fusion of the vesicle with its target membrane.

A Common Mechanism

Do all transport vesicles arise much as COP-coated vesicles do? The first hint that they might was the discovery that assembly of the clathrin coat is also triggered by the attachment of GTP-bound ARF molecules to membranes. This finding emerged in Rothman's laboratory and brought us full circle to where we began.

Yet when most vesicles in cells are viewed in a microscope, they do not seem to have coats. This observation could suggest that the budding mechanism we have described reveals nothing about how vesicles other than COP-

ELECTRON MICROSCOPY has revealed that coatomer controls the budding of COP-coated vesicles. When ARF alone is allowed to bind to Golgi membranes, no buds develop; the membranes remain flat (*left*). But attachment of coatomer to membranes that have already been bound by ARF leads to massive production of coated vesicles (*right*).



FORMATION OF COP-COATED VESICLE on a Golgi membrane is followed quickly by removal of the coat. The vesicle begins to develop (*1*) when a protein called ARF, carrying a smaller molecule called GDP, contacts the membrane. Soon (*2*) an enzyme replaces the GDP with GTP—a change that enables ARF to recruit a complex of proteins called a coatomer (*3*). Assembly of ARF and coatomer units on the membrane causes a bud to form (*4*) and then pinch off (*5*) from the membrane as a coated vesicle. The coat must be removed for the vesicle to fuse with its target organelle. Removal occurs when conversion of GTP to GDP, by release of a phosphate group (*at right in 5*), causes ARF and coatomer molecules to fall away (*6* and *7*).

and clathrin-coated types form. Such a conclusion would be premature, however. Because coats are removed rapidly after a vesicle buds from an organelle, standard methods of observation generally fail to detect them. Indeed, the history of this field to date can be summarized as follows: Investigators initially conclude that a given type of vesicle lacks a coat and thus buds by a novel mechanism. Then the coat is discovered after the budding of that type of vesicle is reproduced in a cell-free system, which allows the transient, coated state to be trapped and identified.

Among the vesicles recently found to require a coat are those that carry proteins from the endoplasmic reticulum to the Golgi apparatus. Experiments conducted by Randy W. Schekman of the University of California at Berkeley, Orci and their co-workers have demonstrated that a close relative of ARF plays ARF's role, in that it initiates the budding process. Further, a different set of proteins, called COP II proteins, take the place of the coatomer.

There is now little doubt that all nucleated cells use a common set of principles for forming transport vesicles. The process is so effective that it was retained when simple, single-cell species evolved into ever more complex multicellular creatures. First, ARF or a close relative initiates formation of the coat by taking up GTP and binding to the membrane. Next, this initiator recruits additional components, which assemble into a domelike coat. In so doing, they induce the attached region of the membrane to bud off as a vesicle. Later, the GTP molecules trigger removal of the coat to allow for fusion with the target organelle.

If the coat is important only for budding, a process that is basically the same in every organelle, why does a cell need many kinds of coats? The most likely reason is that the coat selects the cargo to be packaged into a vesicle. In some cases, the cargo proteins may reside in the membrane itself and bind directly to the coat. In other cases, the cargo may be linked to the coat through an intermediary—a receptor—that spans the membrane. Use of separate coats would allow different kinds of cargo to be shipped out from a single point of origin and from different departments.

Other questions remain as well, many concerning fine details of coat formation. For instance, precisely how do ARF and similar molecules attach to membranes, and which exact interactions enable coatomer proteins to be anchored to membranes? Recent evidence suggests that certain lipids (fats) may help ARF recruit coatomer to the surface and that cargo proteins themselves might influence assembly of a coat. We would also like to learn more about the specific activities of the various coatomer subunits. Nevertheless, it is satisfying to know that many central steps in protein transport have now been defined. At the same time, it is humbling to note that no city organized by human beings can match the elegant efficiency with which cells of every stripe convey their laborers from place to place.

The Authors

JAMES E. ROTHMAN and LELIO ORCI have collaborated for more than a decade. Rothman is chairman of the cellular biochemistry and biophysics program at Memorial Sloan-Kettering Cancer Center. In 1976 he received his Ph.D. in biological chemistry from Harvard Medical School. Before joining Sloan-Kettering in 1991, he was professor of molecular biology at Princeton University and, before that, professor of biochemistry at Stanford University. He is a member of the National Academy of Sciences. Orci is chairman of the department of morphology and professor of histology and cell biology at the University of Geneva Medical School in Switzerland. He obtained his M.D. at Rome University in 1964 and moved to Geneva in 1966.

Further Reading

A NEW TYPE OF COATED VESICULAR CARRIER THAT APPEARS NOT TO CONTAIN CLATHRIN: ITS POSSIBLE ROLE IN PROTEIN TRANSPORT WITHIN THE GOLGI STACK. L. Orci, B. S. Glick and J. E. Rothman in *Cell,* Vol. 46, No. 2, pages 171–184; 1986.

"COATOMER": A CYTOSOLIC PROTEIN COMPLEX CONTAINING SUBUNITS OF NON-CLATHRIN-COATED GOLGI TRANSPORT VESICLES. M. G. Waters, T. Serafini and J. E. Rothman in *Nature,* Vol. 349, pages 248–251; January 17, 1991.

ADP-RIBOSYLATION FACTOR IS A SUBUNIT OF THE COAT OF GOLGI-DERIVED COP-COATED VESICLES. T. Serafini, L. Orci, M. Amherdt, M. Brunner, R. A. Kahn and J. E. Rothman in *Cell,* Vol. 67, No. 2, pages 239–253; October 18, 1991.

STEPWISE ASSEMBLY OF FUNCTIONALLY ACTIVE TRANSPORT VESICLES. J. Ostermann, L. Orci, K. Tani, M. Amherdt, M. Ravazzola and J. E. Rothman in *Cell,* Vol. 75, No. 5, pages 1015–1025; December 3, 1993.

MECHANISMS OF INTRACELLULAR PROTEIN TRANSPORT. J. E. Rothman in *Nature,* Vol. 372, pages 55–63; November 3, 1994.

The Art and Science of Photoreconnaissance

U.S. AIR FORCE

In the 1950s and 1960s, photointerpreters devised ways of extracting valuable information from recondite images. Oftentimes, their work profoundly affected international relations

by Dino A. Brugioni

Every morning for the past 35 years, a Central Intelligence Agency officer has appeared at the White House with a printed intelligence digest and a fistful of images. These aerial and satellite photographs, perhaps half a dozen chosen from the thousands shot the previous day, show current hot spots, with attached explanatory notes. Similar packets are delivered to the secretaries of state and defense, selected members of Congress and other key officials. Over the decades, these images have had a powerful influence on U.S. policy.

Reconnaissance photographs have aided every presidential administration since that of Dwight D. Eisenhower in watching over troops stationed abroad, monitoring disarmament agreements, assessing hostile military forces and planning counterweapons programs. During its early years—the late 1950s and early 1960s—modern photoreconnaissance repeatedly provided timely intelligence, sometimes even helping to bring the superpowers back from the brink of conflict. More often than not, these images prevented dangerous surprises and showed arsenals to be less imposing, and intentions less sinister, than had been thought.

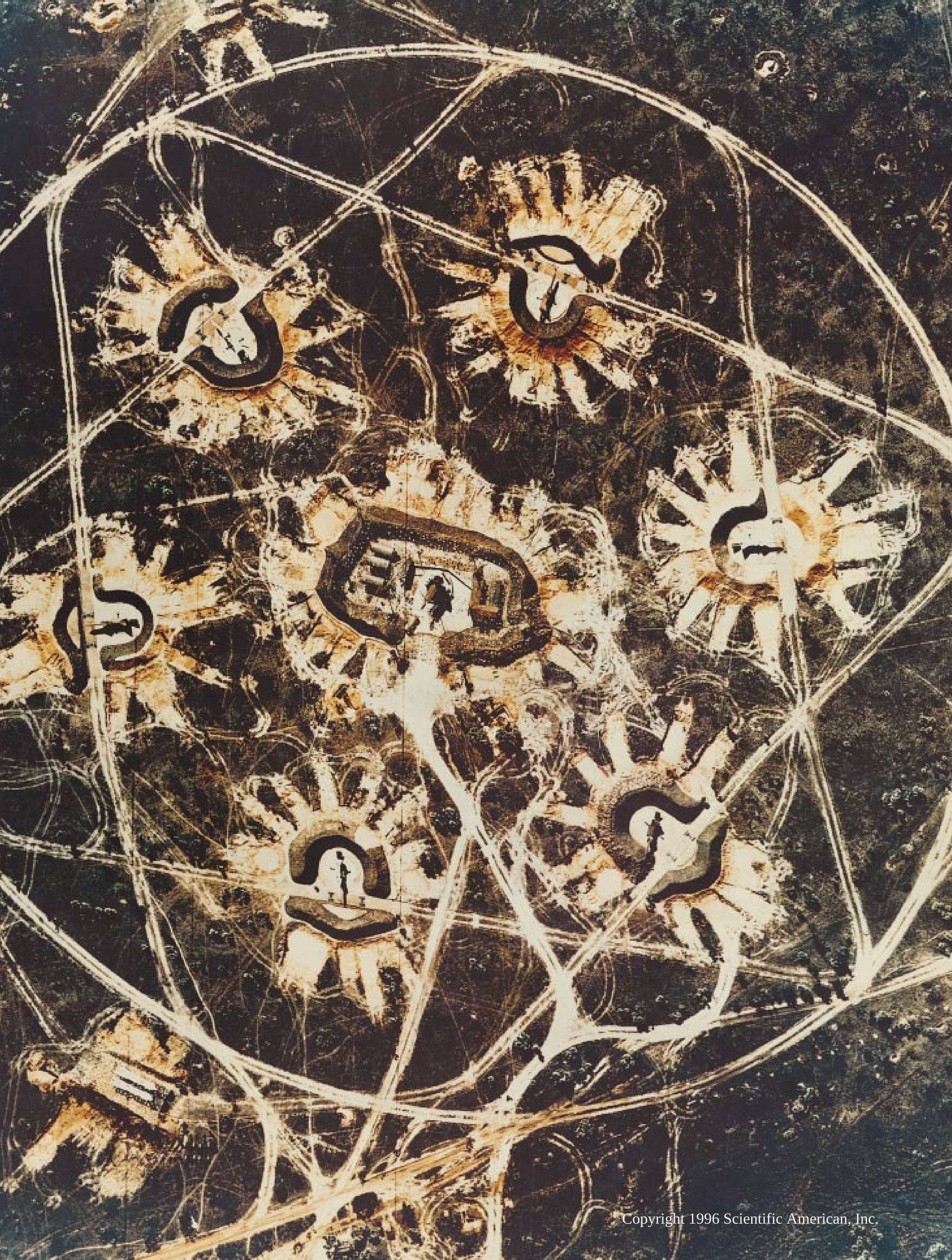
Those of us who handled these photographs every day saw an ever changing and fascinating pageant. Now almost a million of the images from a formative period in photoreconnaissance are being made public. In conformance with an executive order signed on February 23, 1995, the U.S. government is to release more than 800,000 satellite photographs collected by the CIA during the first 12 years of satellite photoreconnaissance, from 1960 to 1972. So far thousands have been released, with the remainder to be made available by August.

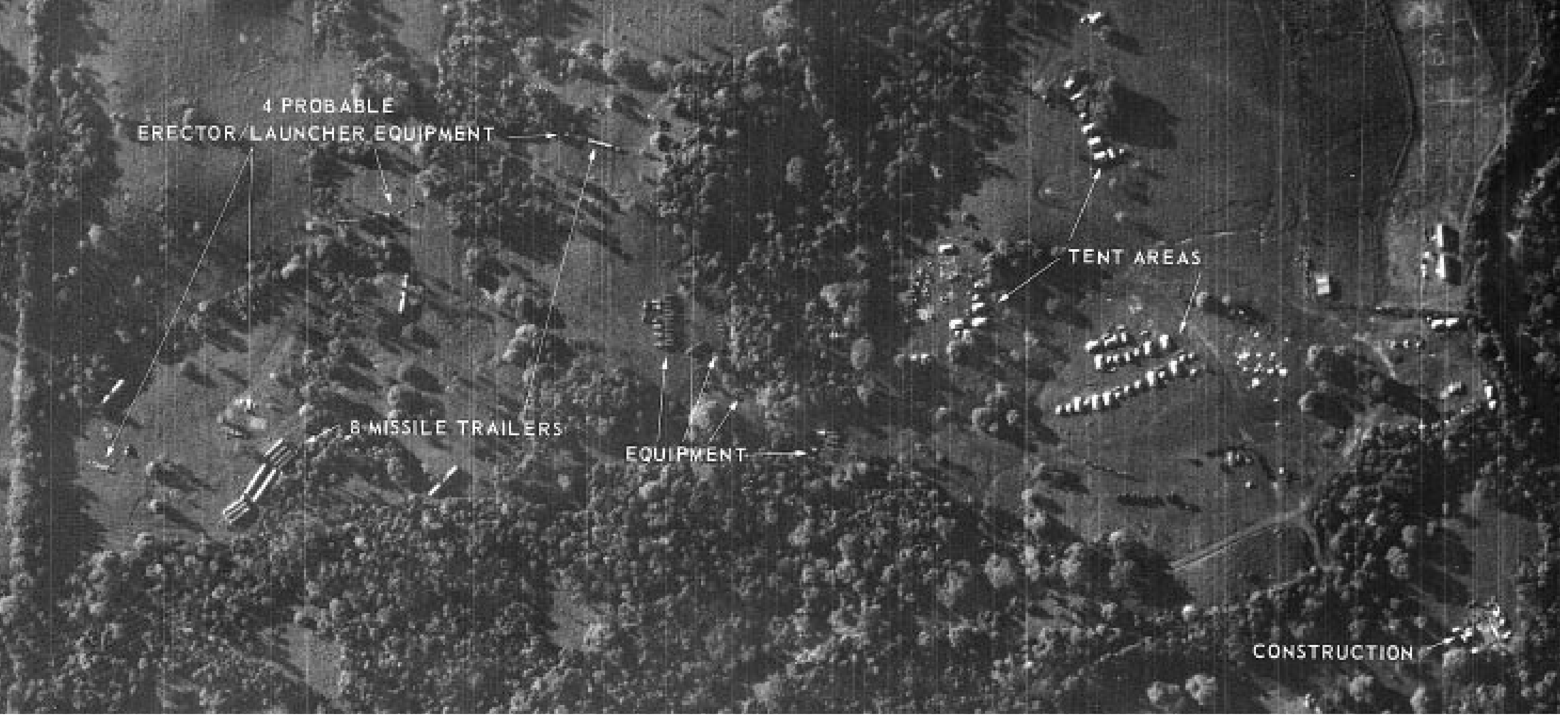
The period covered was a pivotal one, during which the Soviet Union built its first intercontinental ballistic-missile launch sites, the Vietnam War was fought, the number of nations with nuclear weapons expanded from four to at least six, and the U.S. and the U.S.S.R. conducted the Strategic Arms Limitation Talks (SALT). The era began with the construction of plutonium-reprocessing facilities in Israel and continued with the Cuban missile crisis, the detonation of the first Chinese atomic weapon, the space race and the Six-Day War in the Middle East. All these events took place in full view of the U.S.'s eyes in the sky. The photographs they took revolutionized the country's understanding not only of the Soviet Union and China but, arguably, of the entire world.

Continued on page 82

NOVEMBER 10, 1962
LA COLOMA, CUBA

From above, Soviet SA-2 surface-to-air missile sites each resemble a six-point Star of David. In the center is the guidance-and-control radar station; surrounding it are six missiles on launchers (each near a star point). First deployed in the late 1950s and still in use today, SA-2 batteries were designed to shoot down attack or reconnaissance aircraft, so detection of these signatures in images of Cuba made shortly before the missile crisis helped to tip off U.S. photointerpreters that something worth shielding was being installed. The objects being protected by these batteries turned out to be medium- and intermediate-range ballistic-missile sites.





U.S. NAVY

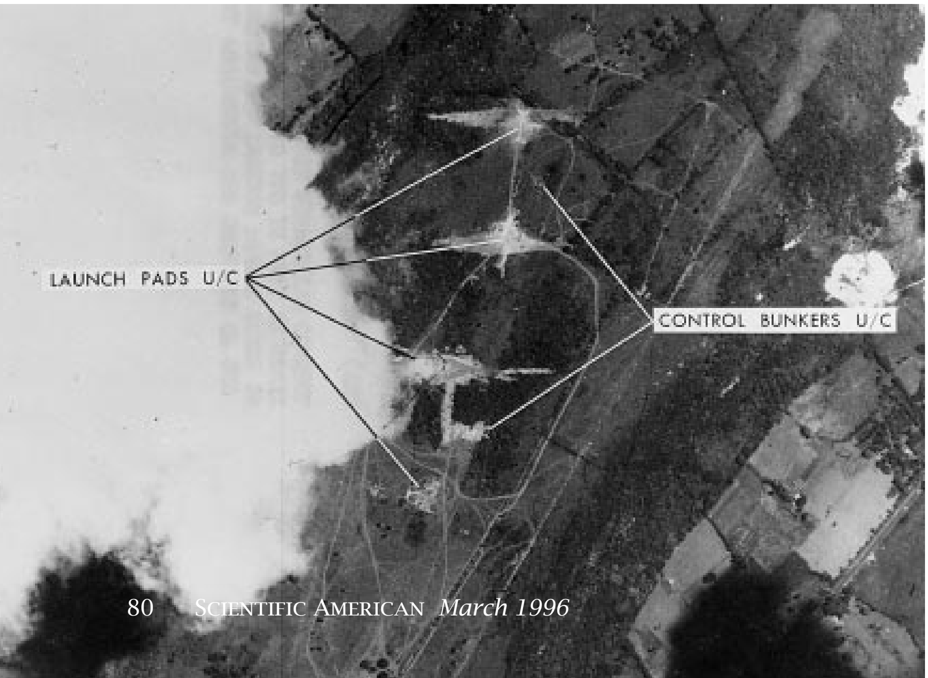
OCTOBER 14, 1962
SAN CRISTÓBAL, CUBA

On the morning of October 16, 1962, a 48-year-old career intelligence officer named Arthur C. Lundahl stood nervously before a lectern in the cabinet room of the White House, presenting this photograph to President John F. Kennedy and his key advisers. Lundahl, then the U.S.'s top photointerpreter, was explaining why this image, taken two days earlier, suggested that Soviet medium-range ballistic missiles had been deployed on the island of Cuba. With a range of 1,800 kilometers, the missiles could reach Washington, D.C., and many other U.S. locales to the south. Snapped by a U-2 spy plane at an altitude of 21,300 meters, the picture shows a cluster of seven missiles, a few launchers and "missile-ready tents," which were used to prepare the missile and warhead for launch.

Carefully spacing his words and looking Lundahl in the eye, Kennedy asked, "Are you sure?"

"Mr. President," Lundahl replied, "I am as sure of this as a photointerpreter can be sure of anything. And I think, sir, you might agree that we have not misled you on anything we have reported to you. Yes, I am convinced they are missiles."

Kennedy ordered enough U-2 flights to permit complete coverage of the island. They revealed that in addition to the medium-range missiles, launchpads for intermediate-range missiles—with twice the range—were under construction. These IRBMs could reach targets anywhere in the continental U.S., except for the Pacific Northwest. During his showdown with the Soviets, Kennedy knew, thanks to photoreconnaissance, that the U.S. had at least seven times as many strategic assets—long-range bombers, ICBMs and missile-launching submarines—as did the Soviets.



U.S. AIR FORCE

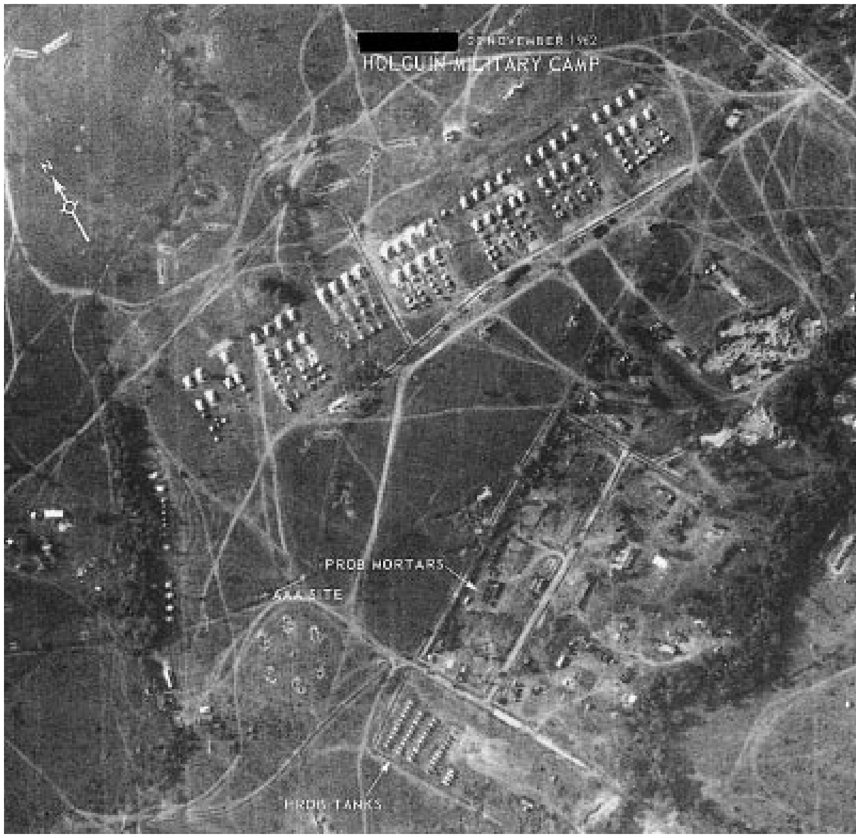
OCTOBER 15, 1962
GUANAJAY, CUBA

During the Cuban missile crisis, photointerpreters relied heavily on the distinct appearance, or "signature," of missile launch sites in the Soviet Union. The unique slash marks indicative of an SS-5 intermediate-range missile site in an early stage of construction had been seen many times in the Soviet Union, so interpreters knew exactly what they were. As soon as they were identified in Cuba, Kennedy was informed that the medium-range missiles already in Cuba would soon be joined by the much longer range intermediate missiles. The president's quick response—a blockade of the island—kept Soviet ships from delivering the SS-5 missiles to Cuba.

U.S. AIR FORCE

NOVEMBER 20, 1962
HOLGUÍN MILITARY CAMP, CUBA

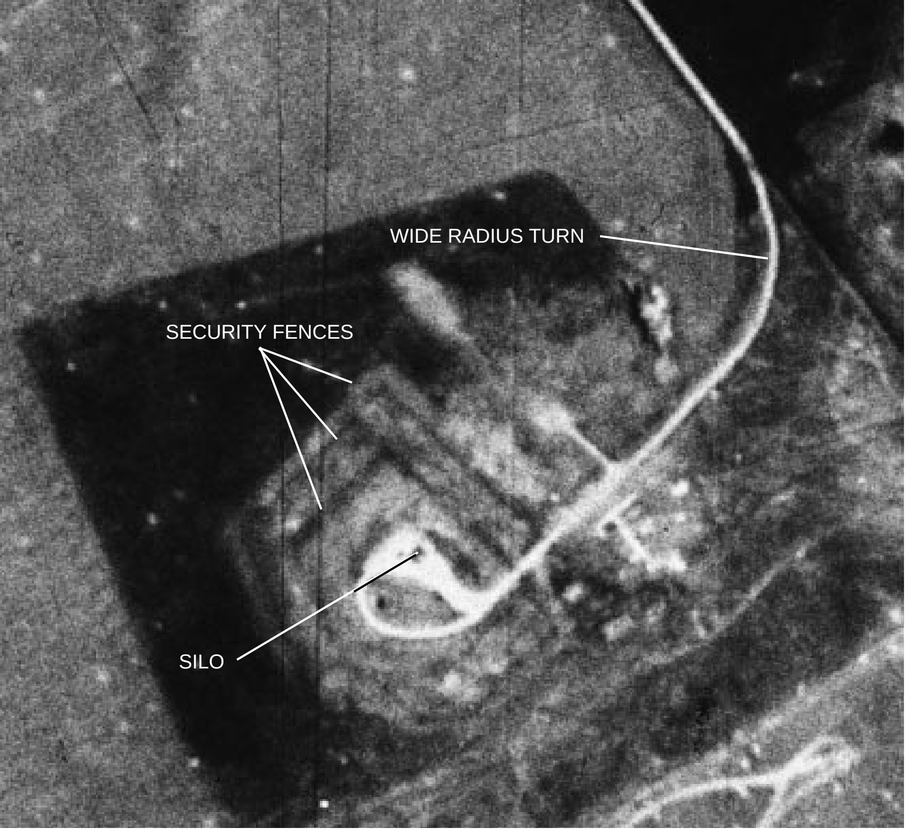
Image analysts developed some unique disciplines, including "tentology," the numerical assessment of a military force by counting the tents that are sheltering it. The number and types of tents at four Soviet compounds in Cuba during the missile crisis indicated that each compound had a force of about 1,500 troops. The precise positioning of the tents and the orderly arrangement of the latest military equipment led interpreters to conclude that these were elite Soviet troops rather than Cuban forces.



U.S. AIR FORCE

AUGUST 18, 1962
ATLANTIC OCEAN OFF CUBA

"Crateology" is the scrutiny of the shape, protrusions and shipping hooks of the crates within which military equipment is shipped. The crates protect the equipment from the weather, salt spray on board ships—and prying eyes. But crates still may be photographed being loaded or unloaded by eyes in the sky (or spies in the port). Even if they are not, careful analysis and photogrammetry, along with the crates' point of origin, often allow the contents of crates to be identified. In 1962 photointerpreters were able to identify crates containing Komar guided-missile patrol boats (shown here in a photograph taken by a low-flying U.S. Navy aircraft). MiG-21 jet fighters, Il-28 Beagle bombers and Il-14 Crate transport aircraft being sent by ship to Cuba were also recognized.



NATIONAL RECONNAISSANCE OFFICE

SEPTEMBER 8, 1967
IMENI GASTELLO ICBM COMPLEX, U.S.S.R.

U.S. photointerpreters were fortunate that the lack of roads in the vast U.S.S.R. forced the Soviets to rely on their rail network for the establishment of ICBM complexes. This reliance on rail required a few additional steps during the construction of the missile sites, which proved of great use to U.S. analysts in ascertaining that an ICBM complex was, in fact, being built. First-class roads, which were wider than ordinary roads and had wide-radius turns to accommodate the long missile transporters, were built to move the missiles to the silos. In this image the silo appears as a dark spot against the light-colored, triangular expanse of concrete, at the end of an access road. Three high-security fences surround the silo, offering more evidence of the site's importance.

Continued from page 78

Arthur C. Lundahl, the first director of the National Photographic Interpretation Center, once estimated that the intelligence community skimmed only about 15 percent of the information contained in these films. The released imagery will open new vistas (literally and figuratively) on the political history of the mid-20th century as well as on subjects ranging from archaeology to zoology.

All the released images, and those yet to be released, were made under a CIA program code-named Corona, the U.S.'s pioneering project in satellite reconnaissance. (The air force carried out a smaller, competing program, called Samos, between 1958 and 1963.) The Corona and Samos satellites were the most revolutionary of the several platforms that formed the foundation of modern reconnaissance in the U.S. The others were the aerial photographic balloons (under a short-lived program known as Genetrix) and the fast, high-flying U-2 and SR-71 airplanes.

President Eisenhower, who championed photoreconnaissance, was briefed on the results of every U-2 and satellite mission. At a time when the surprise attack on Pearl Harbor still resonated, Eisenhower saw the technology as a means of preventing another such traumatic experience in the future. He also seized on it as a way of improving the assessments of the strategic capabilities of the Soviet Union and China.

Indeed, toward the end of his administration, Eisenhower identified photoreconnaissance as one of the few recourses against an unbridled U.S. military-industrial complex that inflated Soviet capabilities to suit its own ends. Although the U.S. spent many billions of dollars over the years on photoreconnaissance, the information that was gathered more than repaid the outlay by saving untold billions in defense dollars. Besides revealing an adversary's true strength, photoreconnaissance allowed more efficient planning of counterweapons programs.

Corona officially began in January 1958. The 13th—and first successful—satellite was launched on August 18, 1960. In all, 94 satellites were successfully orbited as part of the Corona program. The orbiters themselves went through several generations, identified in the most enduring set of code names as Keyhole (KH) 1, 2, 3 and 4. They were followed by seven KH-5 satellites under the Lanyard program and one KH-6 under Argon.

When all went well, the satellites were launched into elliptical orbits ranging from about 100 to 500 miles above the earth. They would take thousands of photographs during 16 or more orbits, each about an hour and a half long. (After 1962, dual cameras were used to make stereoscopic images.) At the right moment, a retrorocket would fire the film capsule back toward the earth. Drifting down under a parachute, the capsule and its precious cargo were snatched out of the sky by an air force plane.

Telltale Signatures

The acquisition of a satellite photograph—as daunting a process as it may be—is only the beginning. Photointerpreters must then detect, identify, describe and assess the objects and patterns in a photograph. The process is painstaking, sometimes tedious and often intuitive. Photointerpreters analyze an image by identifying "signatures," which are recognizable features or patterns of special interest. For example, an interpreter can always spot a military armored unit by its associated series of bowling-alley-like lanes, which are actually tank firing ranges.

These signatures are carefully catalogued in books of photointerpretation keys. Each book covers a particular subject and country, such as the missiles deployed by China. Interpreters use the books, which are updated often, both at headquarters and in the field.

As might be expected, photointerpretation had its baptism by fire during World War II. Allied photointerpreters devised a system based on three distinct phases of information dissemination. During the first, or "flash," phase, information deemed vital in combat was rushed to commanders in the field. Second-phase reporting involved closer study of the photographs, leading to a

NATIONAL RECONNAISSANCE OFFICE

written report. The third phase was similar, but the analysis was even more detailed. With few exceptions, the system continues to this day. Since 1961 the activity has been consolidated at the National Photographic Interpretation Center, which performs the task for the entire U.S. intelligence community.

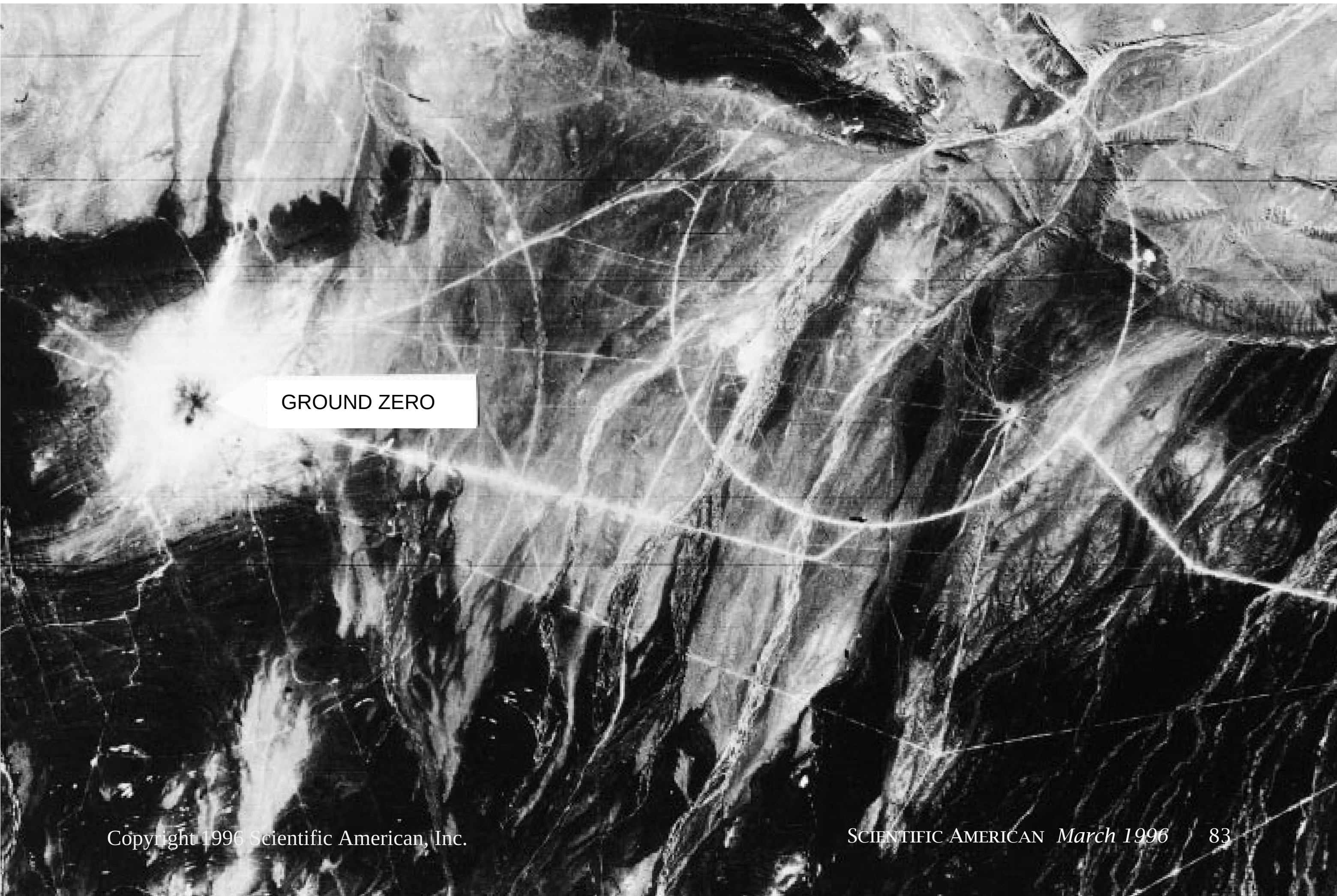
Digital-image storage and manipulation have, of course, transformed photointerpretation. These technologies enable more compact storage of photographs, faster and more controllable access to imagery, and image processing to bring out features of interest. And they allow photointerpreters to call up, for comparison, a variety of past images showing the same or similar sites, perhaps configured for different purposes. Photogrammetrists can determine the dimensions of any object or site in the picture by entering satellite and camera data (such as the altitude and attitude of the satellite and the focal length of its lens) into a computer-based device called a comparator.

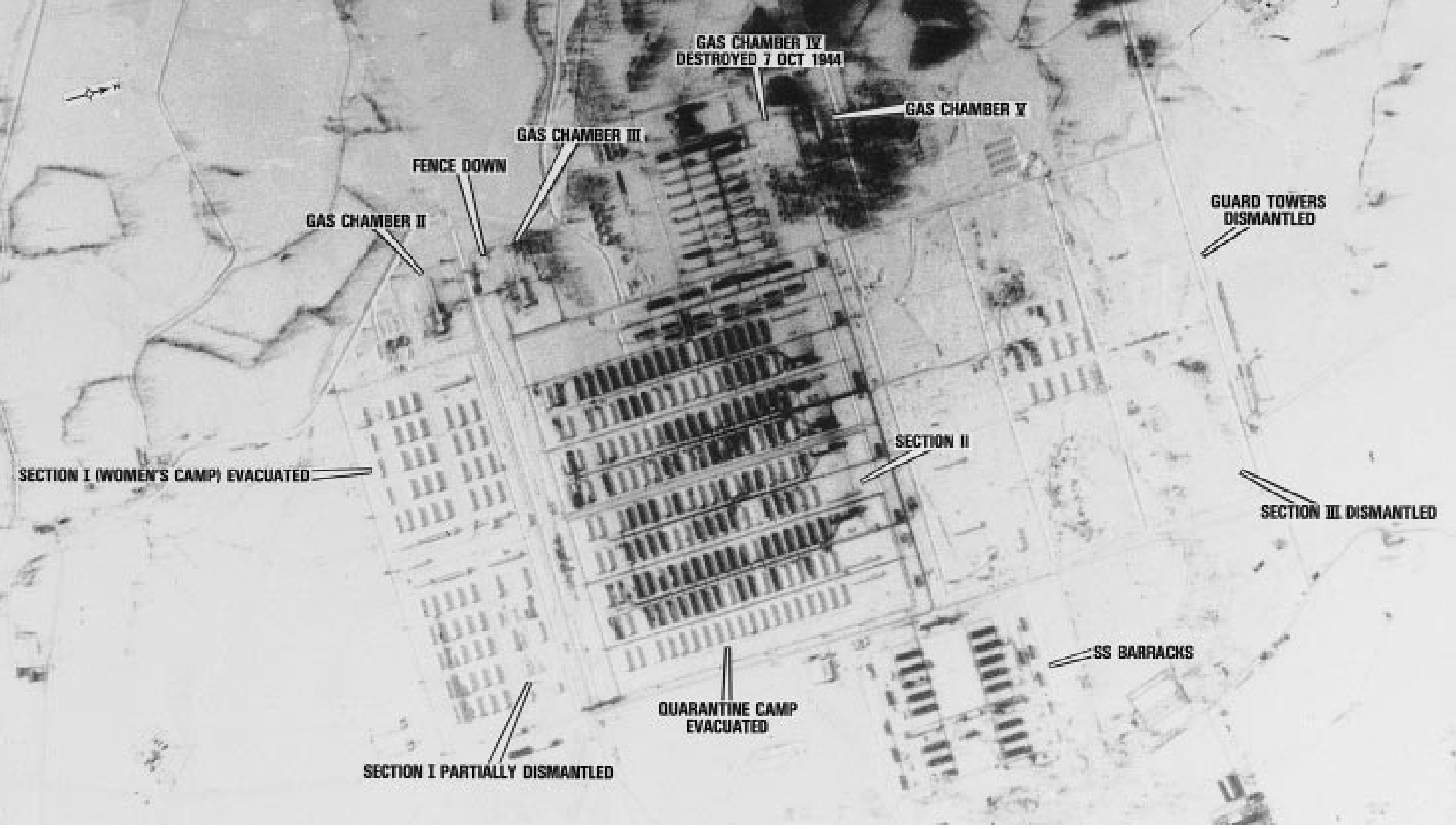
Other important resources of the photointerpreter are human nature and, sometimes, the climate and weather. People in general, and military personnel in particular, live according to hab-

OCTOBER 20, 1964
LOP NUR, CHINA

As more nations began building missiles and weapons of mass destruction, deserts and other remote areas became proving grounds. Analysts always searched photographs of such regions carefully, and occasional discoveries allowed them to forecast momentous events. In December 1961, photographs of the Lop Nur area in the Taklimakan Desert in western China revealed a unique circular road, 4,000 meters in diameter. The road's location and dimensions made it a likely part of a planned nuclear test site; the area encompassed by the road was large enough to enclose an aboveground nuclear test while leaving the road intact. An airfield, barracks and support facilities were subsequently constructed. Still later, a 100-meter tower was erected, and when communications and instrumentation lines were being dug from the tower to electronic vans and bunkers, a test was only weeks away.

Lundahl communicated these developments to the director of the Central Intelligence Agency, John McCone. To keep the Chinese government from reaping a propaganda bonanza from the test, Lundahl noted, President Lyndon B. Johnson might wish to make an announcement about the impending detonation. Johnson deferred to Secretary of State Dean Rusk, who, on September 29, 1964, announced to the press that "for some time it has been known that the Communist Chinese were approaching the point where they might be able to detonate a first nuclear device." Analysis of October 8 photography showed all preparations for the test were complete, with workers and equipment withdrawn from the test area. There was hardly an international ripple when the Chinese did test a 28-kiloton atomic device on October 16. Images made four days later with KH-4 satellites showed the effects of an atmospheric test where the tower once stood.





JANUARY 14, 1945
AUSCHWITZ-BIRKENAU DEATH CAMP, POLAND

Allied reconnaissance in support of the bombing missions against the IG Farben synthetic rubber and fuel plant in 1944 and 1945 sometimes inadvertently produced images of the nearby Auschwitz-Birkenau death camp. The photographs were not analyzed until 1978, when the author and Robert G. Poirier discovered them in U.S. Defense Department archives. This image, taken as Russian troops approached on January 14, 1945, shows the snow that Elie Wiesel, who was in the complex at the time, wrote of in *Night.* The photograph indicates that the gas chambers have been or are being destroyed and that the evacuation of the complex had begun.

Heavy snow cover on the roofs in the women's camp reveals that it has been evacuated. (Most of the women, including Anne Frank, were sent to Belsen; most of the men were later sent to Buchenwald and other camps.) When this picture was taken, the men's camp and the SS barracks were still largely occupied, as indicated by the fact that the heat in the buildings has melted the snow on their roofs. Each barracks in the men's camp housed about 1,000 inmates, so when this photograph was made about 80,000 inmates were still incarcerated (the camp usually held 250,000). After the able-bodied were evacuated, 8,000 sick or emaciated prisoners were left behind and were liberated by Russian troops on January 27.

its, rules and customs. Want to know which are the most important buildings in a military compound? Wait for a snowfall. It has been my experience that the roads to the headquarter buildings are always cleared of snow first, along with the pathways to the latrines. The buildings that are occupied are the ones that are heated—and on whose roofs, therefore, the snow has melted.

Essentially everywhere in the world, Sunday is a day when most people relax, so Sunday mornings (and holidays) are the best time to inventory an opponent's military equipment, when it is in garrison, parked or stowed. Capabilities of ground-force units are best ascertained during training exercises, which are usually held in the spring and summer, when the ground is firm and the troops are unrestricted by heavy clothing or bad weather.

The images in this article illustrate only a minute sampling of the events captured by photoreconnaissance. Among the hundreds of developments detailed were the results of the first major nuclear accident, in 1957 at Kyshtym in the Soviet Union; the 1962 Sino-Indian border war; the effects of the Israeli air strikes on Egyptian, Jordanian and Syrian air forces during the Six-Day War; the Vietnam conflict; the India-Pakistan clashes of 1965 and 1971; the Soviet invasion of Czechoslovakia; and the buildup of military forces in the

AUGUST 29, 1962 AND AUGUST 1–10, 1994 ARAL SEA, KAZAKHSTAN AND UZBEKISTAN

The Aral Sea, whose sources have been dammed and siphoned for irrigation, has shrunk alarmingly over the past 32 years, as shown in these two satellite images. (The recent image (*right*) was pieced together from several photographs made over a week and a half.)

U.S. DEPARTMENT OF DEFENSE

NATIONAL RECONNAISSANCE OFFICE

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

1970s along the Sino-Soviet borders. The Corona program has also provided wide-area coverage of remote and seldom seen regions of the world—including Novaya Zemlya, where the Soviets tested their nuclear weapons; Tibet, which China usurped in 1950; and the Kalahari Desert.

Ravages Revealed

From our unique vantage point, my co-workers and I came to appreciate the exceptional fragility of our globe. Over and over again, we saw sparsely vegetated areas scoured for firewood, the cutting or burning of vast tracts of equatorial forest, the effects of industrial or natural disasters, and the pollution of the skies, streams and rivers. We watched, winter after winter, as pollution blackened the snow around Magnitogorsk, a Russian steel-producing center, to a distance of 150 kilometers. What we saw would have given pause to anyone suffering from the delusion that the earth can benignly absorb whatever ravages humankind serves up.

The release of hundreds of thousands of Corona images presents a rare opportunity to better understand this fascinating, ceaselessly changing blue planet and its inhabitants. I have, for example, tracked the steady northward expansion of the Sahara in images made periodically during the late 1960s. In the same images, I could see where the Romans had put their forts and compare their positions with those of the French Foreign Legion. One December morning in the late 1960s, I selected a series of pictures from one pass of a KH-4 over Nazareth and Bethlehem. Scrutinizing them in a stereoscopic viewer, I followed the Holy Family's biblical journey over each hill and down each valley. I have analyzed the terrain where the Charge of the Light Brigade took place, and I have followed Marco Polo on his travels (to my amazement, many of his descriptions still held).

As Lundahl noted 25 years ago, we have only glanced at many of these images. For future researchers, navigating this uncharted sea of data will prove engrossing. It was, often enough, thrilling the first time around.

The Author

DINO A. BRUGIONI received B.A. and M.A. degrees in foreign affairs from George Washington University. He flew in 66 bombing missions with the U.S. Army Air Corps during World War II. In 1948 he joined the fledgling Central Intelligence Agency and was one of the founders of the National Photographic Interpretation Center, the interagency organization responsible for U.S. imagery analysis. He is the author of *Eyeball to Eyeball: The Inside Story of the Cuban Missile Crisis* and has written extensively on the application of aerial and spatial imagery to intelligence and environmental problems.

Further Reading

THE HOLOCAUST REVISITED: A RETROSPECTIVE ANALYSIS OF THE AUSCHWITZ-BIRKENAU EXTERMINATION COMPLEX. Dino A. Brugioni and Robert G. Poirier. Central Intelligence Agency (ST-79-10001), 1979.

DEEP BLACK: SPACE ESPIONAGE AND NATIONAL SECURITY. William E. Burrows. Random House, 1986.

DRAGON LADY: THE HISTORY OF THE U-2 SPYPLANE. Chris Pocock. Airlife Publishing/Motorbooks International, 1989.

LOCKHEED'S SKUNK WORKS: THE FIRST FIFTY YEARS. Jay Miller. Midland Publishing, Leicester, 1993.

CORONA: AMERICA'S FIRST SATELLITE PROGRAM. Edited by Kevin C. Ruffner. Central Intelligence Agency History Staff and the Center for the Study of Intelligence, 1995.

CORONA: SUCCESS FOR SPACE RECONNAISSANCE, A LOOK INTO THE COLD WAR, AND A REVOLUTION FOR INTELLIGENCE. Robert A. McDonald in *Photogrammetric Engineering and Remote Sensing,* Vol. 61, No. 6, pages 689–720; June 1995.

Electrons in Flatland

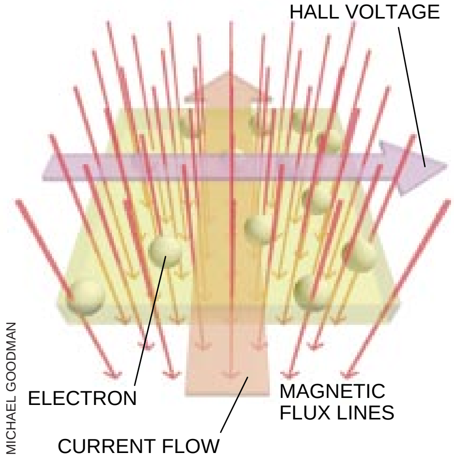
Trapped in a two-dimensional plane, electrons can exhibit the quantum Hall effect, a startling phenomenon now thought to be intimately connected to superconductivity

by Steven Kivelson, Dung-Hai Lee and Shou-Cheng Zhang

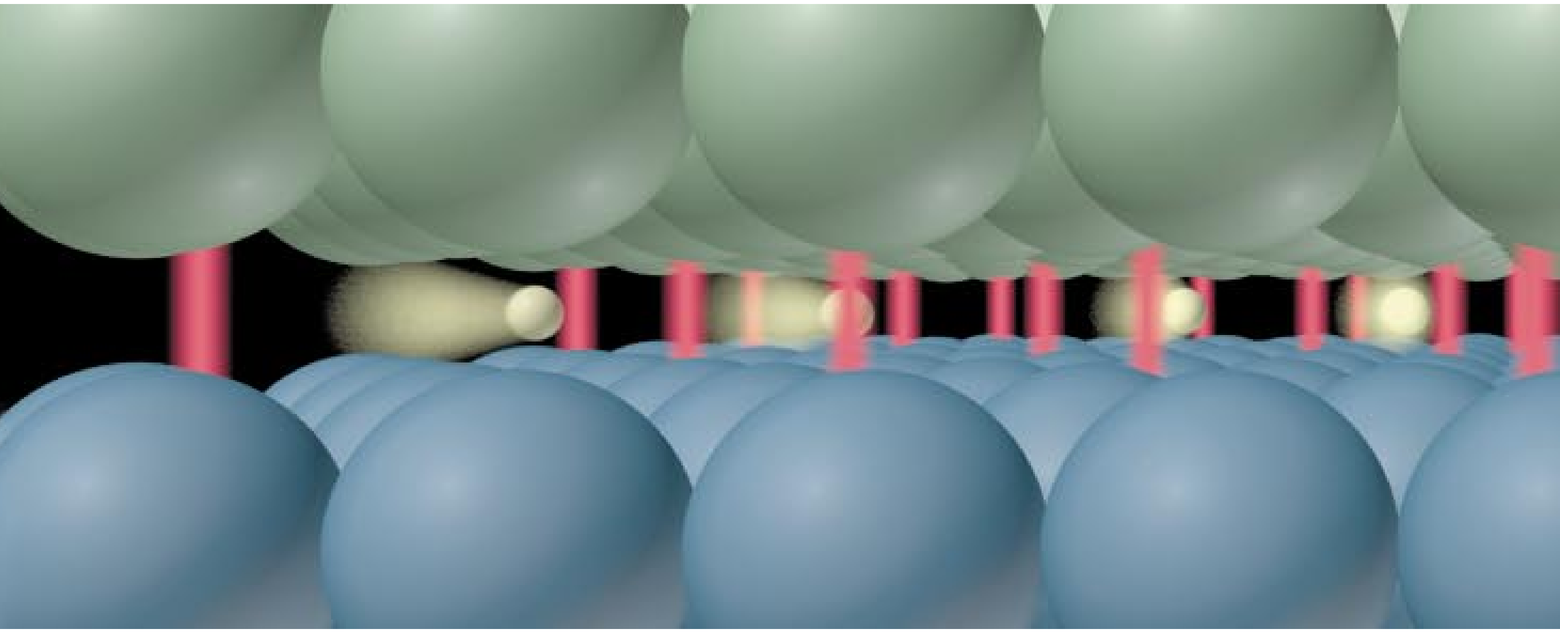
Since the time of the ancient Greeks, a central goal of all scientific disciplines has been to find a minimal set of basic principles that underlie diverse natural phenomena. This reductionist philosophy has succeeded well in some areas, such as high-energy physics—the study of the fundamental particles of force and matter. Here theorists have grouped all particles into a few families and expressed the basic laws of physics in terms of the interactions among them.

The situation is quite different in condensed-matter physics, which is the study of solids and liquids. Research in this century into the behavior of electrons in solids has uncovered various states of matter in which electrons organize themselves in myriad remarkable ways. For example, solids are typically either insulators (they strongly resist the flow of electric current) or metals (they conduct current well but still have a small amount of resistance). Yet under some circumstances, certain solids can enter a superconducting state, in which electric current flows without any resistance at all. The theoretical characterizations of these different states have been as diverse as the states themselves.

That may soon change. Researchers have found a deep connection between superconductivity and another intensely studied subject in condensed-matter physics: the quantum Hall effect. This phenomenon occurs when electrons are subjected to three specific conditions at once: they are trapped at the interface between two semiconductor crystals, so that they can move only in a two-dimensional "flatland"; they are cooled to temperatures near absolute zero; and they are subjected to a high magnetic



QUANTUM HALL EFFECT takes place in the plane between two semiconductors cooled to near absolute zero (*below*); the atoms on the surface of the semiconductors are depicted as green and blue spheres. When a magnetic field (*red lines*) is applied, electrons in a current (*yellow, traveling into page*) are redistributed so that more electrons are on one side (the right) than the other. This redistribution of electrical charge produces a measurable voltage (the Hall voltage) and conductance perpendicular to the current flow (*top view at left*). The quantum Hall effect refers to stepwise increases in conductance as the magnetic field rises.



field. The magnetic field causes the electrons to drift sideways to the direction of the current flow. As a result, a sideways voltage, or force that pushes electrons, develops. If the magnetic field increases, this voltage also increases, but not linearly; rather it increases in a precise, stepwise fashion. This phenomenon is the quantum Hall effect and is considered the signature of a new, distinct phase of matter.

When the quantum Hall effect was discovered in 1980, physicists recognized that the properties of electrons in this extraordinary state differ fundamentally from those in all other known states of matter. The latest explorations in this field, however, have uncovered a striking relation between the quantum Hall effect and the more familiar phenomenon of superconductivity. Studies of this link have even led to predictions of other new phases of matter, which experiments have recently confirmed.

Although the quantum Hall effect may have no immediate practical value, its study has fostered the development of new concepts and theoretical tools. These tools are likely to have broad implications for physics, in much the same way that the theory of superconductivity has helped advance elementary particle physics and that the study of phase transitions has greatly enhanced the understanding of the early universe.

Explorations of the quantum Hall effect also afford a glimpse into the astonishing ways in which the subatomic world operates; these studies are thereby spurring theorists to formulate a more complete view of the physical universe. Moreover, the principles involved may well prove important in future generations of semiconductor microelectronic devices. As these devices are made ever smaller, they will eventually reach dimensions at which quantum mechanics and the interactions among electrons will become crucial to design considerations.

Discovery of the Quantum Hall Effect

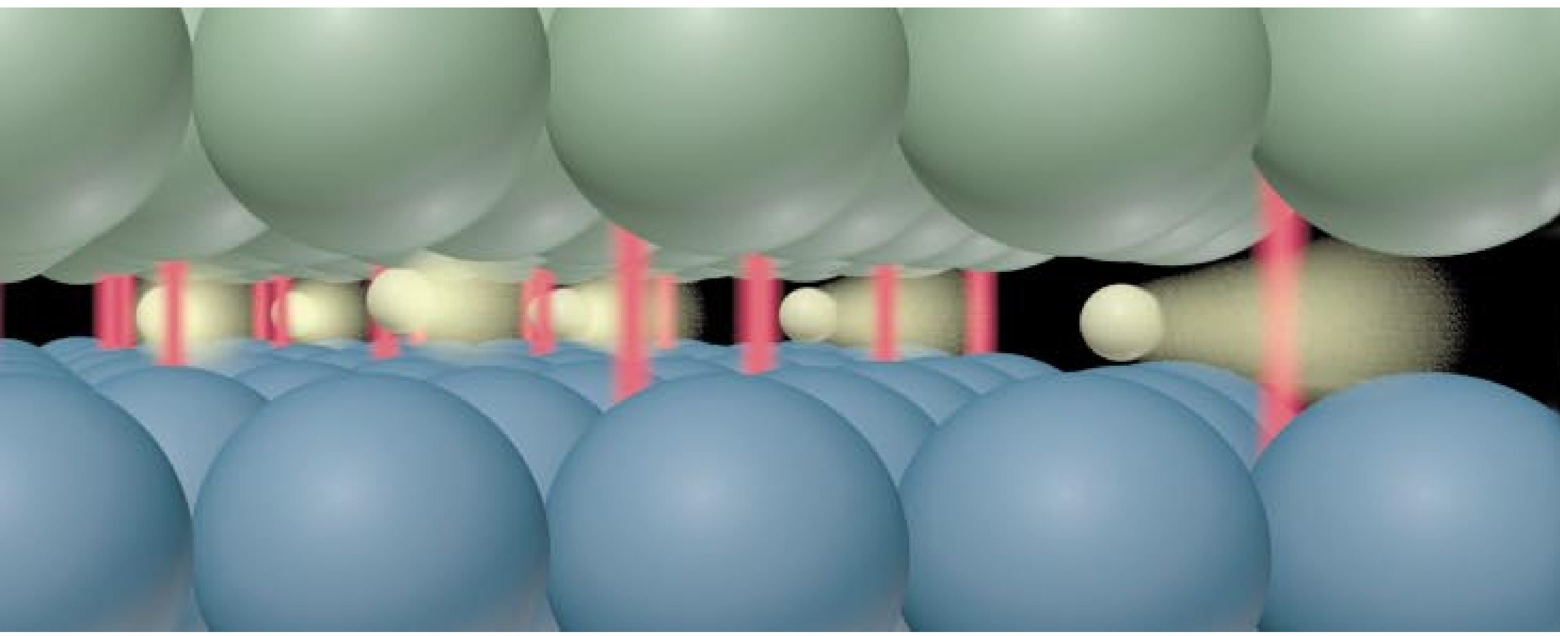
The quantum Hall effect turns out to be an unusual manifestation of a well-known, more general phenomenon in electrical conduction that was discovered by the American physicist Edwin H. Hall in the 19th century. When a voltage is applied between the ends of a wire, a current begins to flow. If the wire is then subjected to a magnetic field, the flowing electrons experience a sideways force. This force redistributes the electrons nonuniformly—more electrons end up on the right side of the wire and fewer on the left.

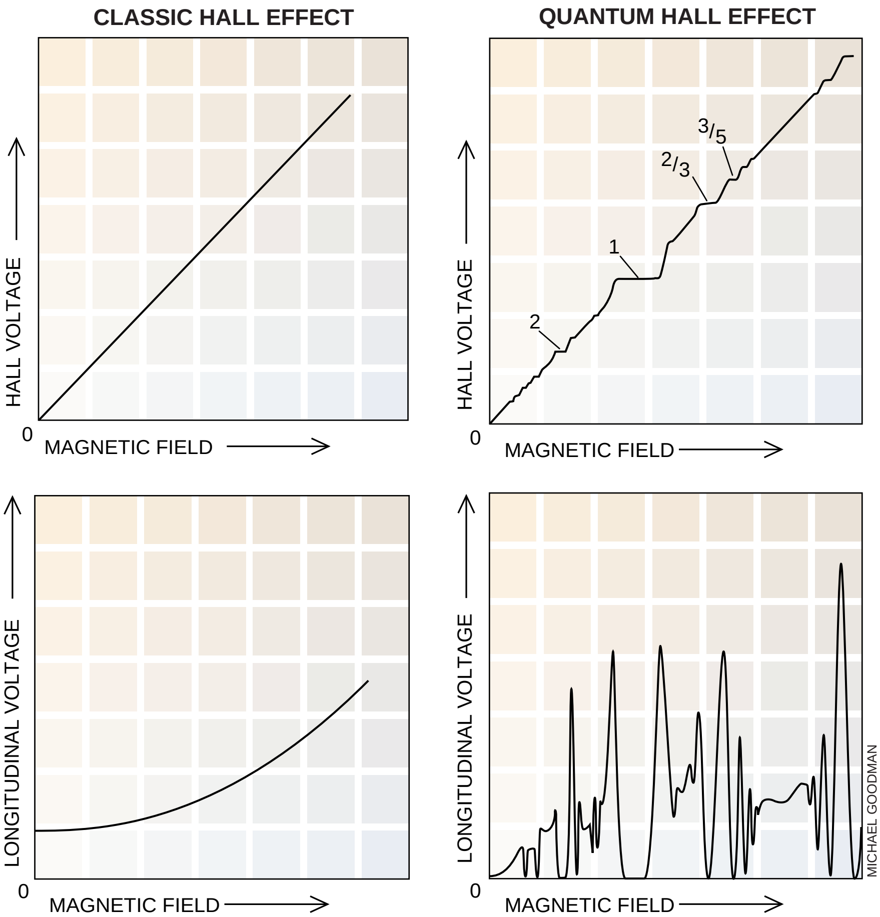
The nonuniform distribution in turn produces an electrical voltage perpendicular to the current direction. To detect this sideways voltage (the Hall voltage), one simply attaches the probes of a standard measuring device, such as a voltmeter, to the sides of the wire. For a fixed magnitude of current, the Hall voltage rises smoothly with the strength of the magnetic field. This phenomenon is now commonly referred to as the classic Hall effect.

In 1980 Klaus von Klitzing, then at the High Magnetic Field Laboratory of the Max Planck Institute in Grenoble, Michael Pepper of the University of Cambridge and Gerhardt Dordda of Siemens Research Laboratory in Munich discovered that under special circumstances, the Hall effect does not obey the usual rules. Thanks to advances in semiconductors, it is possible to trap a collection of electrons between two crystalline semiconductors such that the electrons can move only in a single plane. When the investigators chilled these trapped electrons to within a degree or two Celsius of absolute zero, they found that the Hall voltage did not rise smoothly as the strength of the magnetic field increased.

Instead the Hall voltage rose in steps, with values that did not vary at all over a small range of magnetic-field strengths [*see illustration on next page*]. In addition, the longitudinal voltage—that is, the voltage necessary to maintain the flow of current—nearly vanished when these plateaus in the Hall voltage were reached. In other words, the electrons in flatland became "perfectly conducting." (They are not technically superconducting; superconducting electrons can also expel a magnetic field. These perfectly conducting electrons do not.)

Perhaps more astonishing is that at each plateau, a quantity called the Hall conductance was found to have a special value. The Hall conductance is the ratio between the amount of longitudinal current and the value of the Hall voltage. Von Klitzing and his colleagues concluded that at each plateau the Hall conductance equaled an integer multiple of the quantum of conductance, a unit equal to 1/25,812.8 inverse ohms (conductance is the inverse of resistance). The quantum of conductance is represented as e^2/h. (The e stands for





SIGNATURES of the quantum Hall effect are seen when certain measurements are compared with those relating to the classic (nonquantized) Hall effect. In the classic Hall effect the voltage in the sideways direction (the Hall voltage) varies smoothly with the magnetic field; in the quantum version the voltage plateaus at conductances equal to certain integer and fractional multiples of a fundamental constant (only a few multiples are noted). Parallel to the current, the longitudinal voltage varies smoothly with the magnetic field in the classic effect; in the quantum version the voltage disappears when the Hall voltage plateaus.

the charge of an electron; *h* is Planck's constant, which relates the frequency of a light ray to the smallest amount of energy it can carry.) For his discovery of this "integral quantum Hall effect," von Klitzing won the 1985 Nobel Prize in Physics.

In 1982 Daniel C. Tsui, now at Princeton University, Horst L. Störmer of AT&T Bell Laboratories and Arthur C. Gossard, now at the University of California at Santa Barbara, encountered another unexpected property of the quantum Hall effect. They discovered that the Hall voltage plateaued more often than was originally thought. It leveled off at specific fractional values, such as $^1/_3$, $^2/_5$ and $^3/_7$, of the quantum of conductance. The name of this phenomenon, not surprisingly, is the fractional quantum Hall effect.

So far no experiment has revealed any deviation at all between the measured Hall conductances and the quantized values. They are the same to within at least one part in 10 million (seven decimal places); indirect evidence suggests they are the same to within at least one part in 100 billion. Because of this accuracy, the National Institute of Standards and Technology has adopted the quantum Hall effect as the standard to calibrate resistance-measuring devices.

Magic Filling Factors

Why does the Hall conductance adopt these "magic" values? Investigators spent years trying to solve this puzzle. The answer, as will be seen, lies with the strength of the magnetic field impinging on each electron.

To understand the solution, one needs to know three things about how physicists describe magnetic fields. First, quantum mechanics represents the amount of magnetic field acting on a sample in terms of a unit called the magnetic flux quantum. One way to picture a flux quantum is to imagine it as an arrow. To measure the strength of the magnetic field, one simply counts the number of flux quanta—arrows—poking through a sample in a given area.

Second, an important quantity related to the magnetic-field strength is called the filling factor. This is the number of electrons in a sample divided by the number of magnetic flux quanta penetrating the sample. When the filling factor is one, there is one flux quantum per electron; when the filling factor is $^1/_3$, there are three flux quanta for each electron.

Third, there is a correlation between the quantized values of the Hall conductance and the corresponding filling factors (known as magic filling factors). When the filling factor is 1, the Hall conductance is found to be $1e^2/h$; when the filling factor is $^1/_3$, the Hall conductance is $^1/_3\ e^2/h$, and so on.

Robert B. Laughlin, now at Stanford University, first explained the plateaus in the Hall conductance using separate, idealized mathematical models for the integer and the fractional filling factors. His groundbreaking explanations (and those of others) relied on wave functions: mathematical functions that describe everything there is to know about the state of quantum particles.

Successful as Laughlin's approach was, it left several questions unanswered. It relied on certain simplifications that are difficult to apply to real-world materials, which contain many imperfections. Wave functions are abstract, so Laughlin's explanations were rather tough to picture. His approach did not indicate whether any relation exists among the quantum Hall effect and other kinds of electronic activity in solids. Finally, considering the similarity between the integer and fractional quantum Hall effects, it would seem they should be treated on an equal footing and not separately.

By exploiting a precise mathematical analogy between the quantum Hall effect and superconductivity, we have developed a new way to comprehend the quantum Hall effect. Besides unifying two seemingly disparate phenomena, the analogy enables physicists to apply knowledge of superconductivity to the quantum Hall effect. This approach complements Laughlin's, and it embodies many of the insights gained from his work. But it is distinctly separate, focusing on the real, macroscopic observables of the physical system rather than on the hard-to-visualize, microscopic properties of an ideal system.

The first steps in this new direction

occurred in 1987. Steven M. Girvin and Allan H. MacDonald, now both at Indiana University, recognized that the wave functions used to explain the quantum Hall effect could be viewed as representing the superconducting state of a new, imaginary type of particle, called a composite boson. A similar observation was made somewhat later by Nicholas Read, now at Yale University.

Bosons and Fermions

Bosons are one of two families into which physicists group all particles on the basis of their "statistics," or group behavior. The wave function describing a collection of bosons remains the same when two of the particles exchange places. The other family of particles is the fermions; their wave function changes sign (from positive to negative, or vice versa) when two particles are swapped.

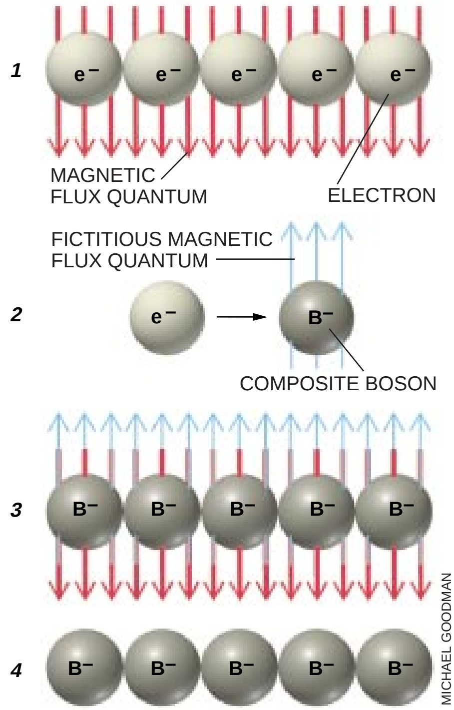
Electrons, protons and neutrons are all fermions. An atom, which contains all three, can also be treated as a single (composite) particle. Whether it is a boson or fermion depends on the net number of its constituents: if this number is odd, the atom is a fermion; if even, the atom is a boson. The isotope called helium 4, for example, contains two electrons, two protons and two neutrons, making it a boson. In contrast, the isotope helium 3 has two electrons, two protons and only one neutron; therefore, it is a fermion.

Bosons and fermions differ in many ways. Most relevant to this article are the rules governing the occupation of quantum-mechanical states. Fermions obey the Pauli exclusion principle, which forbids two fermions from occupying the same state—essentially, they cannot be in the same place at the same time. This rule does not apply to bosons; many bosons can exist in exactly the same state.

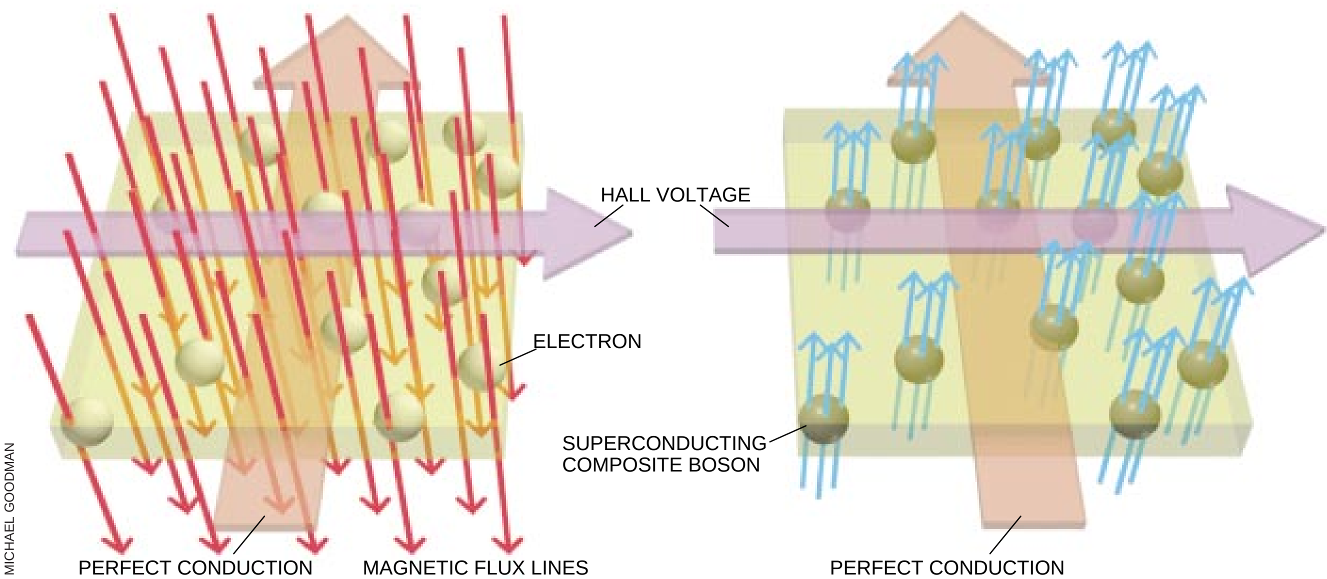
These two fundamentally different properties of fermions and bosons account for many observations in physics. A good example is the dramatic difference between a superconductor and an ordinary metal. Electrical conduction in ordinary metals can be readily understood in terms of the properties of fermions (specifically, electrons); in contrast, superconductivity is a property of bosons.

How can this be, considering that the electric current carriers in all solids are electrons, which are fermions? The answer is that in the superconducting phase, the electrons bypass the rules for fermions by pairing up. Each pair then acts as a boson; all these pairs can condense into the same quantum state to produce superconductivity. In the ordinary metallic state, however, electrons retain their single, fermion identities. As fermions, they exist in different states, as demanded by the Pauli exclusion principle, and fail to superconduct.

The theory that uses composite bosons to explain the quantum Hall effect



COMPOSITE BOSONS can represent the electrons in the quantum Hall effect. For example, at what is called filling factor $^1/_3$, there are three flux quanta (a measure of magnetic-field intensity) for every electron (*1*). The authors envisioned this condition with composite bosons, or charged particles having three (fictitious) magnetic flux quanta (*2*). Orienting the fictitious flux against the real magnetic flux (*3*) eliminates the magnetic field each boson "sees" (*4*), easing the modeling of the quantum Hall effect.



EXPLANATION OF QUANTUM HALL EFFECT with electrons at filling factor $^1/_3$ (*left*) follows once the composite bosons cancel out the external magnetic field (*right, where the canceled field has been left out for clarity*). Cold, charged bosons in the absence of a magnetic field become superconducting, accounting for the perfect conduction in the longitudinal direction. The Hall voltage develops because of induction: the moving fictitious magnetic flux produces a sideways voltage.

Comparing Superconductivity with the Quantum Hall Effect

SUPERCONDUCTIVITY	QUANTUM HALL EFFECT
Paired electrons (Cooper pairs) are the basic carriers of charge	Composite bosons are the basic carriers of charge
Perfect conduction	Perfect conduction in longitudinal direction
Persistence in a magnetic field and in materials with imperfections	Quantized Hall plateaus, whose values persist over a small range of magnetic-field strengths
Electrons exclude a weak magnetic field	Electrons resist change in density at a fixed magnetic field, a property called incompressibility
Flux quantization (a property of superconducting rings, which must enclose an integral number of magnetic flux quanta)	Fractional charge (the quantization of electrical charge in units that are fractions of the electron's charge)

was introduced in 1989 by two of us (Zhang and Kivelson), along with T. Hans Hansson, now at the University of Stockholm. Loosely summarizing, we proposed that electrons moving in two dimensions in a strong magnetic field are mathematically equivalent to a collection of composite bosons in a much weaker magnetic field. Under particular circumstances—namely, when the electron filling factor reaches a magic value (specifically, 1, $^1/_3$ or $^1/_5$)—the magnetic field that composite bosons experience is actually zero. In that case, we argued, composite bosons would, under a broad range of circumstances, become superconducting. We then showed that when composite bosons become superconducting, they give rise to the quantized Hall conductances.

Along with Matthew P. A. Fisher, now at the University of California at Santa Barbara, one of us (Lee) logically extended this theory to account for all the other, more complicated quantized Hall plateaus, such as $^2/_5$ and $^3/_7$. These works formed the basis of the subsequent investigations that the three of us carried out to study the quantum Hall effect under a variety of conditions.

Electrons as Composite Bosons

The theory of composite bosons is based on a mathematical equivalence between electrons moving in two dimensions and a collection of bosons carrying a bundle of fictitious magnetic flux. It turns out that in order for the composite boson to mimic the electron's Fermi statistics, each boson must carry an odd number of fictitious magnetic flux quanta. (A more rigorous justification for this representation appears in *Scientific American*'s area on America Online.)

Perhaps an example best explains the effects of fictitious magnetic flux. Consider one of the magic filling factors at which a plateau in the Hall voltage appears—say, $^1/_3$. This filling factor means that there are three quanta of real magnetic flux per electron. Let's consider each electron not as a fermion but as a composite boson bound to three quanta of fictitious flux. Now point these three flux quanta in the direction opposite to the external magnetic field. The net flux seen by the bosons is the sum of the real and fictitious fluxes. Because we have pointed the fictitious flux so that it cancels the real flux, the boson sees no net magnetic flux whatsoever. At low temperatures, bosons in the absence of a magnetic field are known to superconduct, so we expect the same to be true of cold composite bosons at filling factor $^1/_3$.

Why should the superconductivity of composite bosons imply perfect conduction in the current direction and a quantized Hall conductance in the perpendicular direction? The first part is easy. Because the composite bosons are superconducting, no voltage is needed to sustain the current flow; one thus finds perfect conduction.

The second part is more subtle. Recall that each flowing composite boson carries an odd number of fictitious magnetic flux quanta. Therefore, if the bosons are flowing, fictitious magnetic flux quanta must flow with them. But moving magnetic fluxes (even fictitious ones) generate an electrical voltage perpendicular to the flow (this property is known as Faraday's law of electromagnetic induction). Furthermore, this sideways voltage is proportional to the total amount of fictitious flux flowing through the sample every second. So for filling factor $^1/_3$, the magnetic flux current is three times the electric current. That in turn accounts for the Hall conductance being equal to $^1/_3$ of the quantum of conductance.

From this viewpoint, the only difference between the various magic filling factors—be they 1, $^1/_3$ or $^1/_5$—lies in the number of fictitious magnetic flux quanta each composite boson carries. In addition, the quantized Hall conductances (such as 1, $^1/_3$, $^2/_5$ and so on, multiplied by e^2/h) depend only on the ratio between charge and flux in the composite boson and not on the details of the material in which they are observed.

The model using composite bosons also explains why the Hall conductance remains unchanged even when the filling factor deviates slightly from a magic value. Consider the situation in which the electron filling factor is slightly more than $^1/_3$. In that case, the fictitious flux cancels the real flux only partially, and the composite bosons experience a small net magnetic field. But like a real superconductor, the composite boson superconductor can tolerate a small magnetic field. Consequently, the Hall conductance is unchanged within a finite window around filling factor $^1/_3$.

The analogy between superconductivity and the quantum Hall effect goes much further. For instance, the ability of a superconductor to repel a magnetic field translates into the ability of the electrons displaying the quantum Hall effect to resist any change in the overall area they occupy (the Hall effect electrons are said to be "incompressible"). Other, more obscure aspects of superconductivity also have direct analogues in the quantum Hall effect.

A Map for Flatland

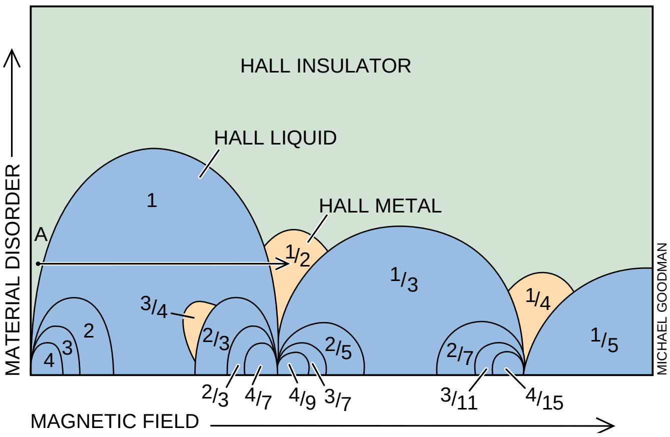
Using the composite boson theory, the three of us studied the quantum Hall effect under a broad range of circumstances. The outcome of that study is represented in a so-called phase diagram. Physicists often use a phase diagram to summarize the behavior of a material under various conditions. For example, at different pressures and temperatures, a collection of water molecules can become a liquid, ice or steam. A diagram illustrating these phases can be constructed to indicate the physical state of the water molecules over a range of pressures and temperatures.

Rather than pressure and temperature, the phase diagram for the electrons in flatland uses as parameters the strength of the magnetic field and the degree of imperfection, or disorder, in the semiconductor crystals that trap the electrons. We obtained such a diagram from the known phase diagram of superconductors. Mapping information from the superconducting phase diagram produced a beautifully nested structure [*see illustration at right*].

The theory of composite bosons in the quantum Hall effect also led to the prediction of an unexpected state, in which the electrons adopt the properties of an insulator and a metal simultaneously. The prediction of such a Hall insulator was confirmed in a recent experiment carried out by Hong-Wen Jiang and Kang-Lung Wang of the University of California at Los Angeles and Scott T. Hannahs of Florida State University. When they increased the degree of imperfection in the semiconductors beyond a certain point, a very large voltage became necessary to sustain the current flow. The need for more voltage increased steadily as the temperature fell toward absolute zero—characteristics of an insulator. In contrast, the Hall voltage remained independent of temperature and increased with the strength of the magnetic field—characteristics of a metal.

Experiments by Jiang, Tsui, Störmer and Loren N. Pfeiffer and Ken W. West of AT&T Bell Laboratories and others have revealed another surprise—this time, near filling factor $^1/_2$. In that case, physicists discovered that electrons acted largely as though they were in an ordinary metal and not in a magnetic field. Among other features, the Hall conductance was not quantized but depended linearly on the magnetic field.

An intriguing explanation for this so-called Hall metal relies on the idea that an electron can be viewed as a composite fermion. A composite fermion resembles a composite boson, except that it carries an even number of fictitious magnetic flux quanta and, as a result, obeys Fermi statistics. Several researchers have put forward such ideas, based in part on a concept first introduced by Jainendra K. Jain of the State University of New York at Stony Brook. These workers include Read, Bertrand I. Halperin of Harvard University, Patrick A. Lee of the Massachusetts Institute of Technology and, independently, Vadim Kalmeyer, formerly at the IBM Almaden Research Center, and one of us (Zhang).



NEW STATES OF MATTER for electrons in flatland are shown in a phase diagram. At a given magnetic field and level of disorder (*point A*), the electrons act as a "Hall insulator" (*green*), which has both insulating and metallic characteristics. At higher magnetic fields, the electrons turn into a "Hall liquid" (*blue*)—that is, they show the quantum Hall effect—and then become a "Hall metal" (*beige*). The numbers signify the integer and fractional values of the quantized Hall conductance.

The virtue of both the composite bosons and composite fermions is that seemingly strange behavior of electrons in flatland can be related to the familiar behavior of composite particles. A question often raised is whether these composite particles are real or are, like quarks in high-energy physics, useful constructs that cannot be isolated and studied individually. This debate has sparked considerable research, but definitive results have yet to be garnered.

Sixteen years after its discovery, the quantum Hall effect remains one of the most exciting areas of research in condensed-matter physics. The rich variety of phenomena has provided a testing ground for many theoretical ideas. A global picture has emerged that unifies the understanding of these phenomena and others in condensed-matter systems. Yet despite the progress, critical issues remain unsolved. For example, the Hall insulator and the Hall metal are still far from completely understood. How other properties of electrons, such as their spin, fit into the picture is also not fully known.

In an earlier *Scientific American* article on this subject, in 1986, Halperin remarked that "the true importance of the quantized Hall effect does not lie in any… applications, but rather in the new insight physicists have gained into the peculiar properties of systems of electrons in strong magnetic fields and into the hidden regularities implied by the mathematical laws of quantum mechanics. Nature may well hold in store other surprising states of matter that none of us yet imagine." Ten years later physicists have found some of those states, and we look forward to discovering more.

The Authors

STEVEN KIVELSON, DUNG-HAI LEE and SHOU-CHENG ZHANG collaborated in helping to draw the connection between superconductivity and the quantum Hall effect. Kivelson, who received his doctorate from Harvard University, held positions at several institutions before taking on his current professorship of physics at the University of California, Los Angeles. Lee earned his Ph.D. from the Massachusetts Institute of Technology and worked at the IBM Thomas J. Watson Research Center before joining the faculty of the University of California, Berkeley. Zhang, currently an associate professor at Stanford University, received his Ph.D. from the State University of New York at Stony Brook and previously held positions at the University of California, Santa Barbara, and at the IBM Almaden Research Center.

Further Reading

QUANTIZED HALL EFFECT. Bertrand I. Halperin in *Scientific American,* Vol. 254, No. 4, pages 52–60; April 1986.

ANYONS. Frank Wilczek in *Scientific American,* Vol. 264, No. 5, pages 58–65; May 1991.

THE CHERN-SIMONS-LANDAU-GINZBURG THEORY OF THE FRACTIONAL QUANTUM HALL EFFECT. Shou-Cheng Zhang in *International Journal of Modern Physics B,* Vol. 6, No. 1, pages 25–58; January 1992.

GLOBAL PHASE DIAGRAM IN THE QUANTUM HALL EFFECT. S. Kivelson, D.-H. Lee and S.-C. Zhang in *Physical Review B,* Vol. 46, No. 4, pages 2223–2238; July 15, 1992.

Caribbean Mangrove Swamps

Despite their ubiquity and prominent position between land and sea, these tropical ecosystems still hold countless surprises for researchers

by Klaus Rützler and Ilka C. Feller

One perceives a forest of jagged, gnarled trees protruding from the surface of the sea, roots anchored in deep, black, foul-smelling mud, verdant crowns arching toward a blazing sun, the whole mass buzzing with insects. These are the first impressions a visitor develops when approaching one of the most common sights on tropical shores—a mangrove swamp. Here is where land and sea intertwine, where the line dividing ocean and continent blurs. In this setting the marine biologist and forest ecologist both must work at the extreme reaches of their disciplines.

Naturalists have long struggled to define, in proper ecological terms, the environment of a mangrove swamp. Is it an extreme form of coral reef or a flooded coastal forest? Compared with the tropical timberlands of some continental interiors (which can house as many as 100 species of tree on a single hectare), a mangrove forest appears puny, monotonous and depauperate. Even the relatively rich Indo-Pacific coasts boast only some 40 mangrove species along their entire length. In the Western Hemisphere only eight or so mangrove species can be found. And of this small set just three kinds of mangrove tree are truly common.

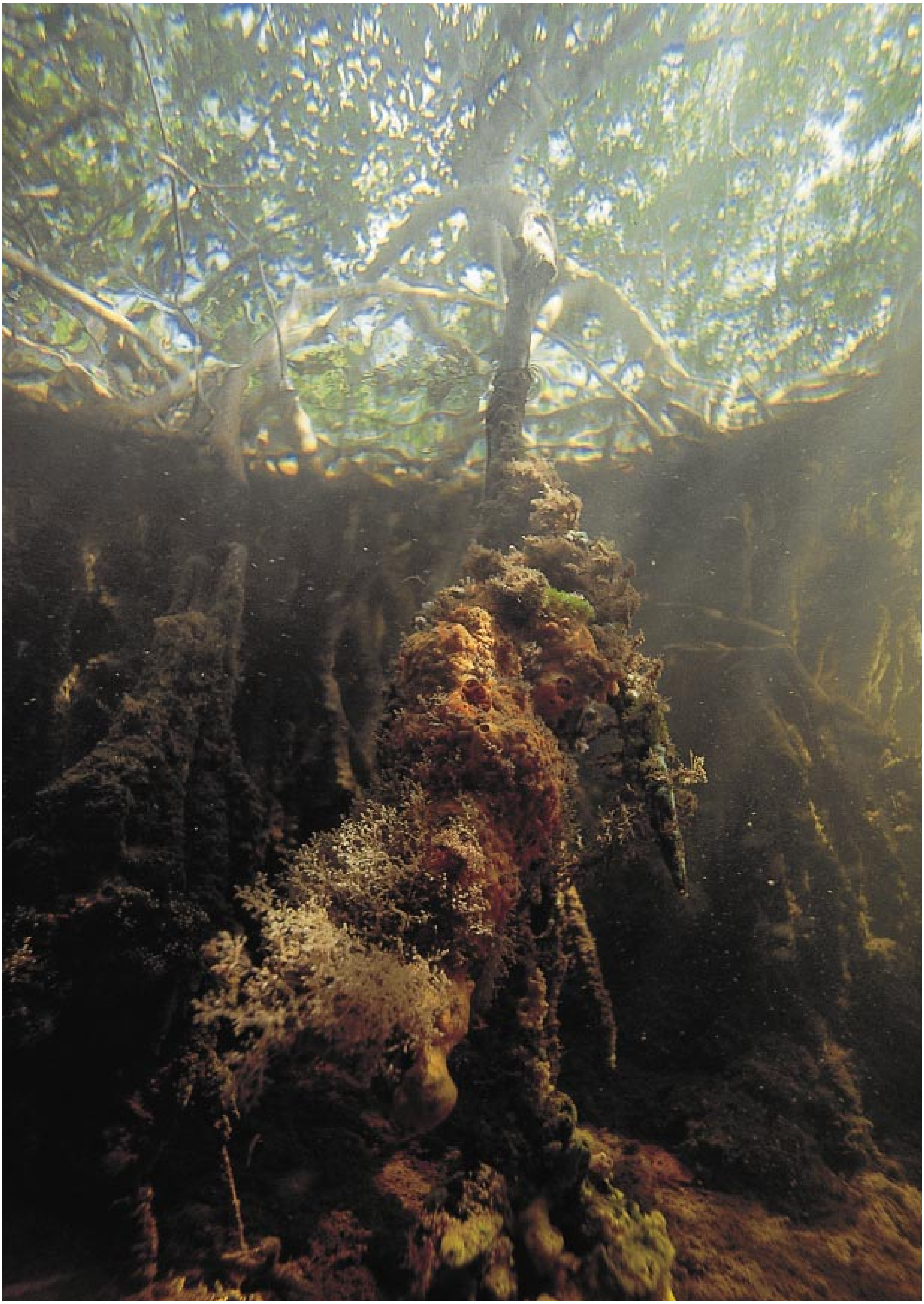
The word "mangrove" itself is somewhat misleading. It is not a formal taxonomic term; rather, it applies to those vascular plants that share a set of physiological mechanisms for living in shallow seawater. For example, mangrove trees are able either to exclude or to excrete salt from their tissues. These plants can also sprout aerial roots that permit the exchange of gas for aerobic respiration. This adaptation allows the trees to survive despite being firmly rooted in oxygen-deprived intertidal soils. Although a salty habitat is not strictly required for growth, mangrove communities develop only near the sea because they cannot compete successfully with freshwater flora.

Mangroves are also limited by a need for rather balmy conditions. As is the case with corals, these trees cannot survive in places where the average water temperature falls below about 23 degrees Celsius (73 degrees Fahrenheit). This requirement causes bands of mangrove and coral to grow largely in parallel swaths along the world's tropical coastlines. But deviations from this simple pattern occur commonly. For instance, coasts that are continuously subjected to large influxes of sediment (such as the shores of western Africa) or areas where deep, cold water rises to the surface (such as the coastal waters off eastern Venezuela) lack coral reefs but can support extensive mangrove swamps. Conversely, some coral islands in the central Pacific lack an accompanying fringe of mangroves, apparently because the floating propagules that serve as seeds for these trees cannot reach such remote isles.

Mangrove swamps typically fall within one of two broad categories of classification: mainland or oceanic island. The former group includes those communities that border continental coasts and are thus permanently sandwiched between salty ocean water and freshwater carried by streams from the interior. Hence, mainland mangroves must usually cope with a pronounced variation in salinity across their width. This situation is quite distinct from that of ocean-

ROBERTO OSTI

TROPICAL SHORES around the world are commonly lined with mangrove swamps (*above*). Salt-tolerant mangrove trees thrive along coasts where the average temperature remains sufficiently warm. Mangroves on the barrier reef bordering Belize (*inset*) are central to the authors' long-term research. These swamps (*opposite page*) host rich communities of plants and animals within the shady tree canopy and around the permanently submerged mangrove roots. The swamps also support a distinctive group of creatures that have adapted to life in the thin zone between high and low tide levels.



CHIP CLARK *Smithsonian Institution*

JIMMY SMITH *Islands from the Sky*

CARRIE BOW ISLAND (*foreground*) and Twin Cays (*center*) have served the authors as a permanent base and natural laboratory. Straight, dashed lines on the bottom of the lagoon mark scars in the sea-grass bed where oil-exploration crews used explosives to make seismic surveys during the 1960s.

ic-island mangroves, which form on shallow banks or in lagoonal areas well separated from the mainland. They are normally less affected by freshwater carried laterally than by the intermittent shifts in salinity that result from intense evaporation or from frequent tropical downpours.

Although interest in mangrove biology reaches back in history at least as far as Alexander the Great's expeditions to the Arabian Sea, scientific knowledge of this intriguing ecosystem is still rudimentary, and key questions remain largely unanswered. Are mangrove communities as rich and productive as other tropical environments? Is their role in protecting juvenile fish indeed as important to commercial fisheries as many people have speculated? Do mangrove swamps serve to protect coastlines from erosion? Although researchers have made detailed observations of many different mangrove swamps around the world, a huge gap exists in the understanding of how the different components of such intricate natural systems work together.

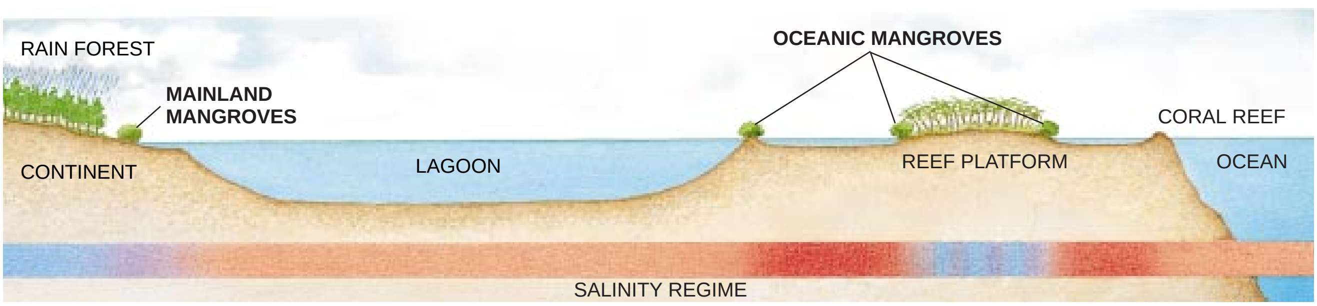
A Cay is Key

Concerns about the limited understanding gleaned from isolated studies of mangrove swamps motivated us to mount a long-term research campaign at one place. We chose to examine the biodiversity and ecology of a locale that was relatively accessible—the spectacular barrier reef off the coast of Belize. We are fortunate to have been able to conduct our fieldwork there from a permanent station situated on a tiny coral island about 10 miles offshore. One of us (Rützler) discovered this site on a memorable morning in February 1972. It was something of a serendipitous find during an excursion with a colleague, Arnfried Antonius.

We had chartered a small boat out of Belize City, some 50 miles to the north, and were looking for a passage through the shoals—one through which we had traveled several times before. But our boat's crew was unfamiliar with the local waters and misguided us in our search for the break in the reef. As we motored seaward, we could hear the splash of waves breaking on the shallow coral, but neither of the two islands in front of us matched what we remembered from our previous visits. To our astonishment we noticed several buildings on the smaller island ahead and decided to make a closer examination of this curious display of civilization within a rather remote wilderness. A few moments later we tied up our vessel on the island's concrete dock and walked toward the largest house. No inhabitants greeted us (other than a few mildly disturbed pelicans), but hanging above the main gate was a sign: "Welcome to Carrie Bow Island."

Little did we know then that this speck of sand containing three cottages and two outhouses was the property of Henry T. A. Bowman, a citrus planter with a passion for the sea who had acquired the island in 1943 in order to build a summer home for his wife, Carrie. Nor could we then imagine that within eight years the Bowman family's appreciation and support for the natural sciences would transform this remote vacation retreat into a permanent laboratory—one that has since hosted more than 70 scientists from 40 institutions, enabling hundreds of scholarly studies of the surrounding reefs to be conducted.

Having Carrie Bow Island available as a convenient base, we decided to center our examination of mangroves on the nearby Twin Cays, a largely untouched mangrove range covering more than a square kilometer of a shallow lagoon. More than 20 other researchers from the Smithsonian Institution's National Museum of Natural History (and at least as many colleagues from collaborating



ROBERTO OSTI

BELIZE'S BARRIER REEF is separated from the coast by a wide, shallow lagoon. Mainland mangrove communities (*left*) endure a permanent lateral gradient in the saltiness of ambient waters, where freshwater (*blue in band*) gives way to salty conditions (*red*). Mangroves situated offshore (*right*) more often cope with erratic fluctuations in salinity.

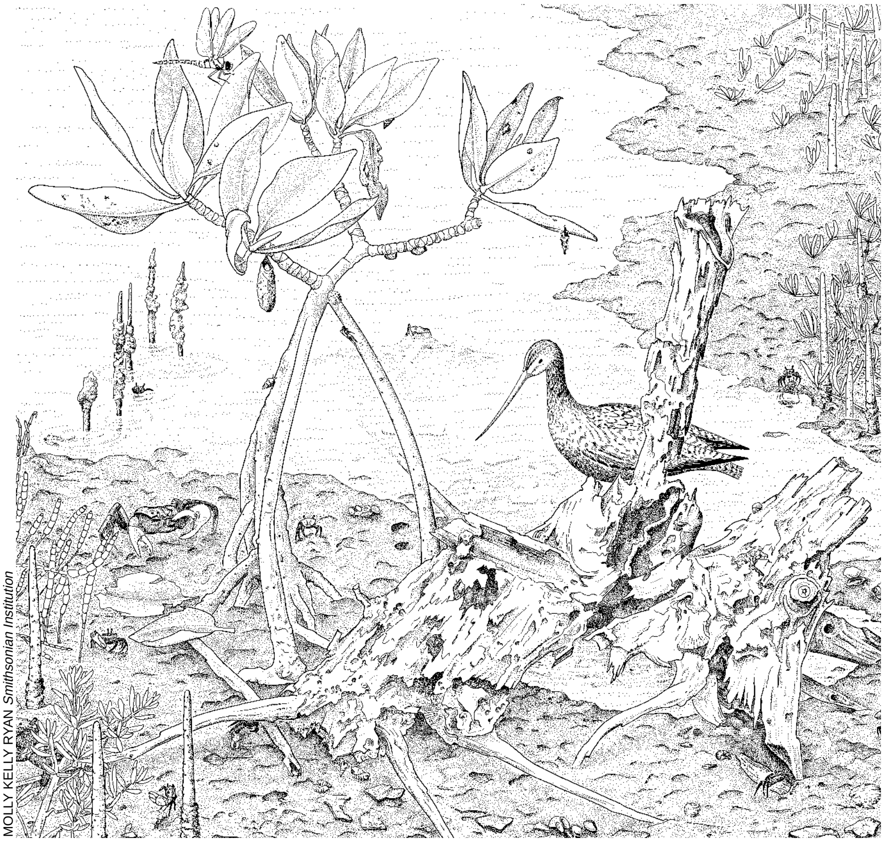
institutions in the Americas, Europe and Australia) also conduct regular research on the mangrove communities there.

A Natural Landscaping

The mangroves of Twin Cays belong to the oceanic-island type. The distribution of such scattered mangrove islands among the patches of reef within the wide lagoon suggested to us that the first trees at Twin Cays may have settled on isolated coral clusters. But work by our geologist colleague Ian G. Macintyre disproved this surmise. Several "vibracores" (geological samples obtained by pushing a vibrating pipe into the unconsolidated sediment bottom) indicated that the mangroves at Twin Cays did not begin growing on coral. Rather the community established itself some 7,000 years ago on what was then just elevated ground. This mangrove swamp has since built a foundation of seven meters of peaty soil as it reacted to rising sea level.

The topography of Twin Cays reflects several thousand years of natural history and provides a testament to the power of the countless storms and hurricanes that have buffeted this tiny forest. For instance, clear evidence of at least one dramatic event in the past remains in the shallow channel that bisects the island and then splits, with one branch that has no outlet. At the bends in the dead-end extension lie deep cuts that have not yet filled with peat or sediment. These submerged excavations indicate that strong currents must have flowed freely through the passage before an unrecorded tempest deposited enough sediment to dam one end of the channel.

The periphery of Twin Cays and its canals is bordered by tall red mangroves (*Rhizophora mangle*) that extend stilt-like roots into the deeper water, beyond the peat bank that supports the trees. Toward the interior, first black mangrove (*Avicennia germinans*) and then white mangrove (*Laguncularia racemosa*) mark zones of shallow floodwater and extended mudflats that develop as the tide moves in and out. Because evaporation under the intense sun can rapidly remove much of the standing water, the brine that remains in pools often contains a high concentration of salt. Karen L. McKee of Louisiana State University has found that hypersaline conditions in the upper reaches of the intertidal zone favor black mangrove seedlings because that species has the greatest tolerance for salt. White mangrove seedlings cannot survive the salinity and the periodic flooding that sweeps this region, and hence these trees are restricted to higher ground. Red mangrove dominates the lower reaches of the intertidal zone not only because it stands on stilt roots but also because its seedlings can better survive the rigors of the fringing environment (hardships that include a dearth of nutrients and an abundance of possible predators).



MOLLY KELLY RYAN *Smithsonian Institution*

SANDY CANUPP

SUPRATIDAL RESIDENTS of a mangrove swamp include a variety of plants, arthropods, snails and insects, and it is no surprise that the marbled godwit (*bird shown in drawing*) has little trouble finding food.

The interior of Twin Cays is lined with numerous shallow ponds and mudflats. Some of these areas are now without growing trees; some places contain eroded stumps of large trees that must have towered there in the past. Other flats are covered by dwarf forms of red mangrove. These tiny trees are barely one meter tall, yet our studies indicate that they may be several decades old. Initially we assumed that the physiological stresses associated with the increased salinity and high temperature of the mudflats accounted for the slow growth of these bonsai-like mangrove trees. But

ILKA C. FELLER *Smithsonian Institution*

CHIP CLARK *Smithsonian Institution*

INTERTIDAL MANGROVE stilt roots are ubiquitously coated by a mixture of algal species called bostrychietum. Such surfaces also host many invertebrates such as the mangrove oyster (*Isognomon*) and mangrove-tree crab (*Aratus*).

by direct experimentation, one of us (Feller) proved that this stunted growth on Twin Cays and similar Belizean mangrove islands resulted from the lack of a single critical nutrient, phosphorus.

Life from Top to Bottom

Naturalists have traditionally placed a great deal of emphasis on the lateral transition of mainland mangrove forests, from the near-oceanic realm of the coast, through the estuarine environment of river mouths, and upstream into regions of freshwater—the domain of the rain forest. Such horizontal variation is not particularly pronounced on Twin Cays. But there is on the island a well-developed vertical stratification that encompasses the forest canopy, the intertidal region and a reeflike zone below even the slackest water. A substantial part of our research has sought to examine the flora and fauna of these three distinctive environments.

Living within the upper forest levels at Twin Cays are countless insects, lizards, snakes and birds. Although the insects are the most abundant group, the ferocity of the sun and lack of freshwater create a harsh environment for them. Only a few types are active during the day. Most insect species avoid the sun's rays by feeding at night or by living entirely within plants. Consequently, the traditional methods for collecting insects—trapping, baiting, or fogging vegetation with pesticides—yield few specimens. But by dissecting plant parts and by collecting larvae and rearing them, we have found that the mangrove insect fauna is much more diverse and ecologically important than previously suspected.

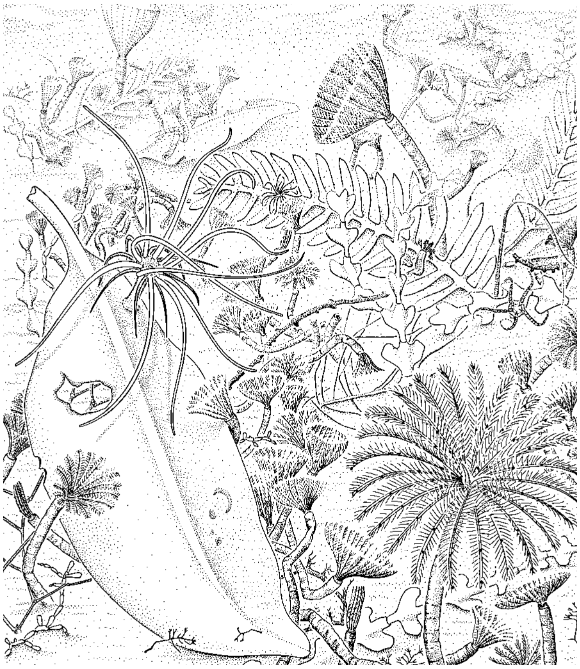
For example, the live branches of red mangrove host several species of specialized wood-boring moths and beetles. Larval stages of these insects feed internally within twigs, creating hollow cylinders of deadwood. After these primary excavators emerge, some 70 other species of ants, spiders, mites, moths, roaches, termites and scorpions use the hollowed twigs for food and as sites in which to hunt, nest and take refuge from the burning sun.

As diverse and interesting as they are, the animal communities suspended above the high watermark are in many ways similar to those found in other tropical woodlands. But the deeper forest environments of tangled red mangrove stilt roots, black mangrove pneumatophores, peat banks and mudflats host inhabitants that are unique to mangrove swamps.

Although the normal tidal range at Twin Cays is only about 20 centimeters, these mangrove forests have distinctive intertidal communities occupying the dank stratum between the high and low tide levels. Aerial roots are typically covered with a combination of red algae that is especially adapted to retaining water when the tide withdraws. Such hard surfaces also support barnacles, oysters and crabs. A rather unusual inhabitant of the intertidal zone is a small fish called the mangrove rivulus (*Rivulus marmoratus*). William P. Davis of the Environmental Research Laboratory of the Environmental Protection Agency and D. Scott Taylor of the Mosquito Control District in Brevard County, Florida, along with several of their colleagues, found that Twin Cays harbors this hermaphroditic species—one of nature's rare examples of a vertebrate that can clone itself.

By far the richest and most densely populated habitat in the mangrove swamp surrounds the subtidal area of the red mangrove stilt roots. Successful colonizers of this space include various species of algae, sponges and anemones. These organisms form biological coatings that cover the stilt roots and offer food and refuge to a variety of fauna, such as mangrove oysters and crabs. Interestingly, our experiments have demonstrated that the roots do not especially attract the "fouling" organisms. After a week of exposure to these waters, any nontoxic material—different woods, plastic or glass—becomes similarly covered, first with a mucous, microbial coating and then by a variety of algae and invertebrates. Mangrove roots seem to benefit from this process: Aaron M. Ellison of Mount Holyoke College and Elizabeth J. Farnsworth of Harvard University have shown that the fouling community protects the trees, at least partially, from the attack of root-boring animals.

The sedimentary bottom of the subtidal swamp is covered with thick stands of sea grass, particularly in well-sunlit channels. In some places on the bottom, algae and jellyfish thrive. The sediment consists of a mixture of sand, mud and detritus that is constantly stirred by the moving water and by the actions of various organisms. Such large bottom feeders as manatees often plow the channel, as do speeding motorboats. Less obvious but much more important, the sediment is continuously being turned over by burrowing animals such as polychaete worms and crustaceans. These creatures act much as earthworms and moles do on land. Peter Dworschak and Jörg Ott of the University of Vienna have shown that some crustaceans can dig and maintain complex burrows of branching tunnels that extend near-



MOLLY KELLY RYAN *Smithsonian Institution*

KLAUS RÜTZLER

SUBTIDAL COMMUNITIES include sabellid worms living among the great quantities of decaying leaf litter (*left*) and the Caribbean fire-sponge (*Tedania ignis*), a beautiful but toxic species (*above*) that can cause severe skin inflammation.

ly two meters into the muddy bottom.

Examinations of algae, aquatic invertebrates and insects on Twin Cays have revealed an astonishing number of new species. Even for relatively well studied groups, such as crustaceans, some 10 percent of the species found there have proved to be novel. We estimate that perhaps as many as 20 to 30 percent of the microbes, algae, sponges and worms living on Twin Cays may also be among as yet undiscovered species.

Biodiversity at Risk

The richness of life contained within the mangrove forest at Twin Cays raises immediate concern for the risk to biodiversity that develops as such delicate environments are lost to human hands. After mangrove forests are cut, it may be difficult or even impossible for them to recover because irreversible changes in the fundamental structure of the ecosystem ensue once the trees are destroyed.

We were able to observe such unfortunate consequences on a black mangrove plot that was illegally clear-cut by fishermen on the western side of Twin Cays. The remaining barren tidal flat was rapidly overrun by saltwort (a small shrub that can tolerate high salinity), reducing the space available for new black mangrove trees, which reproduce more slowly. Where the natural red mangrove fringe was disturbed, currents driven by wind and tides rapidly eroded the peat-rich soil and left a bottom surface on which the seedling propagules had difficulty anchoring.

Some of our field experiments show another impediment that prevents new mangrove trees from establishing themselves after a forest is downed. We found that mangrove seedlings take hold and grow much better in the shade than in the open, indicating that the natural repair of damaged swamps may prove too slow to keep up with the erosion of denuded land.

Such observations demonstrate that mangrove forests indeed constitute extremely delicate natural systems. Although the government of Belize has enacted laws protecting its mangrove species, these trees have nonetheless proved vulnerable because they are situated so directly in the path of commercial development. As on other coasts around the world, people often cut trees and fill low-lying areas to construct housing and industrial facilities.

Because mangrove swamps destroyed in this way would have otherwise supported countless varieties of animals, including juvenile deepwater fish, it seems clear that these losses will increasingly threaten the natural wealth of the ocean, both near and far from shore. But mounting concern about environmental degradation within the world's developing nations has brought renewed scientific attention to mangrove forests. Although much work remains to be done, the research conducted at Twin Cays is slowly but steadily building a body of critical long-term observations. This knowledge should aid future attempts to predict the fate of the earth's fragile tropical shorelines and to foster legislation that might ultimately serve to protect the fascinating mangrove communities found there.

The Authors

KLAUS RÜTZLER and ILKA C. FELLER study mangroves together, but from complementary points of view. Rützler is a marine biologist, Feller a forest ecologist. Rützler began his research in submarine caves of the Croatian Adriatic. In 1963 he earned a Ph.D. at the University of Vienna and in 1965 moved to the Smithsonian Institution's National Museum of Natural History, where he now directs the museum's research program on Caribbean coral reef ecosystems. Feller earned a B.A. as a botanist and scientific illustrator. In the 1980s she worked at the Smithsonian's natural history museum. In 1993, after completing a Ph.D. at Georgetown University, she shifted to the Smithsonian Environmental Research Center in Edgewater, Md., where she is currently a postdoctoral fellow.

Further Reading

THE ECOLOGY OF MANGROVES. A. E. Lugo and S. C. Snedaker in *Annual Review of Ecology and Systematics,* Vol. 5, pages 39–64; 1974.

MANGROVE SWAMP COMMUNITIES. K. Rützler and I. C. Feller in *Oceanus* (Woods Hole Oceanographic Institution), Vol. 30, No. 4, pages 16–24; Winter 1987/1988.

THE BOTANY OF MANGROVES. P. B. Tomlinson. Cambridge University Press, 1995.

Vital Data

by Tim Beardsley, *staff writer*

Later this year, if a biotechnology company called Myriad Genetics has its way, thousands of healthy women in the U.S. will hear doubly bad news. First, a close relative—perhaps a sister—will announce that she has breast cancer. Second, the patient's physician thinks this particular cancer has probably been caused by a mutation that the healthy relative has an even chance of also carrying. The patient has been advised to suggest to all her female relatives that they be tested for the mutation. Women who share it may be urged to consider a prophylactic double mastectomy, as they, too, are likely to develop breast cancer. How likely? Hard to say—the mutations have not yet been thoroughly studied—but the likelihood could be as much as 85 percent.

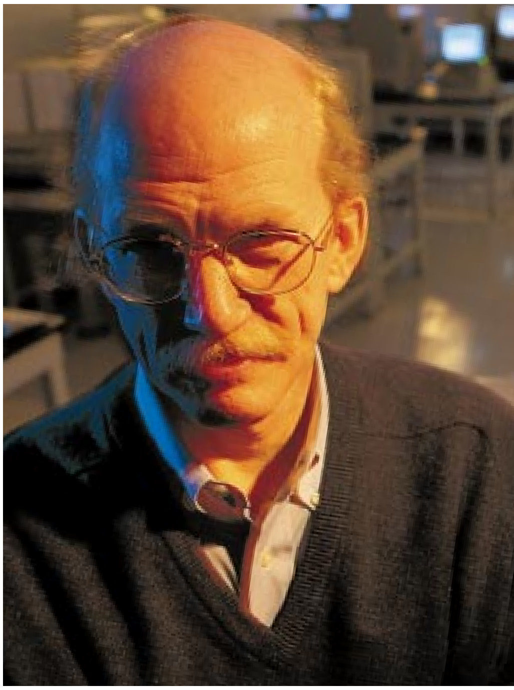
Most people have yet to confront the dilemmas posed by tests for genes associated with serious diseases. But *BRCA1,* the breast cancer gene that Myriad proposes to start testing on a large scale this year, is merely one of what may become dozens that physicians will soon routinely order checked for mutations. The tests are all, to varying degrees, spin-offs of the Human Genome Project, an audacious 15-year, $3-billion federal effort to analyze the human genetic heritage in its ultimate molecular detail.

Although the project was formally launched only five and a half years ago, progress is outstripping expectations. Francis S. Collins, director of the National Center for Human Genome Research (NCHGR) at the National Institutes of Health, thinks the job might be done "maybe even two years early"—by 2003. Technologies developed in response to the effort have already quadrupled the rate of discovery of human disease genes, according to NCHGR's estimate. A medically significant wisp of DNA is now characterized almost every week, and the pace will soon increase again, as the final assault on the genome gets under way. Eventually, a new gene will be sequenced—that is, have the order of its chemical subunits determined—every hour.

The project was launched because it seemed to promise the best hope for ultimately defeating not only diseases long known to be inherited but also others that have a more subtle link with genes, including cancer. Yet the road to that genetic nirvana will be a long one. Early results from one of the most eagerly awaited developments, gene therapy, have been disappointing. And at least one genome biologist, Daniel W. Drell of the Department of Energy, foresees a "vicious legal evolution" of gene-based patent claims, as corporations frantically seek to lock up sequences of DNA with commercial value. The commercialization of humankind's common endowment has prompted protests from activists who see it as an affront to natural dignity. Meanwhile one prominent geneticist has warned that gene-based therapies will be too expensive for widespread use.

Abuses of the new genome science, on the other hand, have arrived already. Compelling evidence demonstrates that children, whom genetic counselors agree should not be tested for certain traits, in fact are being tested in substantial numbers, to their likely detriment. Conversely, because people at risk for a genetic condition are often turned down for either health insurance, life insurance or employment, patients are now declining genetic testing for themselves or their children, even when it would be medically valuable. Others seek testing under false names.

Leaders of the genome project have acknowledged from the start

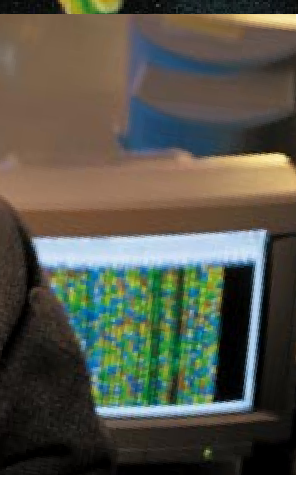


GENOME SEQUENCING CENTER (*top*) at the Washington University School of Medicine, headed by Robert H. Waterston (*bottom*), should begin large-scale sequencing of human chromosomes (*inset, right*) this year.

The Human Genome Project is producing a plethora of information that will illuminate our hidden susceptibilities to disease. The effort could transform medical science. But new dangers are arriving, too

R. JONATHAN REHG

PHILIPPE HURLIN/GLMR *Gamma Liaison*



R. JONATHAN REHG

that human genetics can be used to harm as well as help. They have, indeed, made a point of devoting millions of dollars to studying ethical, legal and social questions. Yet technical gains are close to outrunning the attempts of professional societies and government regulators to guide the technology's use.

Full-Speed Sequencing Ahead

In 1990 planners estimated that finding the sequence of the chemical subunits, or bases, in the whole genome would take several tens of thousands of technician-years. (Writing down the sequence, using one character for each base, would take 390,000 pages of *Scientific American*—with no pictures.) Moreover, sequencing, which will reveal the possible functions and locations of the estimated 100,000 human genes along the chromosomes, could not begin right away. Preliminary prospecting was necessary to describe the lay of heredity's landscape.

Researchers needed a genetic map, essentially a diagram describing how thousands of known marker sequences in the chromosomes separate and recombine between human generations. They also knew they could not get far without physical maps, which show the order along a chromosome of recognizable sequence tagged sites. Using a genetic map, a researcher can compare a given condition's pattern of inheritance with that of marker sequences. That makes it possible to get a quick fix on where a gene that causes the condition might be. Computers can then pin the reams of data emerging from sequencing machines onto the skeleton provided by tagged sites on a physical map. The two kinds of map together make it possible to find quickly genes associated with illness.

Even though Congress has provided less than the $200 million a year the genome project sought, fruitful international collaborations have speeded progress. A good genetic map covering the entire genome was complete by late 1994, and high-quality physical maps now cover 95 percent of the genome. A physical map with tagged sites spaced every 100,000 bases—the ideal—is expected to be complete this year. So the time is ripe for a start on full-scale sequencing.

Efforts to sequence the genomes of other organisms have shown how the process can be speeded and the costs lowered. Several innovative sequencing techniques are now being developed. Collins maintains, however, that radically new approaches will not be necessary to meet the 2005 deadline for a human sequence that is 99.9 percent accurate. Gains in efficiency in the past two years have convinced him that current technology is essentially up to the job.

The Department of Energy, which supports a large part of the U.S. genome project, has several pilot-scale sequencing operations under way. The first large sequencing program was inaugurated at Britain's Sanger Center at the end of last year, funded by the Wellcome Trust. Large-scale sequencing in the U.S. is expected to begin this spring. One likely site is Robert H. Waterston's sequencing center at Washington University.

The oncoming tidal wave of genetic data has not yet affected most people. That will change. Workers are now routinely isolating genetic mutations associated with such widespread illnesses as cancer, Alzheimer's disease and some types of cardiovascular disease. Devising tests for mutations in a known gene has become a comparatively straightforward matter. Genzyme, a biotechnology company in Cambridge, Mass., announced this past fall a diagnostic technology that can simultaneously analyze DNA from 500 patients for the presence of 106 different mutations on seven genes.

When enough is known about the effects of mutations, test results can be a medical boon: they can indicate how likely a person is to develop illnesses and perhaps suggest life-enhancing medical surveillance or therapy. But learning about the effects of mutations requires lengthy study. And genetic data can cause immediate and life-sapping harm. In particular, it can precipitate detrimental psychological changes, and it can open the door to discrimination.

In the past, genetic discrimination has been largely confined to members of families afflicted with rare conditions showing a clear pattern of inheritance. For example, members of families with Huntington's disease, a fatal neurodegenerative disorder that develops in middle age, have long found it difficult or impossible to obtain health insurance. So far "a few hundreds" of people who are at risk of future illness because of their genetic makeup are known to have lost jobs or insurance, Collins states. Most suffered because a family member had been diagnosed with a condition long known to have a genetic basis. But as the number of genetic tests grows, Collins predicts, "we are going to see it happen on a larger scale, since we're all at risk for something."

What's Your Genotype?

Testing for cancer-associated mutations, for example, is at present carried out only in research studies at large medical centers, because interpreting the results is fraught with uncertainty. But some grim facts are clear. In families with hereditary breast cancer—which accounts for less than 10 percent of all cases—mutations in the *BRCA1* gene confer an 85 percent lifetime risk of the disease, as well as a 45 percent chance of ovarian cancer. Some women in such families who have learned that they carry a mutated *BRCA1* have elected to undergo prophylactic mastectomy and oophorectomy (removal of the ovaries)—a procedure that may reduce but does not eliminate the risk of cancer.

More uncertainty arises, however, with women who have a mutated *BRCA1* gene but do not have a family history of breast cancer. For them, the danger is uncertain, but it may be smaller. Nor is it known whether the danger is different for members of different ethnic groups (although Ashkenazi Jews are more likely than others to carry one specific mutation in *BRCA1,* and a gene that can cause neurofibromatosis has more severe effects in whites than in blacks). These and other uncertainties pose an agonizing treatment dilemma. The choice between radical surgery and intensive surveillance—in the form of frequent mammograms—might be crucial. The recent identification of a second breast cancer gene, *BRCA2,* complicates matters even more.

Although further advances in understanding the genome might conceivably one day eliminate such dilemmas, most scientists do not expect them to evaporate in the foreseeable future. For a woman who has already been diagnosed with breast cancer, the significance of a positive *BRCA1* test for treatment is murky. And even with a negative test result, a woman still faces the same one-in-eight lifetime risk that all women in the U.S. do. These factors have led the American Society of Human Genetics and the National Breast Cancer Coalition, an advocacy group, to urge that for now, testing for *BRCA1* mutations be carried out only in a research setting. "We will fight any sale of this test before there is consensus on how it should be used," says coalition member Mary Jo Ellis Kahn, a breast cancer survivor with a family history of the disease.

Such caution clashes with commercial agendas. Myriad Genetics, which is based in Salt Lake City, plans to offer *BRCA1* testing by the end of 1996 to all women diagnosed with breast or ovarian cancer and to their close relatives. If only the diagnosed patients were to accept, that would still be more than 200,000 tests a year. Peter D. Meldrum, Myriad's president, says his company is studying the risks for different populations. Collins, though, calls Myriad's plans "potentially premature," because he says he is not sure that enough will be known about the mutations and the treatment options to justify large-scale testing by the end of this year.

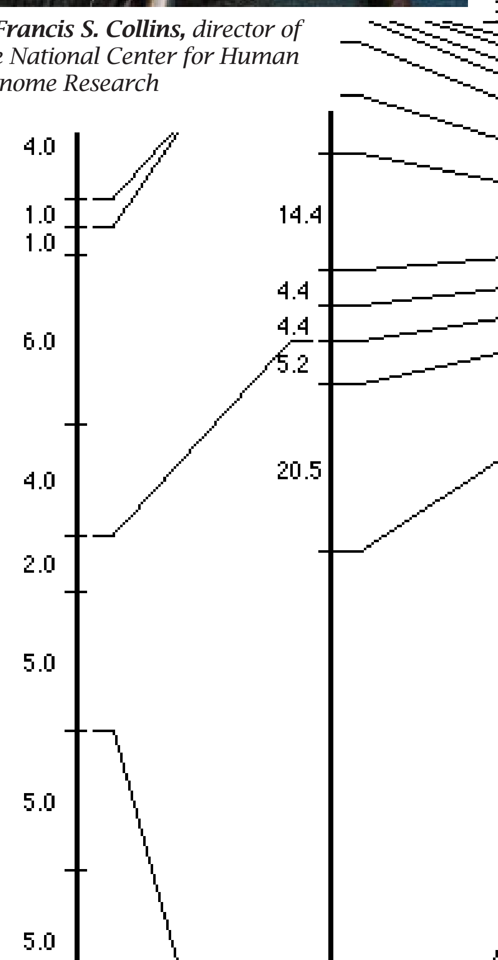
OncorMed in Gaithersburg, Md., is already selling *BRCA1* testing services for use in research protocols for women at high risk. What restrictions should be included in those protocols is a question the National Breast Cancer Coalition is now urgently studying. The company is also testing for mutations associated with a form of colon cancer.

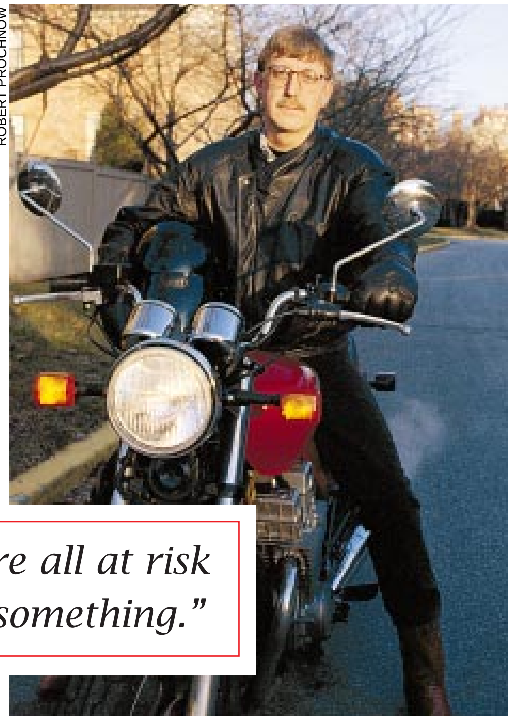
The Food and Drug Administration has historically not regulated genetic testing services. In January, however, a task force on genetic testing created by the genome program was weighing whether to urge the FDA to limit testing for cancer-associated mutations to research uses until there are sound data on the value of treatment possibilities.

Some women who know that they carry a mutated *BRCA1* gene—only dozens, to date—have gone to elaborate lengths to conceal that information from their insurance carriers, says Barbara B. Biesecker of the NCHGR. One fear is that insurers will classify the mutations as a preexisting condition and so refuse to cover treatments related to the "condition." That concern is hardly irrational: health insurance companies often decline to offer policies, or offer inflated premiums, to individuals who have a significant family history of cancer. The National Breast Cancer Coalition was last year rejected several times for health insurance for its Washington, D.C., staff

Chromosome 17

This partial genome map, shown on these and the following pages, was published last December by researchers at the Whitehead Institute



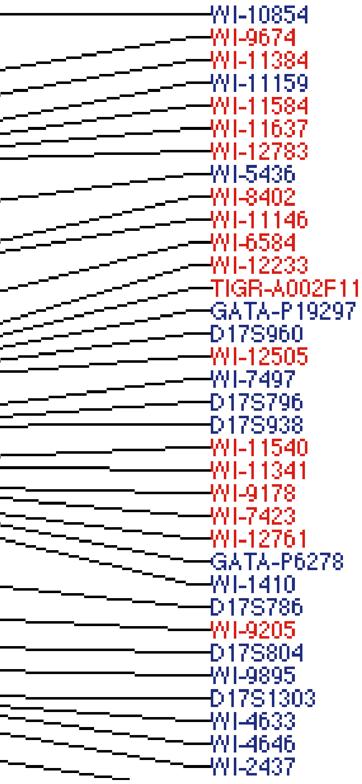


ROBERT PROCHNOW

"We're all at risk for something."

*—**Francis S. Collins,** director of the National Center for Human Genome Research*

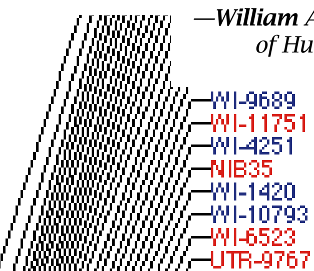
for Biomedical Research and France's Généthon. It can be viewed in full at http://www-genome.wi.mit.edu/cgi-bin/contig/phys_map



ROBERT PROCHNOW

"As of a year ago, we'd isolated over 90 percent of human genes."

*—**William A. Haseltine,** president of Human Genome Sciences*



of eight, because the staff includes some breast cancer survivors. To find coverage, it had to join a larger organization.

The trend toward secrecy in genetic testing seems to be catching on. At a meeting last fall of the American Society of Human Genetics, Thomas H. Murray of Case Western Reserve University asked his audience whether they knew of patients who had requested testing for a genetic trait anonymously or under a false name. Hands shot up all over the room. In many clinical studies, patients are now formally warned that test results could lead to insurance complications if they get into the patient's medical record. Researchers sometimes obtain special legal documents called certificates of confidentiality that prevent the courts from gaining access to data gathered for a study.

Keep a Secret?

Other patients are simply avoiding taking genetic tests, thus forgoing whatever medical benefit they might bring. People with von Hippel-Lindau (VHL) disease, a rare hereditary condition that can cause brain and kidney tumors, often find it hard to obtain health insurance because of the expensive surgeries they might need. Although no prophylactic therapy can prevent the tumors, people with VHL disease can extend their lives by undergoing regular magnetic resonance imaging scans followed by surgical removal of tumors. According to William C. Dickson, research management chair of the VHL Family Alliance, many parents with the syndrome avoid having their children tested for mutations in the recently discovered gene for VHL because they fear that a genetic diagnosis will make their offspring uninsurable.

Parents with polycystic kidney disease, which may be the most common simply inherited, life-threatening condition, also frequently decide for insurance-related reasons not to subject their children to testing, reports Gregory G. Germino, an investigator at the Johns Hopkins University School of Medicine. Approximately 600,000 Americans have the illness—many unknowingly. A gene causing many cases, *PKD1,* was identified in 1994 using technologies developed under the genome project. Testing of *PKD1* can sometimes improve medical therapy for a child, Germino says.

Such reports have prompted alarm among health officials. (Collins admits to being "passionate" on the subject.) A working group on the ethical, legal and social implications of the human genome program, together with the National Action Plan on Breast Cancer, a presidential initiative, recently recommended that insurance providers be prohibited from using genetic information, or an individual's request for testing, as a basis for limiting or denying health insurance.

Currently insurers do not usually ask directly for results of genetic tests; inquiries about the health or cause of death of a person's parents are sufficient to identify many of those at high risk. But insurers may consider genetic data for an individual policy. "They don't ask now about genetic testing, but that will change," says Nancy S. Wexler, president of the Hereditary Disease Foundation, who has herself a 50 percent risk of Huntington's disease.

Because patients with genetic diseases are often reluctant to identify themselves, gauging the extent of discrimination is difficult. But new data strengthen earlier anecdotal reports suggesting the phenomenon is widespread. In one of the first extensive surveys, which was scheduled for publication in January in *Engineering and Scientific Ethics,* Lisa N. Geller of Harvard Medical School and her co-authors describe how they sent questionnaires to people who, though free of any symptoms, are at risk for acquiring a genetically based illness. Of the 917 who responded, a total of 455 asserted that they had been discriminated against after they revealed a genetic diagnosis.

Follow-up interviews by the researchers provided details of health and life insurers who refused or canceled coverage, adoption agencies that required prospective parents to "pass" a genetic test (but on one occasion misunderstood the results) and employers who fired or refused to hire on the basis of a treatable genetic condition or the mere possibility of one. Paul R. Billings of the Veterans Affairs Medical Center in Palo Alto, Calif., one of the study's authors, declares that "the public will reject genetic testing out of fear of discrimination." In a separate study by E. Virginia Lapham of Georgetown University and others, 22 percent of a group of 332 people who had a genetic illness in their families reported having been refused health insurance.

Several European countries have taken steps to prevent abuse of genetic data. Basic medical insurance is not a major concern in Europe, because it is guaranteed by governments. Yet France, Belgium and Norway all have laws preventing the use of genetic information by life and medical insurance companies and by most employers. The Netherlands guarantees basic life insurance, and Germany has some protections.

In the U.S. several states have enacted legislation that limits discrimination based on genetic data. Employment discrimination is prohibited by federal law, and several bills now before Congress would discourage or prevent gene-based insurance discrimination nationally. But their prospects are uncertain.

The potential for psychological harm from DNA testing is receiving growing attention. Because a test may have implications for all the members of an extended family, powerful feelings of guilt and sadness can disrupt relationships. Fear of such consequences may explain the unexpectedly low utilization of a test that has been available for some years to identify most carriers of cystic fibrosis—who are not themselves at risk.

Genetic counselors have formed a strong consensus that because of the potential for harm, children should not be tested for mutations predicting diseases that will not develop until adulthood, unless there are possible medical interventions. That principle rules out testing children for Huntington's, because there is no preventive therapy against developing the rocking motions and mental impairment that characterize the illness. Yet parents do seek testing of their children: in one case, to avoid paying for a college education if the youngster was likely to succumb.

The principal U.S. network of testing laboratories is known as Helix. According to a survey by Dorothy C. Wertz and Philip R. Reilly of the Shriver Center for Mental Retardation in Waltham, Mass., 23 percent of the network's labs technically capable of checking for the Huntington's mutation have done so in children younger than 12 years. More than 40 percent of Helix laboratories had performed tests for patients directly, with no physician involved. Yet the public can easily misunderstand the meaning of genetic diagnoses, Wertz notes. Moreover, she says, many physicians are not well enough informed to be giving genetic advice.

On the credit side of the ledger, it is clear that some patients in families afflicted with hereditary colon cancer, and possibly breast cancer, too, have already made wise medical choices as a result of discoveries facilitated by the genome program. Some with a mutation predisposing them to colon cancer, for example, have had their colons removed as soon as threatening changes started to occur—a procedure that probably saved their lives. Eagerly awaited novel therapies, though, are further in the future.

Collins notes that only six years after a team that he co-led found the gene associated with cystic fibrosis in 1989, drugs developed to counter effects of the mutated gene were already being tested in patients. How soon a definitive treatment will emerge, however, is anyone's guess.

The treatment prospect that has most gripped the public imagination is gene therapy—an approach that would be better described as gene transplantation. But attempts to treat familial hypercholesterolemia, cystic fibrosis and Duchenne's muscular dystrophy have each resulted in failures over the past year, apparently because patients' cells did not take up enough of the transplanted genes. The earlier treatment of adenosine deaminase deficiency described in this magazine by W. French Anderson [see "Gene Therapy," SCIENTIFIC AMERICAN, September 1995] showed at best a modest effect. Last December an NIH review concluded that "clinical efficacy has not been definitively demonstrated at this time in any gene therapy protocol."

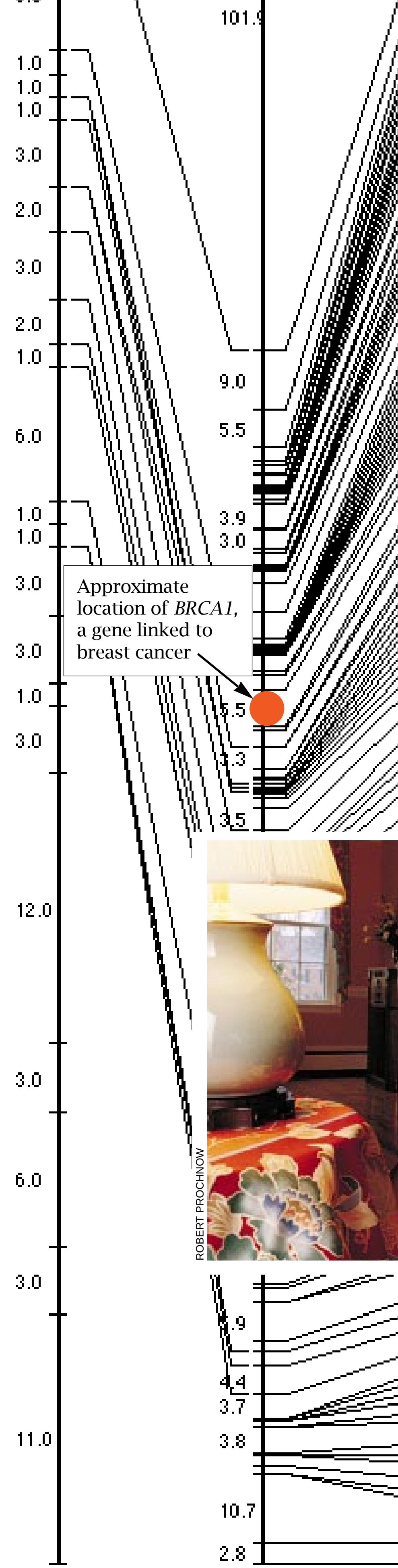
Following the Money

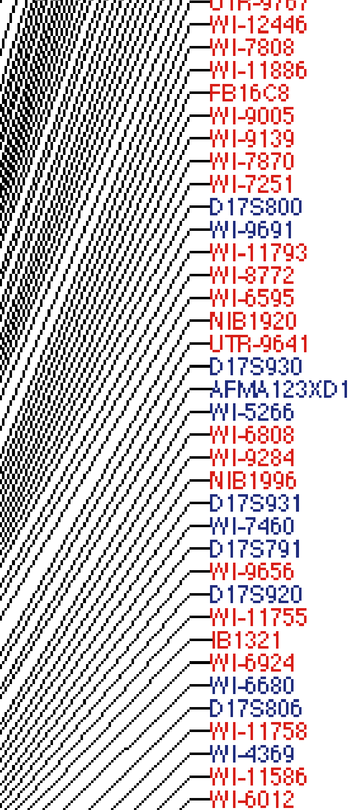
In spite of these problems, the burgeoning genetic revolution is already causing seismic reverberations in the business world. Pharmaceutical companies have staked hundreds of millions of dollars on efforts to discover genes connected to disease, because they could show the way to molecules that might then be good targets for drugs or diagnostic reagents.

One favored shortcut strategy, pioneered by J. Craig Venter of the Institute for Genomic Research in Gaithersburg, Md., analyzes intermediate products—messenger RNAs—created when genes are activated in a cell. The technique produces chemical labels that reveal some basic information about their associated genes for a relatively small investment of effort.

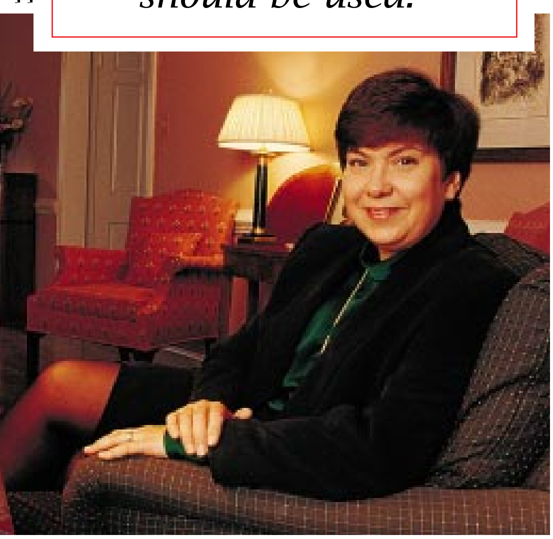
Some scientists were initially unimpressed by Venter's end run around traditional techniques; his approach does not provide the comprehensive information that full-scale sequencing provides. Nevertheless, it has proved its worth. Two companies have embarked on a microscopic gold rush to find lucrative genes using Venter's strategy: Human Genome Sciences (the parent of Venter's organization) and Incyte Pharmaceuticals in Palo Alto, Calif. Both are now busily patenting sequences.

William A. Haseltine, president of Human Genome Sciences, says his company has identified 90 percent of all human genes and has used dozens of

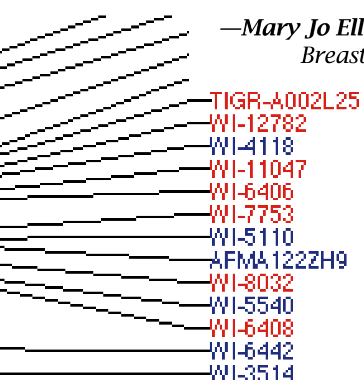




"We will fight any sale of this test before there is consensus on how it should be used."



—***Mary Jo Ellis Kahn,*** *National Breast Cancer Coalition*



them to make proteins with potentially therapeutic properties. Human Genome Sciences has also made "a substantial impact," Haseltine declares, on the drug development program of its principal business partner, SmithKline Beecham. Incyte, for its part, claims to have identified "most" human genes and has found customers for its databases.

The prospect of commercial exploitation of the genome is motivating protests in some quarters. Most of the political flack is being taken by an initiative known as the Human Genome Diversity Project. The diversity project—which is not formally linked to the genome project—aims to study the variations in genetic sequences among different peoples of the world.

Supporters of the diversity project, which was conceived by Luigi Luca Cavalli-Sforza of Stanford University, note that the sequence generated by the genome project proper will be derived mainly from DNA donors of European and North American origin. Studying the 0.1 percent variation among people around the world might yield valuable information about adaptation, Cavalli-Sforza reasons.

Henry T. Greely, a Stanford law professor who is a co-organizer of the genome diversity project, recognizes that data about genetic variation could invite racists to concoct arbitrary rationales justifying discrimination. But he says the project will accept its responsibility to fight such abuses, and he notes that the available data point to how superficial the racial differences are: most of the 0.1 percent of variation in humans occurs among members of the same race, rather than among races.

Still, the project is struggling politically. One of the thorns in its side is the Rural Advancement Foundation International (RAFI), a small campaigning organization based in Ottawa that opposes patents on living things. Jean Christie of the foundation says her group fights the diversity project because it will produce cell lines that can be patented by gene-bagging companies from rich countries.

Greely insists that the project's protocols eliminate the possibility that samples will be used for profit without the consent of donors. Notwithstanding that assurance, a committee of the United Nations Educational, Scientific and Cultural Organization (UNESCO) has criticized the project's lack of contact with indigenous groups during its planning phase. Concerns about exploitation have also been fueled by a controversy over a patent granted to the NIH on a cell line derived from a Papua New Guinean man. Such worries may explain why the genome diversity project has so far failed to obtain large-scale funding.

How the genome patent race will play out is still unclear. Collins says that a scramble to patent every sequenced gene would be "destabilizing," as it would imperil cooperation among investigators. Traditionally, scholars have been free to carry out research unhampered by patents. That freedom cannot be taken for granted, states Rebecca Eisenberg, a patent expert at the University of Michigan. As more scholars forge commercial ties, proprietary interests in the genome may be more vigorously enforced.

Before a gene can be patented, the "inventor" has to know something about its function, in order to meet the legal requirement of utility. Industry, however, controls most of the research muscle that can efficiently discover useful properties. So commercialization seems to be an inevitable consequence of the genome's scientific exploration, as it is for other explorations. Although the U.S. Patent Office recently held hearings to examine questions raised by gene patenting, turning back the clock to disallow such patents, as some critics urge, seems unlikely. And so long as corporate dollars do not stifle collaboration, people in rich countries will probably benefit from the feeding frenzy.

How much the rest of the world will gain, though, is a valid question. Some gene therapies now being evaluated would be tailored to individual patients. But James V. Neel, a pioneer in human genetics at the University of Michigan, warned researchers recently that "individual therapies will be too expensive" for widespread use. Humble interventions such as improved diet and exercise may ameliorate adult-onset diabetes more cost-effectively than genetic medicine, he points out. Neel urged geneticists to pay attention to the deteriorating environment many humans inhabit as well as to their DNA.

Still, the gene race is on. Better medicines will be found; some people will make fortunes, and some will probably suffer harm. But it is a safe bet that although all humans share DNA, not all of them will share in its bounty. The World Health Organization reports that 12.2 million children under the age of five died in the developing world in 1993. More than 95 percent of those deaths could have been avoided, according to the agency, if those children had access to nutrition and medical care that are already standard practice—in countries that can afford them. For many of the world's unfortunates, genetic medicine may always be a distant dream.

Exploring Chemical Bonds

The world is held fast by molecular bonds. They shape every sight, sound and texture. All of life's processes, from respiration to reproduction to repairing DNA, are carried out entirely through interacting molecules. Even consciousness itself is regulated and structured by electrochemical interactions in our brain. Clearly, understanding the strength and nature of these bonds is vital to science. Using the device described here, you can conduct original research into the chemical nature of creation.

When molecules interact, old bonds break, and new ones form to produce various chemical species. If, on the whole, the new molecules are more tightly bound than the old were, heat is given off, and the temperature of the solution containing the reacting molecules rises. If they are less tightly bound, heat is absorbed, and the temperature falls. Measuring how much heat flows into or out of these chemical interactions reveals information about these bonds, such as the amount of energy needed to hold molecules together.

The measurement technique, called titration calorimetry, requires simple apparatus. Place three test tubes into separate Styrofoam coffee cups and gently pack insulation around them. Insulating foam or finely shredded newspaper works well. Label the cups A, B and C. A and C will hold the individual test chemicals dissolved in solution. B will hold exactly the same quantity of solution as A but without the test chemical. You will also need three rubber stoppers: two with two 1/8-inch-wide holes and one with one 1/8-inch-wide hole. Some test chemicals to try initially are solutions of vinegar and baking soda or of concentrated sugar water and vinegar. When you place a little of C's contents into A's, the temperature in A will change. A thermogram, a plot of total energy produced against the concentration of C molecules in A, displays the results graphically.

Two circuits consisting of several resistors and operational amplifiers (op-amps) are necessary to measure and control the temperature difference. One circuit is a sensitive thermometer, which detects the temperature changes that occur during a reaction in A. Small resistors then function as heaters to compensate for the temperature difference, raising the temperature of either A or B as required. The second circuit, called a calorimeter, measures the amount of heat delivered by the resistors, which equals the amount of energy generated or absorbed during the reaction.

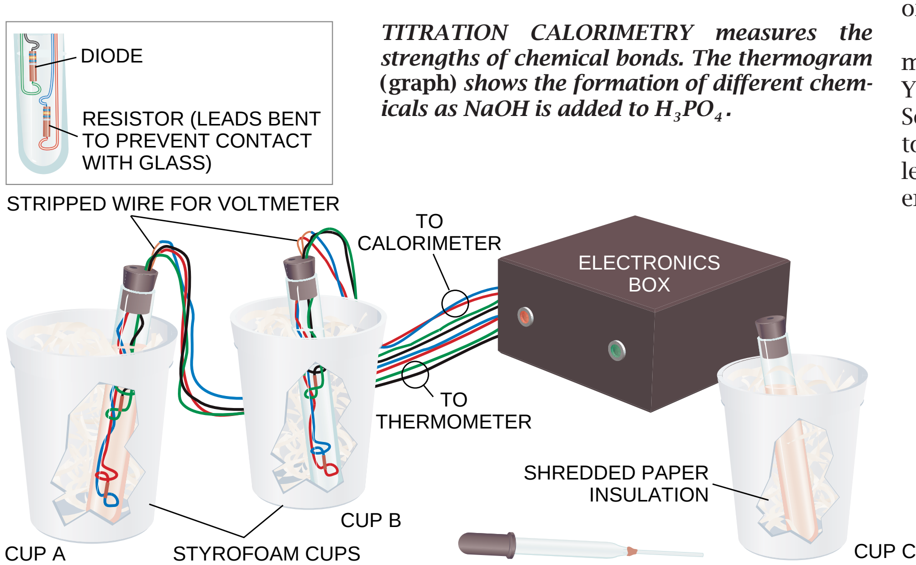
Diodes serve as the key components of the thermometer circuit. When a diode conducts current, a voltage of about 0.7 volt appears across its leads. The precise value depends on temperature. The voltage falls about two millivolts for each degree Celsius increase—about 40 times the signal produced by a thermocouple, the device typically used to record temperature. Indeed, the diode thermometer can distinguish differences as small as 0.01 degree C.

Buy a package of type 1N914 diodes (Radio Shack sells them in packages of 50 for about $2). Select two as your temperature probes. Solder 18 inches of wire to each lead and insulate the exposed wires with latex-based enamel paint. They are connected to an instrumentation amplifier, the heart of the thermometer. This powerful integrated circuit is easy to hook up and costs about $20 from Analog Devices in Norwood, Mass. (telephone 800-262-5643).

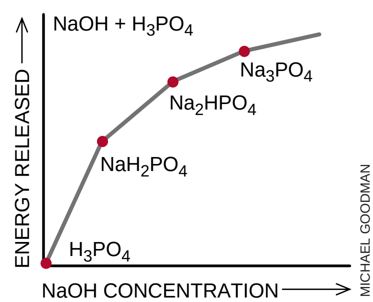
The thermometer first has to be set to zero. Place both diodes in their test tubes. Flip on all three dip switches to set the amplifier's gain to 1,000. Using a voltmeter, adjust the resistor labeled R1 until the voltage at pin 9 reads zero. Throw the dip switches off and adjust R2 until pin 9 voltage is again zero. Flip the switches back on to reset the gain to 1,000.

Next, build the calorimeter circuit. Resistor R3 sets the sensitivity. With the aid of a voltmeter, adjust R3 until the voltage measured at pin 2 of op-amp 3 reads 10 millivolts. Then adjust R4 until pin 3 of op-amp 2 reads –10 millivolts. If your light-emitting diodes (LEDs) turn on when no chemicals are being mixed, use R3 to increase the voltage at pin 2 of op-amp 3 until they remain off.

The simplest possible heaters are made from 10-ohm, 1/4-watt resistors. You will need two, one each for A and B. Solder 20-gauge insulated copper wire to each one and electrically insulate the leads with two coats of latex-based enamel. Bend the wires so that they



***TITRATION CALORIMETRY** measures the strengths of chemical bonds. The thermogram* **(graph)** *shows the formation of different chemicals as NaOH is added to H_3PO_4.*



push each resistor into the center of its test tube. You will also want to insulate them thermally. Check your local hardware store for products such as Dip It—a rubberized solution that sets to form an insulating coating.

For each test tube, thread the wires for the resistor through the same hole in the stopper (the second hole is for the addition of chemicals). Position the resistor about ¼ inch above the bottom of the test tube. Fill the hole with silicone aquarium cement. Finally, strip and solder the leads above the stoppers, leaving the connections exposed. (Designs for other heaters, suitable for organic molecules and strong corrosives, appear on the World Wide Web page of the Society for Amateur Scientists.)

If A's temperature drops by as little as 0.01 degree C when chemical C is added, the circuit will light the green LED in the calorimeter circuit. It also passes a small current through the heater, which gently warms test tube A. When A reaches C's temperature again, the circuit turns off both the heater and the LED. With a stopwatch, you can measure how long the LED remains lit. That measurement, plus the voltage reading across the resistor, enables you to calculate the energy required to return A to its original temperature. If the test tubes are well insulated, this amount will be exactly the same as the energy absorbed by the molecular bonds.

If the temperature of A rises, the red LED lights. The heater in B then activates until B and A are again at the same temperature. The energy needed to elevate the temperature of B to that of A is the same as the energy liberated by the chemical interactions in A. To ensure that A and B always maintain equal volumes, you will need to add the same amount of solution from C into both A and B. Because B does not contain the critical chemical, however, adding the substance to it will not release any chemical energy.

To transfer chemicals from test tube to test tube, you will need a modified medicine dropper. Insert a 100-microliter capillary tube into the dropper and make an airtight seal with aquarium cement. Poke a hole into the dropper bulb with a pin. Capillary action will draw up a precise amount of fluid from B. But be careful. The amount of fluid drawn up will depend on the fluid's surface tension. Measure how far the fluid flows up the tube and, by comparing that height to the 100-microliter mark, estimate the volume of fluid. When transferring the chemical to A, place your finger over the pinhole as you squeeze.

Before recording data, place a small amount of each chemical into test tube A to see which LED lights up. If it is the red one, clip your voltmeter to B's heater. If green, clip it to test tube A's.

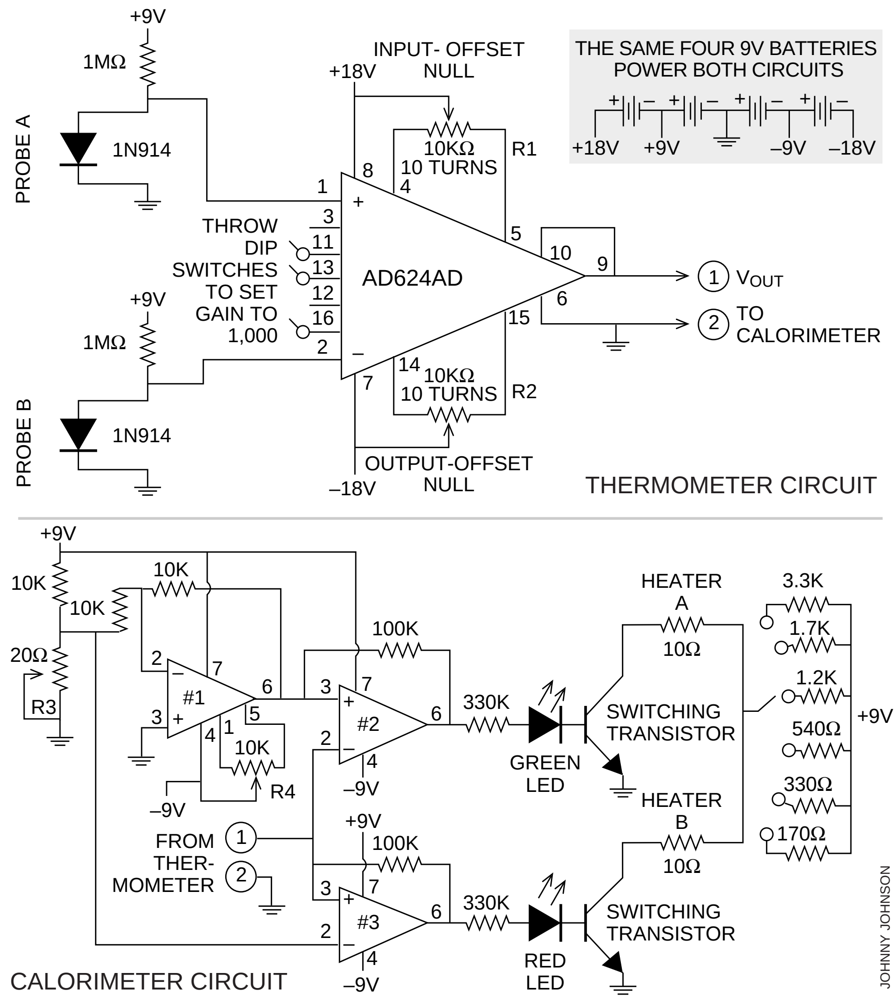
Now prepare your chemicals. Chemical C should be as concentrated as possible. Experiment to find the right concentration for A. Start with a tenth of C's. It should take at least 30 steps to reach the end point of your thermogram (when the LED no longer lights). To find the correct amount, divide 30 by the number of steps and multiply by the original concentration in A. Immerse the containers in water for 15 minutes to equalize their temperature. Finally, pour the solutions into their test tubes.

To take data, insert the capillary tube into C and draw up the solution. Release it into B without dipping the capillary tube into the fluid. Nothing should happen. Repeat the procedure, this time placing C's solution into A. Now the LED should brighten, and your voltmeter should jump. Gently and continuously swirl cup A to make sure its contents are well mixed. Note the voltage and how long the LED stayed on. For greatest accuracy, the LED should remain lit for 100 seconds.

The amount of energy deposited in calories is $0.2390\,(V^2/R)\Delta t$, where V is the voltage across the heating resistor, R is the resistance in ohms of the resistor, and Δt is the time in seconds that the LED remained lit.

In a thermogram, if every molecule of A interacts, the graph's slope will change sharply at the end points. If only a small fraction interact, the end points will be rounded. A changing slope can also indicate the concentrations at which different chemicals form.

To learn how to interpret thermograms and obtain more information about this project, send $5 to the Society for Amateur Scientists (SAS), 4951 D Clairemont Square, Suite 179, San Diego, CA 92117, or download the information from http://www.thesphere.com/SAS/ or from Scientific American'*s area on America Online.*



TWO CIRCUITS form the basis for the experiment. The thermometer circuit detects changes as small as 0.01 degree Celsius. The calorimeter circuit, using type 741 op-amps, measures the heat made or absorbed during the reaction.

MATHEMATICAL RECREATIONS by Ian Stewart

Playing with Quads and Quazars

A friend of mine, G. Keith Still, is a computer scientist who, together with two former soldiers, runs a company that handles crowd control at Wembley Stadium in England. His main professional interests involve simulating crowd dynamics and designing appropriate barriers. But Still is a very inventive person, and he recently told me about a game he came up with several years ago.

The game, which Still calls Quad, is played on a grid—11 cells high and 11 cells wide—from which the four corner cells have been removed. That leaves 117 available cells. The two players each possess 20 pieces, either all red or all black, called quads, and seven (or for a shorter game, six) white pieces called quazars. They take turns placing one of their quads on the board in an unoccupied cell. The aim of the game is to get four of your quads at the corners of a square. The square may have edges that are parallel to those of the board, or it may be tilted [*see top illustration at right*].

To claim a win, a player shouts "Quad!" when placing the fourth and final corner of a square into position. Players do, however, sometimes overlook squares that have formed accidentally. When a player notices an unacknowledged square, he or she may shout "Quad!" at the start of his or her next turn. Any oversights must be rectified while the game is still in progress.

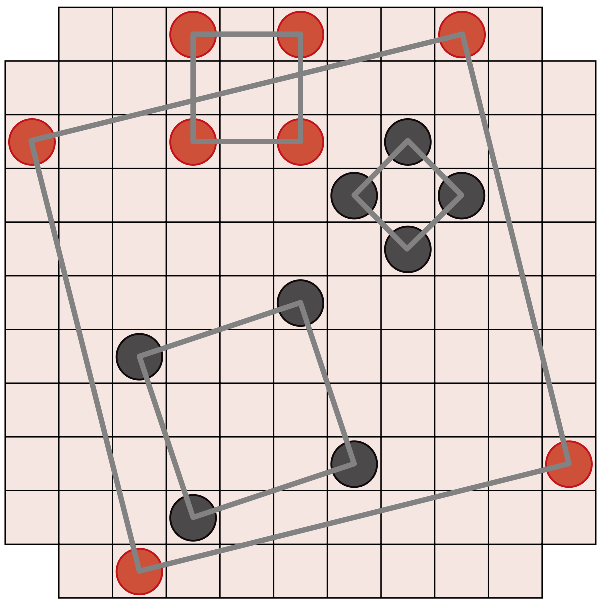
The bottom illustration at the right shows one way in which black can place a series of quads so that at each stage red is forced to make a certain move to prevent black from winning in the next round. This kind of forced play does not make for a very interesting game, which is where the quazars come in. Quazars are used purely for blocking and so do not count toward the formation of squares. They are, after all, white for both players. There are several rules for playing quazars. First, it must be your turn to move. Second, you may play as many quazars as you wish—up to your limit of seven—but you must play them before you play your quad. And you must play a quad to complete your turn.

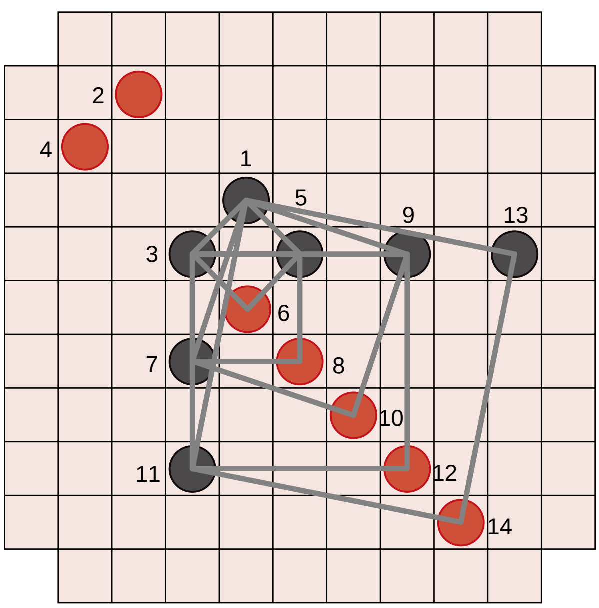
Finally, there are two technical rules. If on the last move of the game, a player leaves a position that would have forced a win on the next move had there been any quads left to play, then that player is declared the winner. If the game ends without either player forming a square, the player having the most unplayed quazars wins. If both players have the same number of quazars, the game is a draw.

Because a huge number of possible squares can be formed, Quad can be surprisingly complex. It is, for example, relatively easy to make a "double square" move—one in which two potential squares are created simultaneously. By potential square, I mean a square that has three of its four corners already occupied. Your opponent can block the completion of several squares in one move using quazars. So the double square move proves to be a good strategy, because it forces your opponent to use up quazars more quickly.

Experience shows that placing a quad in the central cell makes for a reasonable opening move. After that, you need to be on the lookout for potential or accidentally finished squares having unusual orientations. You must also watch for overlapping squares that might lead to double square moves for your opponent. The illustration on the opposite page takes you through a game; potential squares are highlighted as they arise.

If you enjoy Quad, there are any number of variants to play under different circumstances. Young children, for example, typically find the game more manageable on a smaller board, in which case the number of quazars should be reduced accordingly (five quazars for a 10-by-10 board, four for a 9-by-9 board and so on). So, too, if three players wish to compete, each should use only four quazars (and 20 quads, say, black, red and perhaps yel-





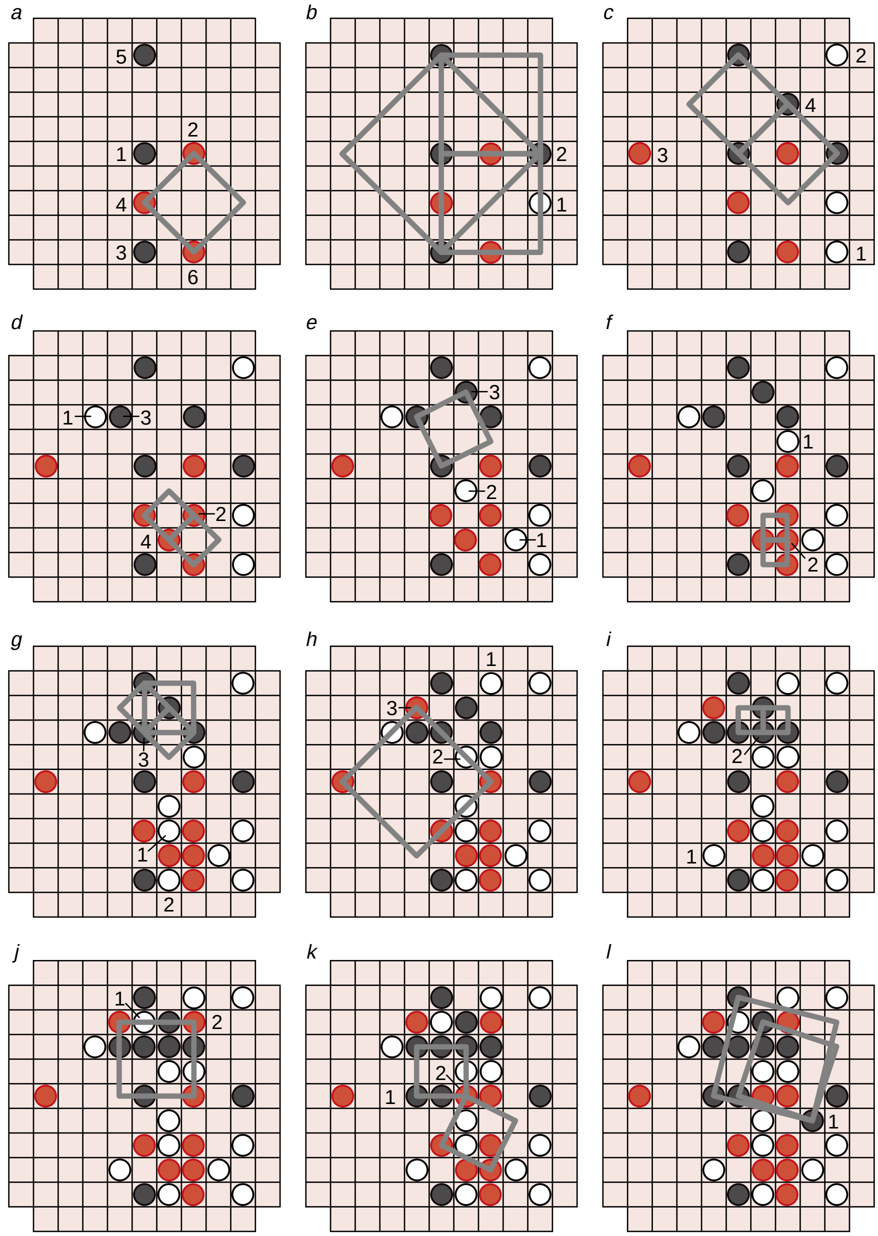
JOHNNY JOHNSON

***WINNING SQUARES can be any size* (top). *So black can make a sequence of moves that, in the absence of quazars, keeps red on the defensive* (bottom).**

low). Four players should each get three quazars, and five or six players should each have two.

Still also describes several additional games. Quad Trek is similar to the standard two-player game, but each player has only six quads and six quazars. After all the quads are played, a turn consists of picking up a quad of your color and placing it elsewhere. Quazars may be played at any time, as in the standard game, but once played they cannot be moved. In Quad Duel, each player has quads of two colors, or even more, and plays one quad of each color per turn. A player may shout "Quad!" only when he or she has formed a square from four like quads.

Each player in a game of Quad Rapid gets six quads and six quazars. At each move he or she may either place a new quad or move an old one. Quazars are played as usual. Quad Bridge requires four players, who form two pairs. Players sit opposite their partners around the four sides of the board. One pair have black quads, the other red. Play proceeds clockwise, and no talking is allowed. You must work out your partner's strategy before your opponents do. But you are permitted to send signals—say, by waving your arms, banging your head against the table or leaping up and down. If readers find other exciting versions of Still's game, send them in, and maybe they will take their turn in "Feedback."



SAMPLE GAME shows all potential squares as they arise in play. The numbers indicate the order in which pieces are put on the board. After 10 moves, red has run out of quazars, and black wins by forcing a double square move.

JOHNNY JOHNSON

Feedback

The October 1995 column prompted a number of letters clarifying the history of the "Morse-Hedlund" tripleless sequence, 0110100110010110. . . .

Jeffrey Shallit of the University of Waterloo in Ontario wrote: "Nowadays this sequence is usually attributed to the Norwegian mathematician Axel Thue, who wrote about the repetition problem in a series of papers beginning in 1906. He also proved it was overlap-free, a stronger property. The application to chess was, to my knowledge, first observed in Marston Morse's abstract in the *Bulletin of the American Mathematical Society,* Vol. 44, No. 9, page 632; September 1938. For a humorous view of Morse's application, see 'A mathematician gives an hour to chess,' by Donald MacMurray in *Chess Review,* Vol. 6, No. 10, page 238; October 1938. Recently several mathematicians observed that the sequence was implicitly contained in an older paper, 'Mémoire sur quelques relations entre les puissances des nombres,' by E. Prouhet in *Comptes Rendus des Séances de l'Academie des Sciences,* Vol. 33, No. 8, page 225; August 25, 1851."

I. J. Good of Virginia Polytechnic Institute and State University noted that Machgielis (Max) Euwe, World Chess Champion 1935–1937, invented the same sequence in "Set Theory Observations and Chess" in *Proceedings of the Academy of Science Amsterdam,* Vol. 32, pages 633–642; 1929. He adds, "This article provoked me to invent (in 1943 or '44) the 'reflection order' for the teleprinter five-unit code: for details, see my article 'Enigma and Fish' in *Codebreakers,* edited by F. H. Hinsley and Alan Stripp (Oxford University Press, 1993). This code is now called the Gray code and was independently invented and patented by F. Gray for analog-to-digital conversion." *—I.S.*

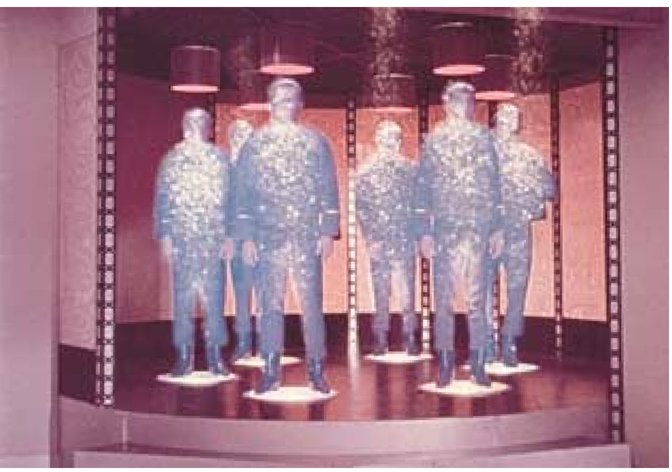
Next-Generation Education

Review by Michio Kaku

THE PHYSICS OF STAR TREK, by Lawrence M. Krauss. BasicBooks, 1995 ($18.50).

Almost 20 years ago Nobel laureate I. I. Rabi, addressing a symposium at Columbia University, chided the audience of physicists for failing to communicate the excitement and drama of their discipline to the general public. To add insult to injury, he charged that even lowly science-fiction writers had done more than professional scientists to impart the romance and fascination of science to the next generation. Nothing has done more than the Star Trek phenomenon to show how right Rabi was. Since the debut of the first *Star Trek* television series nearly 30 years ago, an entire generation of young scientists has been weaned on the series and its spin-offs, fluent in the palace intrigues of the Klingon Empire as well as the periodic table and spectrum of subatomic particles. In my classes, I have found Star Trek to be a powerful tool for attracting students to physics and helping them grasp its concepts.

Albert Einstein once said that imagination is more important than knowledge. If he were still alive, I suppose he might have written *The Physics of Star Trek* himself. Sadly, it has taken three decades for someone else to take on the task: Lawrence M. Krauss, chairman of the physics department at Case Western Reserve University and author of several other, more sober-sounding science books. What took so long?



EVERETT COLLECTION

TRANSPORTER is a thought-provoking but physically implausible technology.

Just five years ago a book bearing such a provocative title might have met with more than a few snickers and shaking heads within the scientific community. When the venerable American Museum of Natural History dared to mount a modest Star Trek exhibit last year, the museum was heatedly denounced in a scathing op-ed piece in the *New York Times,* for allegedly trivializing science. But a lot has happened in the past few years.

The cold war ended, and with it came a reevaluation of government support for physics. The watchword of the 1990s is "downsizing"; "keeping up with the Russians" no longer justifies our science budgets. The Superconducting Super Collider was canceled, and funding for the National Science Foundation is under the knife. At the same time, enrollment in college physics programs is plummeting, and physicists in once sacrosanct laboratories at AT&T and IBM are finding their resources suddenly curtailed. In these lean times, the future vigor and vitality of physics depends more than ever on our convincing the general public of the value of basic research. Scientists must reach out to largely indifferent, and often scientifically illiterate, taxpayers and persuade them to pay our bills.

In sum, it is high time for us to look at Star Trek—which in repeats and new iterations reaches tens of millions of living rooms every week—in an entirely new light.

I am glad to report that *The Physics of Star Trek* is up to the task, succeeding admirably in keeping the science airtight and honest while still respecting the mythology and traditions that surround the series. The book is easy to read, faithful to the physics, full of Star Trek trivia and always entertaining. There is something for nearly everyone, from the person who casually catches an occasional episode to the addicts (like myself) who have seen every episode and can quote lines of dialogue from memory.

Krauss realizes that separating fact from fiction is no mean feat, especially when he is attempting to forecast what physics will be like in the 23rd century and beyond. Even a genius like Isaac Newton, if asked to extrapolate 300 years into the future, would surely have missed most of the subsequent great revolutions of science. And much of what is modern physics today would have been dismissed as pure rubbish a mere 50 years ago. As Arthur C. Clarke once said, "Any sufficiently advanced technology is indistinguishable from magic."

Krauss deftly spots a number of obvious mistakes (in any century) that appear in Star Trek. Some of these errors are real howlers, such as the sound heard when explosions occur in the silent vacuum of space. Others boldly defy known laws of physics: starships racing faster than light, for example, or wriggling through cracks in the event horizon around a black hole. (In all fairness, the special-effects people at Paramount Pictures probably are aware that they are engaging in scientific violations, but they also know that those noisy explosions pump up the level of excitement on the television screen.)

Nevertheless, Krauss gives creator Gene Roddenberry proper credit for correctly incorporating many ideas from modern physics, such as antimatter (first predicted by P.A.M. Dirac in 1930 and discovered by Carl Anderson two years later). When antimatter meets matter, the two annihilate each other and produce large quantities of energy, a reaction that could in principle form the basis of an incredibly efficient propulsion system. The only problems facing antimatter engines are the enormous economic and technical ones: producing large quantities of antimatter would bankrupt any modern nation.

But of course, what every Trekker really wants to know is the physical basis of the warp drive and the transporter, which have forever altered the English lexicon (even the most Trek-averse must have heard the phrase "Beam me up, Scotty"). Could such devices ever exist? Most people, I find, naturally assume that the warp drive is less likely to conform to the laws of physics than the transporter. Surprisingly enough, as Krauss points out, the opposite is true.

Scores of solutions to Einstein's equations allow for the possibility that, given enough energy, one might be able to punch a hole in space and even bend time into a pretzel. Creating a "warp" in space-time could permit faster-than-light transportation without breaking known physical laws. (Stephen W. Hawking once offered "experimental proof" that time machines were impossible. The evidence, he claimed, is that we do not see tourists from the future, so time machines will never be invented. In a preface to *The Physics of Star Trek,* Hawking finally drops that tongue-in-cheek "proof" and admits that time machines are a theoretical possibility.) Lately it seems that almost every month some physicist announces a new design for a time machine in the pages of the quite serious journal *Physical Review D.*


PARAMOUNT/EVERETT COLLECTION

PHASER requires hydrogen-bomb power in a handheld device.

Even if time travel and warp drive are possible, however, we may have to wait quite a while before they are realized. The amount of energy required to open up holes in space, known as the Planck energy, equals 10^{19} billion electron volts; that is about 10^{16} times more energy than the most powerful particle accelerator can currently achieve. So do not believe anyone who claims to have invented a time machine in his or her basement.

And just because the theory of relativity does not preclude time travel does not necessarily mean that it can be done. A complete resolution to the question of whether wormholes can be used to tunnel through space and time requires going beyond Einstein's theory, to an as yet nonexistent quantum theory of gravity. At the instant that one enters a wormhole (or a time machine), quantum fluctuations might build up, which would cause the space-time tunnel to collapse in on itself. We will probably have to wait for a "theory of everything"—such as superstring theory, my own specialty—for the final answer.

At least for time machines and warp engines, there is a mathematical basis with which to frame the debate. But for transporters, which instantly "beam" people from one place to another, it is hard to know where to begin the discussion. For physicists, the thought of disassembling a person's molecular pattern, shooting it across space and then reassembling it in another location gives us headaches. Krauss struggles at length to convey the sheer magnitude of the problem. Even the simplest aspect of the transporter—storing the information describing the identity and location of every particle in a human being—would require 10^{28} kilobytes of memory. In comparison, the information in every book ever written would take up only 10^{12} kilobytes. The process of disintegration and reintegration is sure to be considerably more difficult.

Krauss does a yeoman's job of exploring every possible avenue, but he finally admits that physicists are at a complete loss to explain how a transporter might work. (One sneaky way out, which he does not explore, is to imagine the transporter as a kind of short-distance, personal warp drive.) Scotty will probably not be beaming anyone up, even in the 23rd century. Similarly elusive is the "holodeck," a kind of recreational fantasy room seen in the newer Star Trek television shows. This device can create at will completely convincing, three-dimensional objects, people and environments. Krauss speculates that the holodeck might use a combination of transporter and hologram technology, but that, too, is a real stretch.

Scientists' most common pet peeve with Star Trek is actually a matter of biology, not physics: the irritating fact that all the aliens look obviously human. I suppose the Screen Actors Guild contract requires that all aliens must be played by union members. And the numerous spacefaring civilizations appear to be at the same technological level, give or take a few years of development. The odds of this happening are astronomically small. After all, only a few thousand miles (and a few hundred years of technology) separated Cortés from the Aztecs, and look at the mismatch of their encounter! Perhaps, taking that example into consideration, it may not be so wise to advertise our presence to other life-forms in outer space.

If I fault Krauss for anything, it is only for minor lapses. I think he is a bit too generous about stretching the bounds of physics in some areas and too conservative in others. For example, he is too kind to Star Trek when discussing phaser weapons. Lasers in the terawatt-power range or beyond, which could quickly vaporize an enemy, are certainly conceivable; the problem is that Star Trek requires that these devices have a handheld power supply, which is hard to imagine. The only known portable power pack that could supply the necessary amount of energy is a hydrogen bomb—not a practical solution for a handheld phaser!

Similarly, Krauss avoids the question of force fields, which are probably not possible, at least not within the bounds of the four known fundamental forces in the universe. (Gravity is attractive, not repulsive, and is extremely weak. Electromagnetic fields are not repulsive to all forms of matter—plastics, for instance. And the weak and strong nuclear forces operate over distances far too short to be of much use.)

In other areas, *The Physics of Star Trek* does not go far enough. Many of the most imaginative episodes of Star Trek deal with parallel dimensions and alternative universes, which Krauss only briefly discusses. Recently there has been considerable scientific interest in the burgeoning field of quantum cosmology, which holds that our universe may coexist with an infinite number of parallel universes, like a bubble floating in a "multiverse" of other bubbles. Most of these parallel universes are probably "dead," containing unstable particles unable to gather together into living organisms, but some of them conceivably resemble our universe, except with a twist of slightly different physics. Perhaps in one of these universes, everything is the same, except that we went to a different college, married a different spouse or chose a different career. I suppose we will have to wait for the publication of a book entitled *The Physics of the Twilight Zone* to clarify matters.

Ultimately, I judge *The Physics of Star Trek* a success because it provides such a lively way to engage the public in thinking about physics and about how scientists attempt to understand the world. As Hawking says in the preface, "Science fiction like Star Trek is not only good fun but it also serves a serious purpose, that of expanding the human imagination."

MICHIO KAKU is professor of physics at the City University of New York, where he specializes in superstring theory. He is the author of Hyperspace *(Anchor Books/ Doubleday, 1995).*

Paleontological Predictions

Review by Meredith F. Small

DOMINION, by Niles Eldredge. Henry Holt and Company, 1995 ($25). **THE SIXTH EXTINCTION: PATTERNS OF LIFE AND THE FUTURE OF HUMANKIND,** by Richard Leakey and Roger Lewin. Doubleday, 1995 ($24.95).

One of the hallmarks of human consciousness is the ability to evaluate the past and to consider the future. Paleontologists have unearthed a fossil record that establishes our evolutionary lineage extending back about five million years. And now some of those very same scientists are looking forward to consider the ultimate fate of our species. Although such projections might step beyond the boundaries of paleontological science, there is no astrology or alchemy here, no reading of tea leaves. The predictions are based on lessons learned from natural history.

Niles Eldredge's *Dominion* is a personal essay on the human career written by a well-established paleontologist at the American Museum of Natural History. The author's main focus is the overwhelming growth of the species; there are now 5.7 billion *Homo sapiens* on the earth, and, according to Eldredge, the globe is groaning under the weight. Before culture, overpopulation was not a problem. Various hominid ancestors evolved and became extinct in Plio-Pleistocene Africa, along with other large mammals—antelopes, elephants and pigs. Around 1.7 million years ago, armed with fire and tools, our ancestors moved out of Africa and across the planet; wherever they went, large mammals disappeared. Culture, in the form of technology, enabled our kind to migrate relentlessly, reproduce without check and then kill and eat at a rate far beyond the bounds of any local ecosystem.

The shift to sedentary agriculture 10,000 years ago is, to Eldredge, the watershed moment when humans became finally and totally disengaged from nature. Once settled, our forebears used their cultural acumen to build cities, weave complex societies and construct civilizations. We now see this ability as a kind of dominion, and many feel we have a right over the earth and over all creatures in all habitats. As Eldredge puts it, "We told Mother Nature we didn't need her anymore." More frightening, he thinks, is that in recent years humans have moved even further away from the natural processes of life by becoming globally linked—economically, politically and culturally. No other species has defied nature in this way, and Eldredge suspects there will be hell to pay.


BOYD NORTON

ELEPHANT KILLED by poachers is one example of human "dominion" over nature.

Dominion further explains that our fine consciousness, a trait that appeared as an afterthought to our puzzle-solving large brain, is now a selfish feature: other animals are aware of those creatures with which they share their ecosystem, but humans think only of themselves. Our self-absorption, he believes, will be our ultimate downfall. "Our culturally imbued skills have been so successful that we think we have won—we have sought and gained dominion over nature," Eldredge writes. But the victory is only so much smoke and mirrors. Culture may have changed our relationship with nature, and it might now define the way in which we live, but humans remain as subject to the natural order as any other beings. In fact, we are also highly dependent on nature. Instead of recognizing this dependence, however, we exploit it. We are depleting the soils in which we grow our food, cutting down the forests that cleanse the air and polluting the waters that provide everything from drink to sushi. If the rate of human population growth continues, there will soon be no more resources left for us to exploit.

There is a certain contingent—prevalent among those who happen to be comfortable and well fed—who believe that science and technology will save us from the brink. Eldredge does not suggest it, but such thinkers might do well to take a quick trip to an impoverished Third World country, where the process of environmental and personal degeneration is already under way. Every day the majority of our kind deal with a magnitude of poverty and deprivation that is beyond the experience of the industrial West. So far technology has not saved them.

Eldredge's book attempts to deflect some of our inner-directed self-consciousness outward, to take dominion not over the earth but over ourselves. He believes the only way to that goal is through population control, and he is probably right. The problem is that natural selection has programmed us, the animal side of us, to pass on genes; Eldredge's earnest call for a return to harmony with the world glosses over the difficulty of redirecting not only our culture but also our evolved instincts. After millions of years of mating and reproducing, of consuming and destroying, it would be a miracle if the human species could suddenly slam on the breaks and manage itself.

Richard Leakey, scion of the famous family of paleoanthropologists, and science writer Roger Lewin offer a take on the human future that is no less depressing, one grounded in an even broader evolutionary and ecological framework. The authors explain that,


BOYD NORTON

KENYAN ELEPHANT tears up a tree for food, engaging in a natural cycle of renewal.

over the history of life on the earth, there have been five mass extinctions and that, given the extent of human interference, we are probably in the middle of a sixth. The first five might be thought of as improvements—or at least as productive prunings; after each mass extinction, the slate of life was wiped rather clean, and new species emerged and flowered. Humans owe their existence to the mass extinction 65 million years ago, when dinosaurs died and a small, primitive mammal happened to survive. The present extinction, however, is not the result of the usual causes of climatic change, geological forces or asteroid impact. Instead one of the creatures on the earth is rolling the dice, and no one can predict the outcome.

The strength of *The Sixth Extinction* lies in its fluid presentation of the history of life as the interplay of natural selection and extinction. Leakey and Lewin come down squarely on the side of Stephen Jay Gould and others who believe that chance, rather than natural selection per se, has guided the important evolutionary events on this planet. Mass extinction, they argue, is an unpredictable event. Who can say when an errant asteroid will collide with the earth and wreak havoc with ecosystems? Survival of the fittest has nothing to do with this kind of evolution; it is just a matter of a species being in the wrong place at the wrong time.

Some might counter that even within the realm of chance, those who make it, by definition, have been "naturally selected." In any case, no one can fault Leakey and Lewin for their exhaustive and clearly written explanations of both the history of life on the earth and various extinction theories. *The Sixth Extinction* gathers data, theories and writings of various scholars from Gould to Edward O. Wilson to support a cogent argument about chance and extinction as necessary parts of life; quaintly, Cuvier's catastrophism is back. Although Leakey is the paleontologist, and the book is written in his voice, those familiar with Lewin's other books will recognize his journalistic skills and clear style in every sentence.

The more surprising aspect of this book is Leakey and Lewin's insistence that humans are a kind of pinnacle of evolution. Most biologists dismiss this notion as ahistorical or as religious doctrine; bacteria are more widespread than humans, insects more evolutionarily successful and horseshoe crabs have been around longer than just about anything else. Going against the politically correct grain, the authors point out that there is a progression from simple to complex over evolutionary time and that humans are surely the most complex beings of all. Their twist on the Great Chain of Being does, however, dismiss the self-satisfied notion of human superiority: "We cannot seek solace in believing that the evolution of *Homo sapiens* was inevitable. But the evolution of a mental capacity like ours, and the emergence of a degree of self-awareness like ours, probably were inevitable and predictable at some point in Earth history. We just happen to be the species in which they became manifest." It is a kinder, gentler pinnacle.

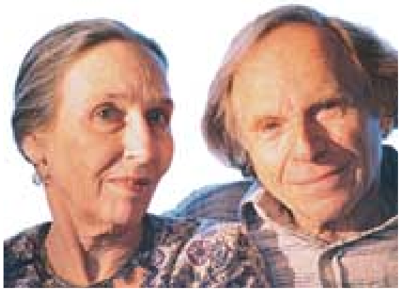
From that height, Leakey and Lewin advance their thesis that humans need to steer the earth away from ecological destruction. Leakey's former position as head of the Kenya Wildlife Service has given him some in-the-trenches experience with trying to save at least one patch of the globe. He believes in the preservation of biodiversity, for both economic and aesthetic reasons, but acknowledges that attempting to preserve a perfect template for biodiversity is impossible. The authors lucidly convey that viewpoint in a chapter discussing efforts to save the elephant from extinction. During a recent dry period, elephants in some of Kenya's parks were destroying woodlands, prompting calls to kill off part of the elephant population. But the apparent destruction was actually part of a longer-term ecological cycle. In times of drought, elephants tear down trees and bushes, opening up new grasslands where zebras and antelopes thrive. Interference in the form of "conservation" would only have derailed the natural process in which elephants lent a tusk.

And all the ecological modeling in the world cannot pinpoint the perfect sustainable community that can exist alongside needy or greedy humans. For now, Leakey and Lewin suggest, the best we can do is change our approach to nature and revise our attitude about our place on the earth.

Humans, even armed with an attitude of dominion, cannot find or create a "balance of nature," because there is no such thing; the world is constantly in flux, with or without us. All we can do now, these two books remind us, is hold back, want less, scale down and be part of the change rather than the cause of it. If we cannot control ourselves and if we cannot consciously learn to reconnect with nature, the sixth extinction will be in our name, a self-immolation.

MEREDITH F. SMALL is associate professor of anthropology at Cornell University. Her latest book is What's Love Got to Do With It? The Evolution of Human Mating *(Anchor Books/Doubleday, 1995).*

COMMENTARIES



WONDERS

by Phylis and Philip Morrison

Symmetry and Symmetry-Breaking in South Africa

We have just returned from an eye-opening, two-month sabbatical in South Africa. This month and next, we will report on the singular symmetries—of the land and of the people—that we witnessed during our travels.

Half a day's drive brings you out from the city of Cape Town, near the southern tip of South Africa, to the Cape of Good Hope and back again. The Cape's stony, salt-sprayed heathlands, austere in early spring before the myriad blossoms erupt, are now a nature preserve, wild and romantic under the steep hills that overlook rushing waters. The long fetch sends unending rollers to crash high above the tangled kelp, while seabirds cry plaintively in the half-gale aloft.

Only the smoothly paved roads and a few brief glimpses of hilltop lighthouses and radio masts remind you that five centuries have passed since Bartolomeu Dias, in command of two caravels and a storeship, first sailed around Africa's southern end in 1487, near the close of Portugal's century-long campaign to reach the Indies by the Atlantic seaway. Dias was blown by a gale past the twin points of the Cape of Good Hope and saw the place only on the way back; it was then that he fixed that buoyant name on a site that pointed to a place of safe haven.

It takes just a few more hours to drive from Good Hope to another cape, Cape Agulhas. This, the actual southernmost point of Africa, is a far more commonplace locale, merely the rockbound lowland shore of a windy seaside resort. Someone has put up a house there at the sea's edge, manifestly the southernmost dwelling in all the continent. Well above the water stands an obsolete lighthouse, now a museum, but retaining its old fortlike facade.

Born and bred north of the equator, we pondered the world anew from our southern vantage. Our earth is round. So big a sphere is bound to be a bit imperfect, but its departure from symmetry is minor. On an accurate model globe, sized to fit in the hand, Mount Everest and the ocean deeps would differ only by the thickness of a sheet of paper. That sphericity is a consequence of simple dynamics. Neither the omnidirectional pull of gravity nor the resistant pressures within the rocky interior favor any fixed direction. A small asteroid or moonlet has too little self-gravity to crush itself, so it could resemble a brick, a potato—even a teacup. A large enough isolated solid object, however, can take on no shape but spherical; a teacup the mass of the earth would quickly disintegrate toward the round. At planetary scale, only bulk motion is able to much modify the sphere's blandness. The earth's rotation causes a minor polar flattening, giving rise to the "oblate spheroid" description often found in textbooks.

The strangeness of the antipodes set no problem for Isaac Newton. South Africans and others from the Southern Hemisphere certainly do not stand on their heads any more than we northerners do. Everywhere on the globe the direction of gravity is relative, the vertical direction determined locally by the plumbline. Here in the south, apples drop just as near the parent bough as they do in Newton's Woolsthorpe, England, home. But a suitably distant, space-dwelling observer would judge that an upright stance here in South Africa appears almost parallel to the horizontal floor of our New England house. (It was fun a few years ago to persuade the locals in Delhi that to point directly to the unseen snowy summit of Everest they should aim not upward but a little *downward* into the ground.)

Because the vertical is local, everyone—every stay-at-home, every traveler and you, too, reader, wherever you are—is on top of the world. Nothing much extends above you, save the air. (Local departures, whether some high mountain across the valley or a mere porch roof, cannot really mar the fact that you stand farther out along your own earth radius than any substantial piece of the planet.) Spherical symmetry unites us.

Up and down may be subjective

terms, but the sun has no such democratic leanings. It is where it is, and we on the earth must live with the consequences—witness the shift of seasons we experienced as we traveled from the North Temperate to the South Temperate Zone. Here on a November afternoon in the Eastern Cape, the white apple blossoms have all been shed, and green apples are daily ripening. Our long flight had turned autumn overnight into spring. A reference book shows the smooth annual wave of average daily temperatures in the nearby port city of East London, with a crest every January and a trough every July.

Take for comparison two U.S. cities, Charleston, S.C., and San Diego. Each is a coastal city at latitude 33 degrees north. The port of East London, at 33 degrees south, matches their distance from the equator, but East London's annual cycle of heat and cold is shifted in time by six months. Here is a perfect opportunity to bury the widespread misconception that the changing seasons are caused by the varying distance to the sun as the earth moves through its orbital ellipse. The southern summer heat and the northern winter cold occur at one and the same time, in two places equally far from the sun's hearth.

Our seasons reversed because of the way the sun's rays strike our round world. In North America, we left a September sun sinking southward in the autumn sky. That same sun was venturing steadily higher here, bringing lengthening hours of daylight and the warming of spring. Sunshine graphically draws the antisymmetry of day and night (again, more clearly observed from space, as when we gaze upon the phases of the moon), a 24-hour rhythm built into all life as profoundly as is the worldwide symmetry of gravity.

Another antisymmetry, though a smaller and more complicated one, shows up in the record of seasonal temperatures. The wavy, annual plot of warm and cold days in East London matches closely in shape the two American ports (allowing for the six-month offset), but it is not identical. In the U.S. cities, the winters are cooler and the summers warmer, by a noticeable three or four degrees Fahrenheit on average. Like East London, the American cities are shore cities, but each one lies at the edge of a broad continental landmass.

The blunt tip of South Africa, in contrast, is a seagirt peninsula, practically an island, immersed in the strong winds and currents of a world ocean that lies both to the east and to the west. South Africa may no longer be so isolated politically from its neighboring countries, but geographically it still stands apart from the bulk of the African continent.

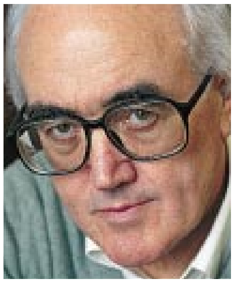
The temperature difference between north and south is a major effect of a grand tempering medium, the Southern Ocean, the weather held in Neptune's control. In the Northern Hemisphere, land and sea run nearly neck and neck for dominance, but in the south the water clearly gains the upper hand. The South Temperate Zone includes only about a tenth of the total land on the earth. Its northern counterpart spreads over fivefold that dry area. Slow to warm and reluctant to freeze, the ocean moderates the extremes of climate.

Two tremendous ocean currents, westward cold and eastward warm, join around Cape Agulhas. It is the balance of these two that keeps East London's mercury on a more even keel than in its northern counterparts. (Currents are familiar causes of climatic disparities; Ireland, bathed in the temperate waters of the Gulf Stream, shares the latitude of Labrador but not at all its bitter, icy winter.) The Southern Ocean grants great leeway for the water to find its course. From Cape Agulhas we looked to the south and faced no inhabited land, only the Antarctic ice cap 3,000 miles away. Yet in the wide and deep waters between, life dwells abundant, displaying itself here magnificently in the months of southern spring.

It is then that baleen whales—specifically, southern right whales—swim out of plankton-rich feeding grounds lying in the cold ocean halfway toward the Antarctic ice. They journey to the western Cape region, to bear their young in the calm bays here. In this entire ocean, no shallow waters are nearer, and none are warm earlier in the season.

One late October day in 1995, an airborne naturalist reported seeing more than 300 whales during a single day's flight along the western Cape shore, a record count. Driving seaward along clifftop and foreshore a few days afterward, we scanned the waves in anticipation. Within two days we saw perhaps 100 whales dive, spout or shake a dark heavy fluke for a moment high over the whitecaps. Most exciting were the several whales we could so clearly see, lolling awash in the slow smooth swell just outside the line of breakers offshore at Pearly Beach and Gansbaai. Each mother weighs as much as eight large African elephants; their calves often swim close beside.

The popular game parks, with their lions, elephants, rhinos and giraffes, rightly thrill most visitors, but the wonder of viewing life free in the wilderness could not be stronger than here in the Cape, poised at the roaring edge of this, the only globe-encircling ocean.



CONNECTIONS

by James Burke

Satisfied Customers

I suppose the modern department store, with its money-back-guaranteed merchandise, is one of the great examples of industrial democracy in action. Thanks to mass production and distribution, I can go back to the shop and get a free replacement for the cup that I found a flaw in last week. It was one of those willow-pattern things. Genuine Wedgwood. An ironic term, really, because Wedgwood's original stuff was fake.

Josiah Wedgwood was a potter who started his career repairing Delft chinaware (fake porcelain, first made for the Dutch middle classes who couldn't afford the sky-high prices of the real thing coming in from the Far East). Then, in 1769, he graduated to crafting his own stuff (fake Greek pieces, first made for the English middle classes who couldn't afford the sky-high prices of the real thing coming in from southern Italy).

The source of Wedgwood's inspiration was an amateur archaeologist and site robber by the name of Sir William Hamilton, who had been appointed British Minister to the Court of Naples in 1764, not long after the first systematic excavation of the nearby ancient city of Pompeii. So there was a ton of classical bits and pieces available for what might charitably be referred to as "collecting." Hamilton's collection grew so big that he published catalogues, one of which influenced Wedgwood.

From time to time, Hamilton would return to England to sell his latest haul of antiquities to biggies such as the Duchess of Portland or the British Museum. On most of these occasions, the sales agent was his nephew, a ne'er-do-well named the Honorable Charles Greville. Now, there must have been something astir in the Hamilton blood, because Sir William's mother had seduced the Prince of Wales, and in 1785 he himself took over Greville's mistress (to "save the boy the expense"). The lady in question was a strapping lass 35 years Hamilton's junior, who called herself Emma Lyon and who was into "attitudes" (posing, in diaphanous outfits, as various classical Greek and Roman personages).

Emma may have learned the trick while working as an "attendant" for James Graham, one of the era's greater electricity quacks. Graham boasted an impeccable scientific background from the University of Edinburgh, where he had studied under such medical greats as Joseph Black, the discoverer of latent heat. Electricity at the time was something like cold fusion today: nobody quite understood it, but people supposed that it might do miracles. They knew that an electric current (produced by rubbing glass with a silk cloth or by touching a Leyden jar) could cause dizziness, a quickening of the heart rate and spots before the eyes. So maybe electricity was good for the health.

Graham, in fact, claimed electricity cured everything. At his London Temple of Health (and its elegant, Adam-designed premises), the elite took mud baths and shocks while surrounded by scantily dressed, nubile maidens—Emma was one for a while—and six-foot-tall bouncers. Graham knocked the London demimonde out cold with the star of his show: the amazing "magnetico-electrico-celestial" bed, guaranteed to fix infertility and almost anything else that ailed you. The line of the credulous ran all the way around the block.

Back in Naples, Sir William set Emma up in a plush villa, where she continued to assume attitudes. Not surprisingly, her posing turned out to be just the thing to catch the attention of a prominent navy type who had been at sea for too long (that and possibly the fact that, as he later noted, Emma never wore underwear). The sailor in question was the hero of the day, Horatio Nelson, whose charms were so renowned that when he sailed into Naples there was female fainting all around. He met Emma in 1798; quicker than you could say "Admiral of the Fleet," she was his mistress, and they were canoodling on the island of Malta, where the commissioner running the place was another old sea dog, Captain Alexander Ball, who had once saved Nelson's ship and life. But that is another story.

In those days, Malta was a strategic hot spot in the conflict between Napoleon and the rest of Europe. Malta gave

Nelson control of the Mediterranean sea-lanes and hence secured contact, through Egypt, with British India. Disrupting that contact was why Napoleon was after Malta, too. So the island was full of intrigue, as well as Russian, French and Turkish spies. There were also a few Americans (resting up after their war against Tripolitania) who had their own transatlantic reasons for undermining the Brits.

All this international hugger-mugger meant that when Ball was not entertaining Nelson and Emma, he was busy writing secret dispatches, night and day. And because Ball was better at navigation than prose, the dispatches were being edited, day and night, by his new rewrite man. This latter player was a passing opium addict and Romantic maven called Samuel Taylor Coleridge, who had arrived on the island in 1804, on the run from his wife and his habit.

Coleridge had journeyed to Malta to recover his health and financial well-being. After nearly two years, neither goal had been achieved, so the poet headed back to London via Rome, where he met and was painted by an American artist named Washington Allston. The two soon became close friends, and on a later visit to England, Allston introduced Coleridge to his protégé, a young American whose aim in life was to create one of the murals for the Capitol Rotunda in Washington, D.C. Alas, the job never came his way, although he did become the rage of New York's art world, founded the National Academy of Design, and made portraits of such movers and shakers as General Lafayette and De Witt Clinton. In 1829 this young painter headed once again for Europe, where he gradually came to realize that his future might lie elsewhere than on canvas.

On the return trip, in 1832, he came up with the idea that made him so much more famous than did his art that you probably still haven't figured out who this guy is: Samuel Morse. Six years of development later, Morse was only about the sixth inventor to produce a telegraph, but his version hit the jackpot for at least two reasons. One was the Morse Code. Nobody is totally sure that he didn't snitch it from his partner (and supplier of free hardware) Alfred Vail. Be that as it may, compared with the complicated, telegraph-and-printer models developed by the competition, Morse's technique was a breeze. It needed just a simple contact key (to send simple groups of five on-off signals), required only a single operator, worked over low-quality wire—and was cheap.

The other reason for Morse's success was also financial. Back then, railroads often ran both ways on single tracks (this saved money), and they frequently crashed (this lost money). Operators urgently had to find a way to instruct trains, coming in opposite directions, when to move and when to wait. From its first such use on the Erie Railroad in 1851, the telegraph did just that. But it also did a great deal more.

By the mid-1850s the Erie employed over 4,000, and the rail network was growing like topsy. In 1860 the company had around 30,000 miles of track, and things were threatening to go off the rails. The problem was that railroad companies served as many different enterprises all at once: shops, terminals, rail-track engineers, marshaling yards, warehouses and engineering units. Moreover, their materials, personnel and money were spread across thousands of miles. And the nature of the business meant that, from time to time, they had to make instant, system-wide decisions. If the companies were to survive, they needed a radically new kind of command-and-control organization.

Three engineers came up with the solution, making use of the rapid communications capabilities offered by the telegraph. Daniel McCallum (of the Erie), J. Edgar Thomson (of the Pennsylvania Railroad) and Albert Fink (of the Louisville & Nashville Railroad) devised the first business administration organization chart, the idea of line-and-staff management and divisional company structure, as well as the first true cost-per-ton-mile financial analysis. As a result, the railroads were soon able to handle routinely thousands of articles (passengers and freight) at high rates of turnover (getting them on and off trains) at low margins (cheap prices) on a huge scale (all across the continent).

By the 1870s railroad management techniques had helped establish another industry built on the frequent and regular delivery of goods. Like the railroads, these businesses operated on a large scale, at low margins and with high-volume transactions. Like the railroads, their staffs outnumbered the population of many cities. And like the railroads, their organization was departmental. Which is why they became known as "department" stores. Department stores proved a great hit and went on to generate the democracy of possessions that characterizes the modern industrial world.

So thanks in the first place to Wedgwood (whose factory is still operating), everybody today can buy his or her crockery. And if there is something wrong with it, get a free replacement, guaranteed. Which is a practice first introduced, in his London showrooms, by Wedgwood.

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.

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.

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.

Advertising correspondence should be addressed to Advertising Manager, Scientific American, 415 Madison Avenue, New York, NY 10017-1111, or fax (212) 754-1138.

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: SCAinquiry@aol.com.

ESSAY by Anne Eisenberg

Privacy and Data Collection on the Net

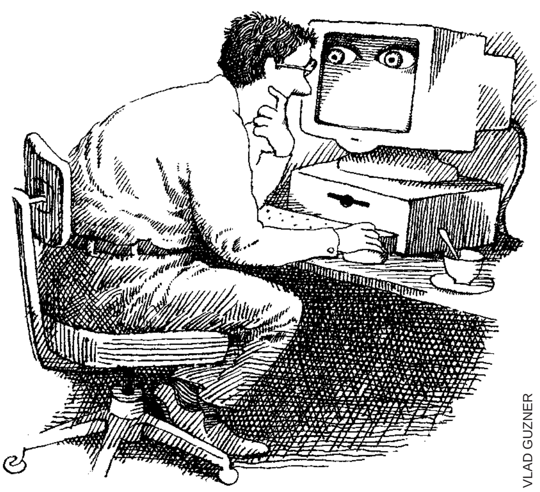
They are known as intelligent agents, those popular new computer programs that sift through the Internet for the things that most interest us. Overwhelmed by all that information on the information highway, we instruct our agents to do a little scouting on our behalf: sort our e-mail, search for the latest articles on knee surgery or shop the World Wide Web for the cheapest compact discs.

There is a small catch, though, to all this convenience. While we or our electronic alter egos are busily looking at Web sites, a good number of the owners and advertisers on these sites are looking right back. All those mouse clicks and keystrokes—which electronic sites we visit, how long we stay and where we go before and after—are not the ephemera they seem. Clickstreams, as they are called, enjoy a digital afterlife in commercial databases, where raw statistics about our on-line behavior are transformed into useful information and then warehoused for future application, sale or barter. These are known as the clickstreams that keep on giving—to advertisers, mass marketers and lucky venture capitalists.

The Net is turning out to be a bonanza for business folk, who will soon outnumber those quaint speakers of Unix who once dominated it. In one telling example, the two young men who originally developed Yahoo (http://www.yahoo.com) as a cheeky cyberspace index (they organized their favorite sites into categories like "New" and "Cool") have been succeeded by businesspeople who know that Yahoo will earn its keep not from its insouciant listings but from ads and data collection.

The Web is proving to be a good hunting ground for companies that want to appeal to highly specialized interest groups—for instance, consumers who start their search for four-wheel-drive vehicles or new software through Yahoo. And with every keystroke, advertisers and marketers will know where the customers are going and therefore where it makes sense to advertise.

Visits to Web sites—referred to as hits—are not the only statistics advertisers collect over the Net. They also receive valuable information from the questionnaires that users voluntarily complete. These forms include queries not only about the users themselves (name, income, e-mail and street address) but also concerning what they like and dislike about the product and even their responses to advertisements for the product. All the answers go into the hopper, where they increase the value of the corporate database of customer needs and preferences. To encourage response, some home pages offer rewards for personal information. One company will lend you an electronic agent to search its publications if you complete a short survey.



VLAD GUZNER

Data mining (a newly popular term once used mainly in the field of artificial intelligence) is the practice of creating programs that autonomously search data not for individuals but for group patterns. Wal-Mart, for example, mined its terabytes of sales data to respond more efficiently to daily changes in customer choice and sales volume. At one time, most companies used only their own, proprietary data for mining, but as computers have become more powerful and search algorithms have improved, many companies have begun buying outside demographic and marketing data to add to their own collections; they sometimes repackage what they have and sell that, too. Inevitably, the Internet will be mined.

Economists think we are headed for an "exponential growth in intrusiveness," as computers become faster, cheaper and smarter. To prevent fraud, to contain medical costs, to promote economies of all kinds, sophisticated databases will flourish and be accessed through networked computers. In principle, such information can help alert insurance companies to false claims or sound the alarm that a credit card has been stolen when a characteristic pattern of charges suddenly changes.

But it is one thing to mine data to find patterns, another to go the opposite way. Connecting records to the people who left them raises an unsettling vision: enter some names, and the computer will tell you all you ever wanted to know about them. If that capability gives pause, consider the credit reports that banks routinely buy from companies such as TRW when customers apply for a mortgage. The folders bulge with the bits and pieces of people's lives: what the applicants spend, where they spend it, the number of drivers in the household and how many accidents they've had.

There are still a few eccentric souls who gamely try to hold on to what lingering shreds of anonymity they possess. They never fill out questionnaires; they give their Social Security numbers only to their bank and to their broker. They encrypt their e-mail; they bypass the supermarket discount card that links identity to purchases; they pay cash for medical procedures they do not want known; and they wait patiently for e-cash to become a reality. Joining these hardy individualists are impressive privacy advocates such as EPIC (Electronic Privacy Information Center, http://www.epic.org) and Net groups like cypherpunks (cypherpunks-request@toad.com), which believe in untraceable communications and in the technology needed to achieve it.

These groups will face a vigorous challenge when computer telephony (the merging of computers and telephones, also known as Internet phones or I-phones) arrives. I-phones will truly be a telemarketer's dream come true.

ANNE EISENBERG (aeisenb@duke.poly.edu) can be found at Polytechnic University in Brooklyn, N.Y.