SCIENTIFIC AMERICAN

APRIL 1990
$2.95

Does a black hole lie at the center of our galaxy?

PCR: turning DNA needles into information haystacks.

The promise of light-water reactors.

Squandered asset: *fire destroys a tropical rain forest, clearing it for conversion to ranching or large-scale agriculture.*

Rocket Full Of

"It's just pure acceleration, like being in a capsule that's under tow from a space shuttle. In a fast corner taken briskly but within reasonable limits, the car just points and goes around, with no dramatics."

–Motor Trend

"A car that will redefine performance in the small sporty coupe class."

–AutoWeek

"If a driver's got any sense of dynamic delights, the back side of the garage door won't see much of this car."

–Car and Driver

"With performance 4WD and hybrid rear suspension, Talon can make a ham-handed driver look reasonably coordinated, and a good driver look great."

–AutoWeek

"This Eagle is ungodly good at flying on the ground."

–Car and Driver

"On low traction surfaces, the car can be steered with the throttle. The shifter is smooth and precise, one of the best we've ever tested. The pedals harmonize perfectly in shifting."

–Sports Car International

"At high speeds, this thing runs as straight and true as the irrefutable laws of inertia demand. This is a terrific, high-speed, over-the-road, all-weather car for any season."

–Motor Trend

"After wolfing down 750 miles in one day with the all-(wheel) drive Talon, we count ourselves as previously underprivileged."

–Car and Driver

7/70 Protects engine and powertrain for 7 years or 70,000 miles and against outerbody rust-through for 7 years or 100,000 miles. See limited warranty at dealer. Restrictions apply.

That's more trees to provide a home in which wildlife can flourish. And more trees to enhance the beautiful recreation areas for which America's forests are famous.

We are committed to renewing America's forests. Not just on Earth Day, but every day.

To help us manage the future of America's forests, mail this coupon or call 1-800-648-6699.

Name ____________________

Address ____________________

City ____________ State ______ Zip ______

1250 Connecticut Avenue, N.W., Washington, DC 20036

AMERICAN FOREST COUNCIL

MANAGING THE FUTURE OF AMERICA'S FORESTS.

...ORP. MEAD CORP. PACKAGING CORP. OF AMERICA POTLATCH SCOTT PAPER CO. SIMPSON TIMBER TEMPLE-INLAND UNION CAMP CORP. WEYERHAEUSER WILLAMETTE INDUSTRIES

SCIENTIFIC AMERICAN

April 1990 Volume 262 Number 4

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017. Second-class postage paid at New York, N.Y., and at additional mailing offices. Authorized as second-class mail by the Post Office Department, Ottawa, Canada, and for payment of postage in cash. Subscription rates: one year $27, two years $48, three years $66 (outside U.S. and possessions add $11 per year for postage). Subscription inquiries: U.S. only 800-333-1199; other 515-247-7631. Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51593.

DEPARTMENTS


TAURUS SHO

THE COVER photograph shows a Brazilian rain forest in flames. The burning of forests in order to clear land for ranching or farming is a major cause, along with excessive and wasteful timber operations, of the accelerating destruction of tropical forests. Both the conversion of forestland to other uses and the overlogging have been encouraged by government policies intended to deal with the severe economic pressures on many developing countries (see "Deforestation in the Tropics," by Robert Repetto, page 36).

THE ILLUSTRATIONS

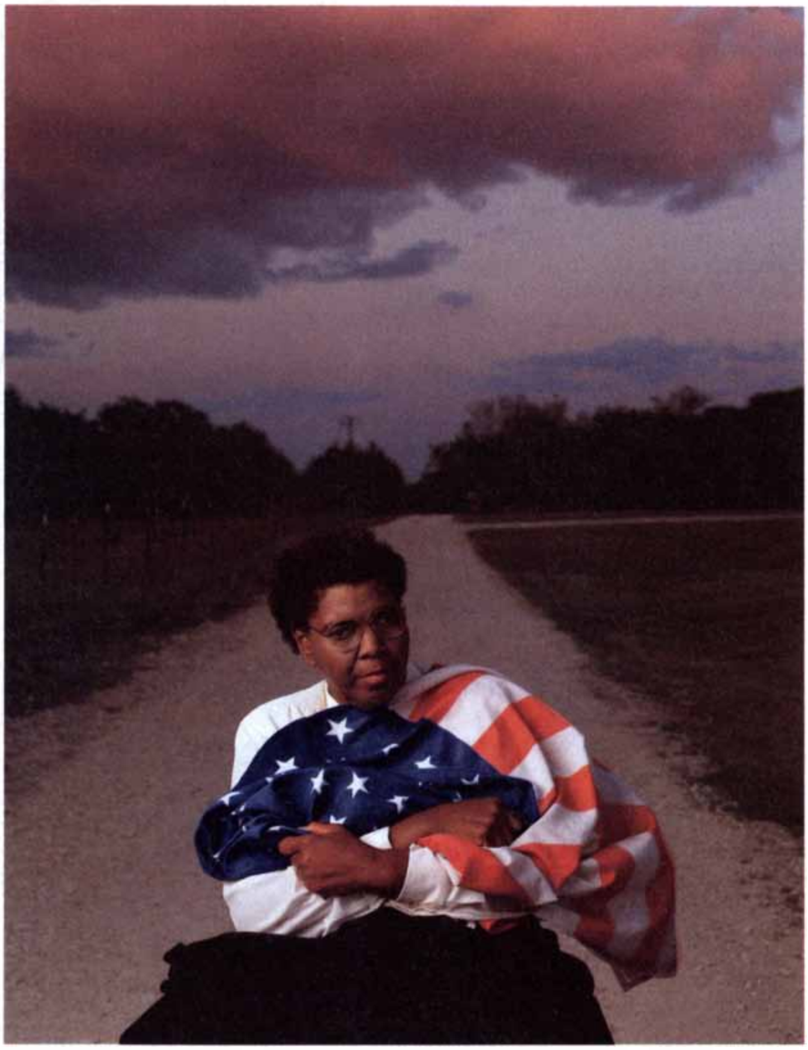
Cover photograph by Jacques Jangoux, Peter Arnold, Inc.

Page	Source
36–37	Jacques Jangoux, Peter Arnold, Inc.
38–39	Ian Worpole
40	Ian Worpole (*top*), Johnny Johnson (*bottom*)
41	Johnny Johnson
42	Edward Bell
47	Very Large Array (Mark Morris, Donald R. Chance and Farhad Yusef-Zadeh)
48	George V. Kelvin
49	Very Large Array (Alain Pedlar, K. R. Anantharamaiah, Ronald D. Ekers, W. Miller Goss, Jaqueline H. Van Gorkom, U. J. Schwarz and Jun-Hui Zhao)
50	National Optical Astronomy Observatories (Ian Gatley, Richard Joyce, Darren DePoy, Ronald Probst and Al Fowler)
51	Hat Creek Millimeter-Wave Interferometer, University of California (Rolf Güsten, Melvyn C. H. Wright)
52–53	George V. Kelvin
54	Johnny Johnson
57	George V. Kelvin
58–64	Michael Goodman
65	Cetus Corporation
67	Barrie Rokeach
68–70	Ian Worpole
71	Richard A. F. Grieve
72	Johnny Johnson
73	National Aeronautics and Space Administration
82–83	George Retseck
84	Johnny Johnson
85–88	George Retseck
91	Lester Lefkowitz
92–94	Tom Prentiss
95	Tom Prentiss (*left*), Robert B. Barlow, Jr. (*top right and bottom right*)
97	Stan Washburn, Anthro Photo (*top*), M. Shostak, Anthro Photo (*bottom*)
100	Patricia J. Wynne
101	Irven DeVore, Anthro Photo
102	Patricia J. Wynne
102D	John E. Yellen
104–105	Patricia J. Wynne
105	M. Shostak, Anthro Photo
107	Freer Gallery (*top*), micrographs by Pamela B. Vandiver
108	Edward Bell (*top*), Andrew Christie (*bottom*)
109	Freer Gallery (*left*), micrographs by Pamela B. Vandiver
111	(1) J. Paul Getty Museum; (2) The Metropolitan Museum of Art (*left*), Brooklyn Museum (*middle and bottom*); (3) Michael Nedzweski, Harvard University Art Museums; (4) Seattle Art Museum; (5, 6) Freer Gallery; (7) Chicago Art Institute; micrographs by Pamela B. Vandiver
112–113	Michael Goodman and Pamela B. Vandiver
114–117	Michael Goodman
119	Charles C. Place, The Image Bank

SCIENTIFIC AMERICAN®

Established 1845

SCIENTIFIC AMERICAN, INC.
415 Madison Avenue
New York, NY 10017
(212) 754-0550




AMERICAN EXPRESS
BARBARA JORDAN

LETTERS

To the Editors:

I thoroughly enjoyed your August issue but noticed a rather important omission in "The Age-of-the-Earth Debate," by Lawrence Badash [SCIENTIFIC AMERICAN, August, 1989]. The author states: "In 1899 John Joly of the University of Dublin devised the only truly new geological technique for measuring the earth's age." His method hinged on calculating the present-day salinity of the ocean and extrapolating backward to determine how long it would have taken for the oceans to develop the current concentration of salt.

It was an ingenious if flawed method, but it was certainly not new in 1899. Sir Edmond Halley (1656–1742), in addition to his many other scientific investigations, was interested in determining the age of the earth. He used the salinity of the oceans, as Joly did, to determine that the world was at least 100 million years old, and he did it several generations before Joly was born.

RICHARD SANTALESA

Flushing, N.Y.

To the Editors:

Lawrence Badash provided us with a very interesting and informative article, but there are a few points that should be clarified. He writes, "The Judeo-Christian tradition also combines the earth's and the universe's birth in a single event." That depends on how one reads Genesis. A literalist would read those verses one way; others, also with a high regard for Scripture, would read them differently. I do not believe that the Bible necessarily says the earth was created at the same time as the rest of the universe. This may put me and some of my scientist friends out of the Judeo-Christian tradition, but I think we sometimes need to rethink traditions in the light of later theological and scientific insights. Even some literalists allow, for the most part, that the "days" of creation may not have been 24-hour days.

Badash also states, "In 1654 John Lightfoot refined Archbishop Ussher's famous calculation of the moment of creation to an ultimate degree of precision: October 26, 4004 B.C.E., at nine o'clock in the morning in Mesopotamia, according to the Julian calendar." Are there any adherents to the Judeo-Christian tradition who accept this unfortunate and embarrassing conclusion? How many were there in 1654, for that matter? This aberration is a deviation from the Judeo-Christian tradition, and to imply otherwise shows the same insensitivity toward Christians that Salman Rushdie's *The Satanic Verses* shows toward Muslims. Authors of geology textbooks and others delight in setting up the work of these 17th-century theologians as a straw man to show how modern science contrasts with the biblical account. The phlogiston theory in chemistry [that combustible materials contain a mysterious element, phlogiston] and the neptunist theory in geology [that the earth's crust formed entirely by sedimentation and crystallization] probably attracted many more followers.

KENNETH J. VAN DELLEN

Department of Geology
Macomb Community College
Warren, Mich.

The author responds:

Readers of *Scientific American* will recognize that history, like science, expands, improves and is even reinterpreted with new information. Mr. Santalesa is correct in saying that I erred in not crediting Halley with the conception of the ocean-salinity method long before Joly reinvented it.

Professor Van Dellen's letter concerns only three sentences of background material that endeavored to encapsulate the biblical-chronology approach to dating the age of the earth. That is a topic worthy of an article in itself, but it was not the focus of the paper I wrote; it is also a topic guaranteed to irritate someone. Van Dellen argues for a broad, rich definition of the Judeo-Christian tradition, and he is, of course, correct.

He is, however, curiously sensitive to my use of Lightfoot's calculation of creation, calling it an aberration and, to my astonishment, likening me to Salman Rushdie. Lightfoot nevertheless provides a striking example of the "precision" of biblical chronology, and reference to him is historically legitimate. Johannes Kepler and Isaac Newton in the 17th century and Joseph Priestley in the 18th are others who made such calculations. To report this does not lessen our esteem for them or their accomplishments any more than does reporting that Kepler prepared astrological horoscopes and Newton spent a great deal of time pursuing alchemy. These were intellectually acceptable activities of learned people of that period.

LAWRENCE BADASH

Department of History of Science
University of California
at Santa Barbara

To the Editors:

As one who has been involved in the World Climate Research Program since its inception, I found "Global Climatic Change," by Richard A. Houghton and George M. Woodwell [SCIENTIFIC AMERICAN, April, 1989] accurate and informative. I was puzzled, however, by the illustrations on pages 42 and 43 showing the effects on Florida and Washington, D.C., of sea-level rises of both 4.6 and 7.6 meters. The captions associate such rises with the breakup of the West Antarctic Ice Sheet. Yet the article itself mentions sea level only in passing.

It is generally conceded that a new, warmer climate will result in some additional rise in sea level, but few contemporary estimates suggest a rise of more than a meter in the next century, and workers expect smaller increases.

Over the past decade the possibility that the West Antarctic Ice Sheet might break up has been subjected to a good deal of specialist attention, and the current consensus is that there is very little likelihood of it occurring, at least within the next century. There also seems to be a consensus that large changes in the volume of the great Antarctic and Greenland ice sheets are improbable on the same time scale.

ROBERT W. STEWART

Centre for Earth and Ocean Research
University of Victoria
Victoria, British Columbia

The authors respond:

Dr. Stewart is correct. Our emphasis in the article was on biotic interactions with climate, but the selection of figures shifted the apparent emphasis. Although unpleasant surprises are a possibility, given the global climatic tinkering now under way, we think a sudden change in sea level of from four to six meters or more is unlikely to occur in the short term we were discussing.

GEORGE M. WOODWELL

RICHARD A. HOUGHTON

Woods Hole Research Center
Woods Hole, Mass.

50 AND 100 YEARS AGO

APRIL, 1940: "Many attempts have been made to explain the present stalemate in the war between the Allies and Nazi Germany. No longer do people call it a phoney war; some, indeed, see in it a ray of hope for the future, a hope that the terrible bloodshed of 1914–1918 may never be repeated. It has been our feeling that the very deadliness and enormity of modern armaments have done more than all the cries of horror by the pacifists during the past decades, if not to stop war, at least to limit rigidly the bloodshed of a modern war between powerful nations."

"Thanks to Hollywood, most of us think of bananas as the easy manna of the drowsy tropics. In fact the banana is one of the most delicate and scrupulously cultivated crops on earth. Every stage requires hard work and tender care. The banana also requires perpetual haste, an average initial investment of about $400 per banana acre, and more man-hours of labor than any other principal harvest of the modern world. To the normal hazards of flood, drought, and hurricane there has been added of late a further and more terrifying threat to this 300,000,000-dollar-per-year industry: the sinister disease known as Sigatoka."

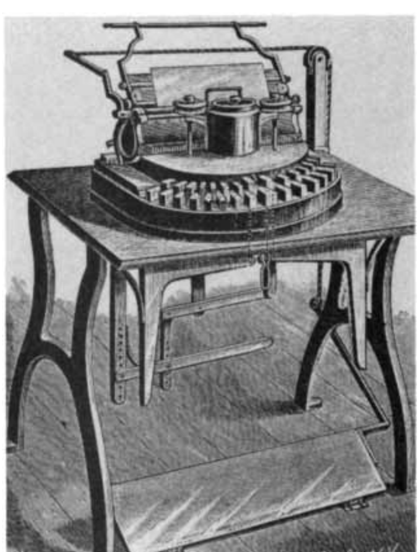
"Photographs that show oil wells afire and wasting thousands of barrels of oil will be dug out, three generations hence, from the musty files of the magazines of 1940 and reproduced again, so that posterity, to whom we shall transmit but a meager supply of petroleum, may curse our prodigality in recklessly wasting a resource which we know to be non-renewable. It is fairly safe to conclude that the world stores of petroleum will last only something like 75 years at the present rate of withdrawal."

"Food is now being given by vein to patients who cannot eat—because of excessive weakness at the time of operation or because the surgeon must cut some portion of the alimentary tract. Water, sugar, mineral salt, vitamin, and fat are five of the six nutrients known to be indispensable in human nutrition. The sixth is protein, until recently the main obstacle in the way of perfecting the administration of food by vein through an extended period. In a series of experiments, dogs were fed a diet free of protein for several weeks. Serum protein was observed to sink to lower and lower levels, and characteristic accumulation of water due to protein deficiency began to swell the tissues. But when amber-colored amino acid-dextrose solution was administered by vein, the amount of protein in the dogs' blood serum rose toward the normal level. Next came humans. In practically every instance, benefit came from amino acids by vein, and water-logging of tissues was diminished."

"Baking soda is now being made in what is said to be the world's cleanest chemical works. The plant has interior walls of glazed tile and functional panels of glass block instead of windows. Its entire frame was electronically welded, as was most of the equipment required for the new continuous process. Produced from start to finish in an enclosed maze of equipment, the bicarbonate of soda will not even come in contact with air inside the plant until it is ready for packaging."

SCIENTIFIC AMERICAN

APRIL, 1890: "The Italian astronomer Schiaparelli, after studying Mercury for seven years, makes the surprising assertion that Mercury, instead of turning on its axis once in twenty-four hours, turns only once in the course of a revolution around the sun. In other words, it always presents the same face toward the sun, behaving in this respect just as the moon does toward the earth. Moreover, Schiaparelli has discovered many marks upon the disk of the planet. A map shows that these marks strikingly resemble features discovered in recent years on Mars. They are elongated streaks running in various directions, and frequently presenting at their points of junction the appearance of an enlargement or knot. Similar streaks on Mars have been assumed to be long narrow seas or water courses."



A knee-operated typewriter

"On the first of this month a further test of the new smokeless gunpowder took place near Paris, under the auspices of Gen. Saussier, on which occasion some regiments of infantry riflemen and batteries of artillery took part. The firing from cannon and small arms was continuous, but without smoke, and the military people were greatly surprised. It is said it will now be necessary to change the uniforms of the French soldiery, to substitute subdued colors for their red pantaloons and other bright accouterments, as the absence of smoke renders the gay figures of the men very conspicuous when in action."

"Dr. Luderitz has recently made a number of observations on the destructive power of coffee upon various microbes. Anthrax bacilli were destroyed in three hours, anthrax spores in four weeks, cholera bacilli in four hours, and streptococcus in one day. It was, however, remarkable that good coffee and bad coffee produced precisely similar effects. He believes that, as previous observers have suggested, the antiseptic effect of coffee does not depend on the caffeine it contains, but on the empyreumatic oils developed by roasting."

"An attachment for typewriters, by means of which the shifting of the characters and the spacing may be effected without using the hands, has been patented by Messrs. Reuben Durrin and Rosecrans Sheldon. To throw the capital characters into printing position the operator presses a knee against one side lever, pressing the opposite side lever when it is desired to space, while to throw the figures into printing position the central lever is pressed by the knee, the latter lever being adjustable to any desired height."

SCIENCE AND THE CITIZEN

Slow Boat to Mars

Can NASA get us to the red planet?

Nine months ago President Bush endorsed a lunar base and a manned mission to Mars as national goals. So far, though, the response has been underwhelming. The public pulse has not quickened, and space experts have yet to agree on a strategy: a study by the National Aeronautics and Space Administration failed to gain the support of either the National Space Council or the National Research Council.

NASA's study, which was rushed out in 90 days and presented to a special panel convened by the space council last November, envisioned a manned lunar outpost by early in the next century. The first manned expedition to Mars would take place by 2020. The agency detailed a range of scientific questions that the "great exploration" should answer, including the origin of the moon and whether life ever existed on Mars.

But the space council (which Bush established last April) felt that NASA's plans were premature. The council's chairman, Vice President Dan Quayle, told NASA to go back and try again, this time soliciting help from outside experts. "The [space] council is facing up to a widespread concern that NASA is not up to the job," says John M. Logsdon, a space-policy analyst at George Washington University and member of the space council's panel. "We have underinvested for a number of years."

Logsdon points out that the space station now planned was not designed to support a Mars mission and that budget pressures have eliminated or delayed the development of key technologies. For example, the space station would not provide "artificial gravity" for astronauts, nor would it initially use a "closed loop" system for recycling water and oxygen.

NASA's scheme also fell flat with a committee established by the National Research Council. The committee's study, scheduled for publication in March, concludes that "the 'great exploration' underestimates engineering and operational challenges."

By some estimates the combined moon/Mars mission would take 30 years and cost at least $300 billion, an amount equal to five Apollo moon

missions. But NASA's work force is now at about half of its peak during the Apollo program, and some members of the research council's committee believe the agency would need at least to quadruple in size to tackle Mars. An administration proposal to boost NASA's budget to $15.1 billion in fiscal 1991, a 23 percent increase, is likely to be challenged.

The research council's committee is also concerned about the long-term health effects of space flight, especially the effects of prolonged weightlessness. After a year in orbit, Soviet cosmonauts lost bone mass and their hearts weakened; a Mars mission might remove astronauts from normal gravity for as long as three years. In addition, the psychological stresses of prolonged isolation have yet to be adequately assessed. Another potentially serious and largely unstudied problem is exposure to solar and galactic cosmic radiation.

A crucial question is whether the Mars mission would use nuclear propulsion or nuclear power once on the Martian surface. Although NASA is participating in the development of a space-based reactor that has been considered for use in the Strategic Defense Initiative, public opposition makes the use of nuclear power problematic. NASA's study assumed that chemical rockets and solar panels would provide power for the first Mars mission.

Nuclear propulsion could shorten the round-trip time to Mars to as little as seven months and has weight advantages, says Arnold D. Aldrich, who heads technology development at NASA. Nuclear power on the Martian surface is "probably necessary," he adds.

Another uncertainty for NASA is the nature of the mission. Is a journey to Mars to be the first stage of a long-term exploration, or is the goal simply to go and come back safely? "How you do a mission is a function of why you're doing it," says Louis Friedman, executive director of the Planetary Society, which has long advocated manned exploration of Mars.

To Friedman and Planetary Society co-founder Carl Sagan, the best way to get a Mars mission off the ground would be to team up with the Soviets. A Mars mission will certainly require a new heavy-lift launcher, for example. The Soviet Union has one available, as well as an automated shuttle. Friedman says the public will support a Mars mission only if it is undertaken as a cooperative venture.

Yet some officials question whether the U.S. would ever enter a collaboration in which the Soviets control key technologies. Others doubt the wisdom of sharing the U.S.'s best radar, propulsion and control systems with the Soviet Union.

The idea of a manned Mars mission has captured the imagination of workers in U.S. laboratories. A team at the Lawrence Livermore National Laboratory, for one, has proposed space vehicles built from modules that would be inflated in space. That scheme, it claims, could put men on Mars for $10 billion. Many engineers, however, dismiss the cost as unrealistic. Parts of the Livermore plan are "unworkable," adds John McElroy, an aerospace engineer at Texas A&M University.

Still, NASA has invited interested parties to throw in their suggestions. And in spite of reservations about NASA's current plans, many aerospace experts think a revitalized agency could pull off the Mars mission—but only if NASA receives stronger political direction and support. —*Tim Beardsley*

Right to Lie?

Studies disprove claims about abortion's dangers

Opponents of abortion often profess deep concern for the health of women. In literature and speeches, they suggest that abortion poses greater physical and psychological risks than childbirth. John C. Willke, president of the National Right to Life Committee, recently reiterated these claims to SCIENTIFIC


Liqueur
Grand Marnier
Liqueur

AMERICAN. He said his belief in them rests on his experience as an obstetrician and on anecdotal reports, primarily from others in the antiabortion movement. "There are no good, hard statistics on this," he said, "but there are none on the other side either."

In fact, government statistics clearly show legal abortion to be safer than giving birth. Data compiled by the National Center for Health Statistics from 1981 to 1985 indicate that abortion was 11 times less likely than childbirth to lead to a woman's death. Researchers at the Centers for Disease Control reported in 1982 that women undergoing abortions are 100 times less likely to have complications requiring major abdominal surgery than women bearing children.

The question of psychological harm has been more difficult to assess. Indeed, last year C. Everett Koop, the former surgeon general, judged past studies to be inconclusive. Now a group at the Johns Hopkins School of Hygiene and Public Health has provided "hard statistics" on this issue.

The researchers studied 334 black urban teenagers who entered clinics in Baltimore to be tested for pregnancy. The homogeneity of the group reduced the chances that the study's results would be skewed by other variables, according to Laurie Schwab Zabin, who did the study with Marilyn B. Hirsch. The teenagers were initially interviewed before they or the researchers knew the results of the pregnancy tests. They were then divided into three groups—those who bore a child, those who had an abortion and those who were not pregnant—and tracked for two years.

The investigators report in *Family Planning Perspectives* that those who chose abortion were less likely to undergo adverse psychological episodes (as measured by three separate psychological tests) than either those who bore children or those who had not been pregnant; they were also more likely to remain in school and less likely to become pregnant again.

"The right-to-lifers have been saying that it is terrible to let young women go through abortions," Zabin says. "Our study shows we should lay that ghost to rest. Not only is abortion a medically safe procedure; we now know it is psychologically safe."

Poor urban teenagers represent an important minority of all those who elect to have abortions, Zabin notes. They are the most likely to be affected by laws proposed in a number of states that would limit the use of public funds for abortion, require minors to gain consent from parents or otherwise restrict abortion's practice.

Moreover, such laws could help give more substance to Willke's claims if they hinder women from obtaining timely abortions. Numerous studies show that the risks of medical complications rise significantly when abortions are performed after 16 weeks of pregnancy. *—John Horgan*

Acid Test

A mammoth assessment fails to find all the answers

As Congress wrestles with legislation intended to reduce atmospheric emissions of sulfur dioxide, the principal cause of acid rain, a huge scientific assessment has confirmed that acid rain does indeed have harmful effects. How harmful?

After working for 10 years and spending more than $600 million, the National Acid Precipitation Assessment Program in February released draft summaries of 28 scientific studies on the subject; final versions should be published within months, after comments have been digested. Although intended to be definitive statements, the studies seem unlikely to settle arguments about either the extent of lake acidification or the effects of acid rain on forests.

NAPAP, a federally funded body created by Congress, is no stranger to controversy. In 1987, when it released an interim assessment, environmental groups charged that the assessment distorted the research conclusions. Now the 28 draft summaries, presented at a conference in Hilton Head Island, S.C., appear to be headed for the same fate.

The draft summaries confirm that lakes in the East have become acidified because of emissions of sulfur dioxide from power plants and other sources. Lakes in the Adirondacks, in the Upper Peninsula of Michigan and in Florida have been particularly affected; no lakes in the mountainous West, where sulfur dioxide emissions are far lower, were affected. "Most people accept that some damage has occurred," according to NAPAP associate director Patricia M. Irving. But the studies suggest that during the 1980's acidification was less severe than some predictions: only 4.2 percent of a representative sample of lakes had lost all ability to neutralize acid inputs. The assessments conclude that acid rain has not affected crops and confirm that liming of waters can reverse acidification. They note, however, that the fauna and flora are unlikely to be the same after acidification and liming.

More controversial is the study on forest health, which concludes that "the vast majority of forests in the United States and Canada are not affected by decline." Of those forests that *have* been shown to be in decline, according to the study, only in the case of red spruce at high elevations in the northern Appalachians is there reason to hold acid precipitation partly responsible. With regard to other instances of tree decline and growth reduction, the draft says merely that "natural stresses are important factors." The study identifies regional ozone as the pollutant of greatest concern for its effect on forests.

Kristiina Vogt of Yale University, whom NAPAP asked to review the forest-health summary, says the Canadian data in the summary "do not reflect the forests of Canada and taken out of context would suggest that acid pollution is not affecting Canadian forests"; she also expresses concern that "personal biases are strongly emphasized in some sections."

Canadian scientists at the conference appeared to agree with Vogt, as did some U.S. researchers. "A number of other species [besides red spruce] are also showing more damage than would be expected from the natural stresses that exist in any ecosystem," says Richard M. Klein of the University of Vermont. Elizabeth Dowdeswell of Environment Canada, a government agency, has charged in a letter to NAPAP director James R. Mahoney that the summaries underestimate damage to lakes in Canada and fail to take account of Canadian research indicating that acid precipitation is one probable cause of forest decline.

NAPAP acknowledges that sugar maples are in decline in Canada, but Canadian officials say there is evidence that other species are in decline as well. Moreover, the Canadians take issue with the report's conclusions that "there is no evidence that acidic deposition is an important factor in sugar maple decline." Says Thomas G. Brydges of Environment Canada, "Our scientists are pretty convinced that the cause of forest decline is multiple stresses and that one of those stresses is acid rain." Canadian researchers think acid rain may be affecting trees by causing changes in soil chemistry.

The dissenters may succeed in forcing changes in the summaries, most of which are still under peer review. As for NAPAP, it now moves on to compile an integrated report on strate-

gies for dealing with acid rain, which should be ready in September. But if history is any guide, NAPAP's final word will not be *the* final word on the subject. —*T.M.B.*

Science behind the Wall

East Germany's liberated scientists regroup

American scientists frazzled by the incessant quest for grants and worried about the inconstancy of government support should take comfort from contemplating conditions on the other side of the rapidly crumbling Berlin Wall. There, under the aegis of the Academy of Sciences of the German Democratic Republic, an atmosphere has prevailed that must have suffocated all but the hardiest scientific spirit.

Certainly the academy's dilapidated offices in East Berlin belie its reputation as a former home to such luminaries as Gustav Hertz, Max Planck and Albert Einstein. The offices also symbolize the state of basic research. During more than four decades of the dictatorship of the proletariat, the academy has "done much harm to fundamental science," observes Jens Reich, a researcher at one of the academy's component institutes, the Central Institute for Molecular Biology.

Reich rose to national prominence as a founding member of New Forum, the alliance that played a pivotal role in East Germany's revolution. By his account, political domination of research in Erich Honecker's republic reached a pitch barely conceivable to Western scientists. Reich says that only members of the Communist party could aspire to become institute directors and that even leaders of research groups had to survive vetting by the secret police. (Reich failed the vetting and was eventually demoted from research-group leader to postdoctoral status; he thinks he fell into disfavor because he maintained unauthorized contacts with Western researchers.)

Also damaging to scientific collegiality, Reich says, was the academy's practice of granting foreign-travel privileges only to older investigators who had demonstrated their political dependability. Even the agendas for visits by foreigners were rigidly controlled, he asserts. And academy staff members, Reich notes, helped the secret police monitor contacts.

East German scientists have faced other difficulties. According to Reich, the abysmal research performance of East German industry obliged the academy's institutes to divert much of their effort to noninnovative science, such as making biomedical products. The industrial sector has left another challenge for the technical community: some of the worst environmental pollution in Europe.

Scientists have lost no time in exploiting their new freedom. According to Tom A. Rapoport, a cell biologist at the molecular biology institute, workers at the academy's institutes have in recent weeks stripped the president of the academy of much of his control over research. New councils have been created to advise institute directors. Scientists have also proposed controversial revisions to the academy's constitution: Rapoport expects changes in months or weeks.

Whatever precise course the academy and the country choose, Rapoport is confident that East German scientists, now free to decide their own agendas, will want to increase cooperation with Western scientists. There are already indications that West German grant-giving agencies might provide research support.

Other benefits of East-West exchanges might be forthcoming. Reich, whose expertise is in informatics, visited the National Library of Medicine in Bethesda, Md., earlier this year, his first visit to an English-speaking country ("a dream come true"). The visit convinced him that East German information scientists should contribute to the international effort to map and sequence the human genome. —*T.M.B.*

PHYSICAL SCIENCES

Zap!

Coil guns offer to orbit small cargoes on a regular schedule

Decades ago, in the heyday of pulp science fiction, there was the mass driver: an electromagnetic accelerator that shot vehicles into outer space without benefit of rocket engines. Futurist Gerard K. O'Neill envisioned mass drivers on the moon slinging supplies to space colonies. In the early 1980's, with the advent of the Strategic Defense Initiative, or "Star Wars," the mass driver begot the rail gun: a supposedly cheap, simple version of electromagnetic propulsion that could be used to fire interceptors at rising Soviet missiles. As strategic peace breaks out and Star Wars continues its long slide, the rail gun has begotten a new generation of mass driver, poised to deliver small packages to orbit for a fraction of the cost imposed by rocket technology.

A group of researchers from Sandia National Laboratories, best known for such projects as the maneuvering nuclear warhead, has proposed the construction of an 800-meter-long electromagnetic accelerator at Barking Sands on the Hawaiian island of Kauai. The site, says M. Bill Cowan of Sandia, is perfectly situated for delivering small satellites to polar orbit: the ground track of the launch vehicle would pass largely over ocean, thus alleviating safety concerns. In Sandia's design—a coil gun, not a rail gun—electromagnets would induce a current in a 450-kilogram aluminum armature and accelerate it through a series of coils. The armature would carry with it a 400-kilogram projectile, which would depart the coil gun's barrel at about six kilometers per second. Once above the atmosphere, the projectile would fire a rocket to kick about 100 kilograms of payload into orbit. In contrast, big boosters such as the Titan place only a small percentage of the launch mass in orbit.

What has changed since the days when mass drivers were pulp science fiction? Power storage, for one: the capacitors that will provide enormous bursts of power for the fraction of a second it takes the coil gun to fire have improved more than tenfold in the past five years and could well do so again, Cowan says. For another, Star Wars experiments with coil guns, rail guns and light-gas guns have advanced the technology in many areas. Before the SDI-funded boom in electromagnetic propulsion, the largest experimental systems built were a factor of a million short of what was needed to put payloads in orbit; current systems will merely have to be scaled up by a factor of 1,000.

Certain problems still need to be worked out, Cowan notes. The satellites that the coil gun fires will have to be redesigned to survive acceleration forces more than 1,000 times that of gravity. Although proximity fuzes and other simple circuits have been built to withstand more than 100,000 g's, the standard for modern communications satellites is closer to 10 g's.

Engineers will also have to make progress in aerodynamic shields: current nose cones are designed for ultrahigh speeds only in very thin air, not at sea level. And designers will have to cope with the superheated plasma that forms whenever the armature

15 ties an hour isn't exactly mass production...

But then, we're talking handsewn ties— the only kind Lands' End sells.

There's this quiet little place down in New Orleans, where everybody's busy, but nobody hurries, and the sound of machinery doesn't set your teeth on edge. (Oh, now and then, the sisss of an iron pressing cloth might get your attention, but that's about it, really.)

It's the home of a wonderful little handsewn tie company precisely suited to our homespun heart.

Here gifted needleworkers make the most of finely woven fabrics entrusted to their care from around the world. Fine natural fiber fabrics from Scotland, Italy, China, wherever. In Spring, they're likely to be cottons, silks, and linens. In Fall, the mix accommodates wools.

Once in house, patience works magic.

These elegant fabrics are carefully cut on the bias for maximum resiliency. Tailors with a good eye and sensitive hands mate them to fine brushed interlinings. Handpressing assures edges with that unmistakably proper roll. Then the ties are slip-stitched by hand, and even hand-bartacked at both the wide and narrow ends. This allows the tie to "give" so it drapes more naturally, and holds its shape.

And note this, please: the "keepers" on which we're proud to put our name are firmly stitched on—at all four corners—so they don't pull loose from the back of the tie after the second wearing.

Isn't it time you met our handsewn ties?

From the Crescent City to Dodgeville, Wisconsin, these marvelous handsewn ties arrive in a Mardi Gras parade of styles and colors. We feature them in our catalogs at prices unheard of anywhere. And men in the know are buying them—literally—by the dozens. (We even offer most of our styles in both Regular and Long, to accommodate the more generously-sized of you.)

All handsewn, of course. And all GUARANTEED. PERIOD.® Like every other item in our catalog. Surely you must at least see them, which you can do simply by calling us toll-free at 1-800-356-4444. Or, by sending us the coupon below.

Then, catalog in hand, you can also scan our array of dress clothing to wear with your freshly-minted handsewn tie wardrobe. And shop our assortment of colorful casual clothing and soft luggage, too. All down to earth values because we do business directly, bypassing middlemen.

Let us hear from you, by phone or by mail. Not just so we can sell you a tie. But so we can start building the ties that add up to a permanent relationship.

LANDS' END
DIRECT MERCHANTS

Please send free catalog.
Lands' End Dept. Q-41
Dodgeville, WI 53595

Name ______________________

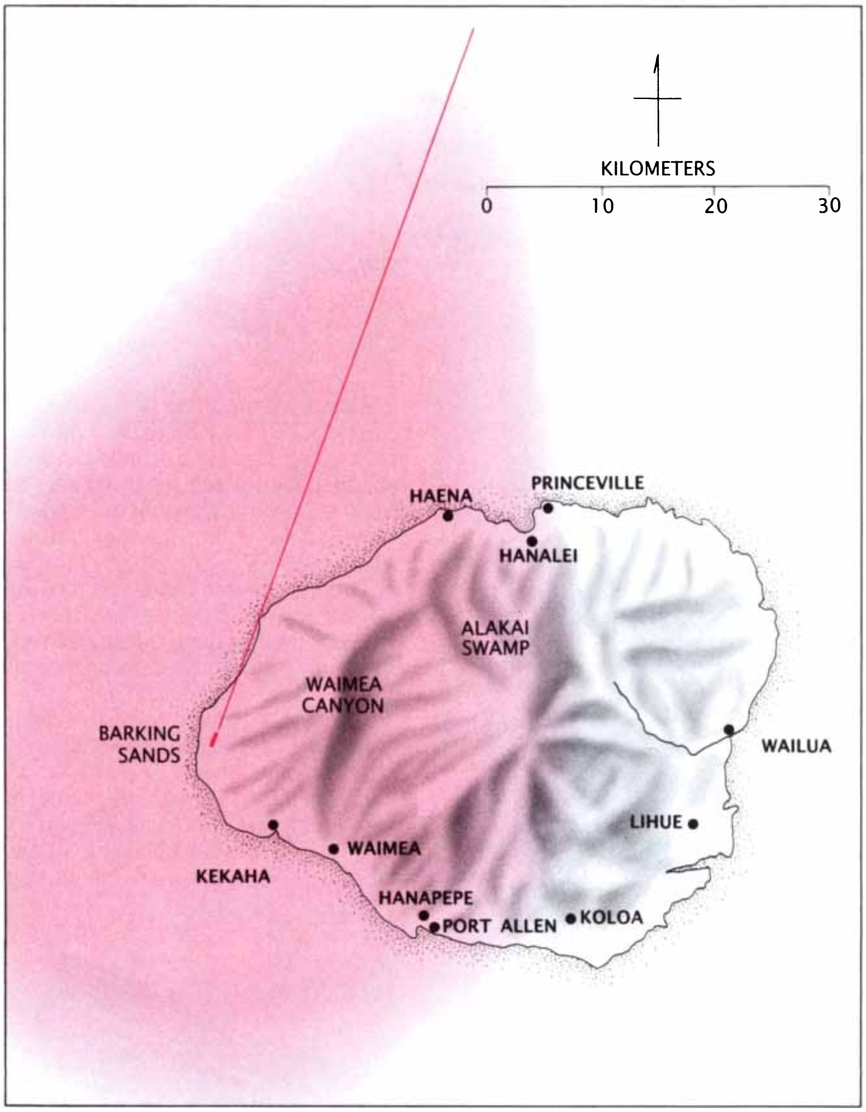
Address ______________________

City ______________________

State ____________ Zip ____________

Or call Toll-free:
1-800-356-4444

Proponents of the electromagnetic launcher want to build a prototype on a Hawaiian island



BARKING SANDS on the island of Kauai is the proposed site for the electromagnetic launcher. In addition to the 800-meter gun, the launch facility* (red) *would also require as much electricity as the island's total present generating capacity. Sonic booms from the launches, occurring as often as every 10 minutes, could be audible as far as 30 kilometers from the line of flight (as indicated by shaded area). The island location has a precedent: Project HARP, a program that fired shells as high as 180 kilometers during the 1960's, sited its gun in Barbados.

bumps against the barrel of the gun at six kilometers per second.

If a couple of years of funding show that such obstacles can be overcome, Cowan says, Kauai could have an operating coil gun by the turn of the century. The estimated $1.3-billion price tag is cheap compared with the $20 billion or more projected for development of successors to the space shuttle. Both could loft roughly the same amount of material into space during their operating lives, although the coil gun would do so in much smaller pieces. The coil gun could put 10,000 small payloads in orbit at a cost per pound significantly lower than that of the cheapest big rockets, according to Miles R. Palmer, a physicist at defense contractor Science Applications International Corporation.

That brings up the question of why anyone might want to put 10,000 small payloads in orbit. Palmer contends that a mass driver could send supplies—such as air, water, food and other acceleration-hardy items—into orbit to provision a mission to Mars or resupply a space station far more cheaply than conventional rockets could. He also envisions a global cellular-telephone network, based on thousands of satellites in low earth orbit, which could attract 20 to 30 percent of worldwide spending for telecommunications by the turn of the century. Perhaps Dick Tracy's two-way wrist radio will join the mass driver in the ranks of pulp visions come to life. *—Paul Wallich*

Gone in a Flash

What happens in the sky when nobody is looking?

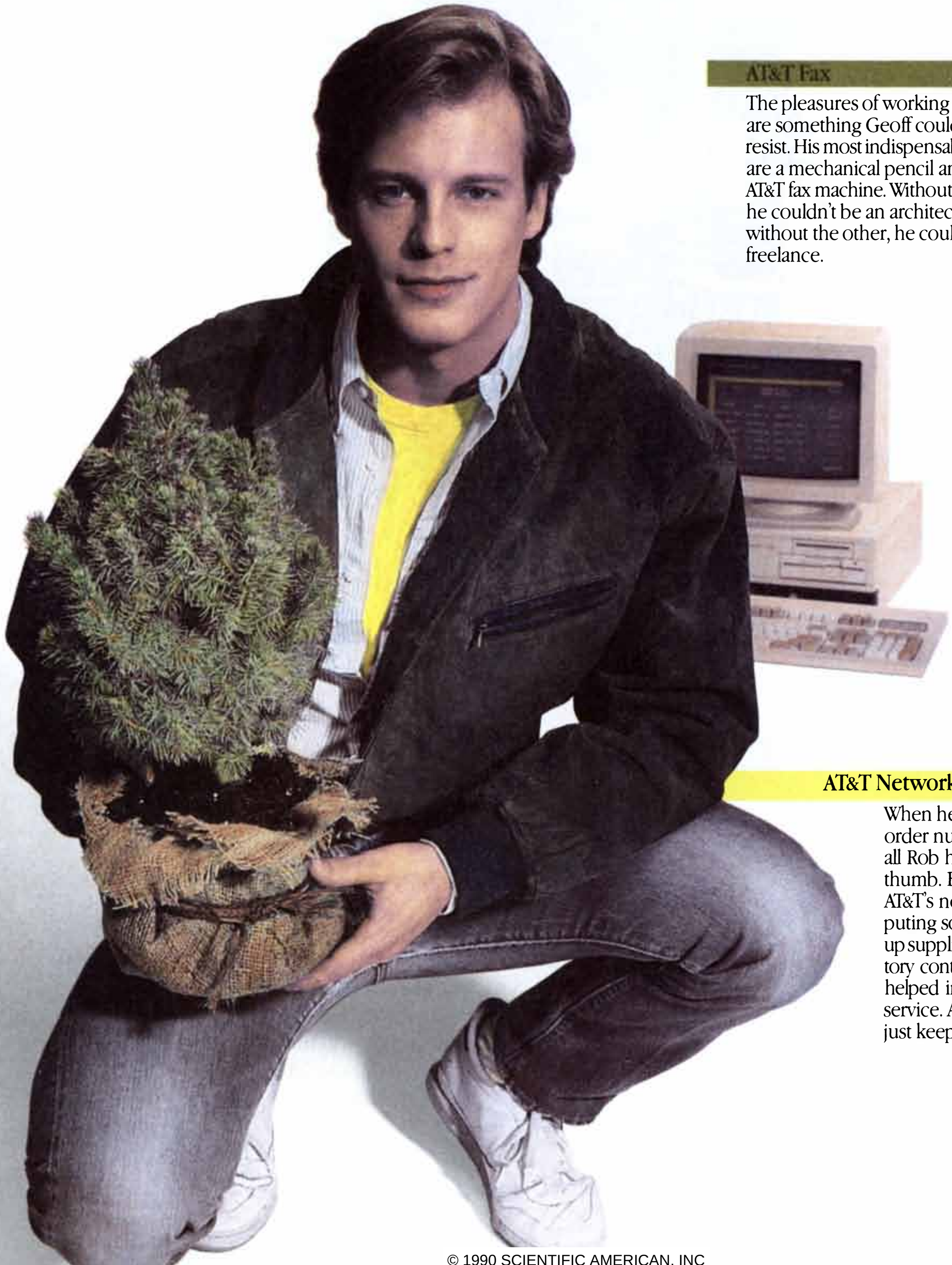
Sometimes an astronomer's job is not unlike that of a teacher with unruly students. Everything appears nice and orderly while the teacher faces the class; the noteworthy things happen behind the teacher's back. Optical astronomers, who mostly study long-duration exposures of a small patch of the sky, have a similar problem. If a faint star appeared—or disappeared—for a brief moment, would anybody notice?

Scott D. Barthelmy and three coworkers at the National Aeronautics and Space Administration's Goddard Space Flight Center hope to answer that question using their new Rapidly Moving Telescope (RMT), which can point within one second toward any patch of the sky and then observe objects in that patch as faint as the 13th magnitude at a resolution of approximately one arc second (1/3,600 of a degree). The RMT has undergone preliminary testing at Goddard and is now being installed at the Kitt Peak National Observatory in Tucson, Ariz. There it will be coupled with the Explosive Transient Camera (ETC), an instrument designed at the Massachusetts Institute of Technology that automatically will scan wide swaths of the heavens for sudden flashes; when a flash is spotted, the RMT will spin around to take a detailed look at the region from which the flash arrived.

The RMT and ETC have been built to look for optical counterparts to gamma-ray bursts, mysterious pulses of energetic radiation that typically last from one to 10 seconds [see "Gamma-Ray Bursters," by Bradley E. Schaefer; SCIENTIFIC AMERICAN, February, 1985]. Bursts are thought to be associated with violent events occurring near dense stellar corpses called neutron stars, but little is known about them.

Once upon a time, a phone was just a phone. And AT&T was just the phone company.

Today AT&T has spawned a whole new generation of computer and communications technologies. Everything you need to get things done in today's world. And tomorrow's.

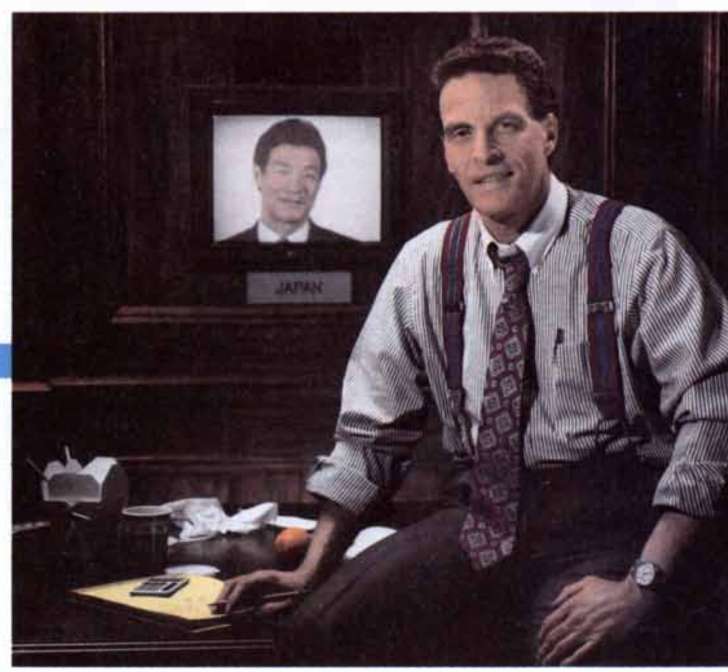


AT&T Fax

The pleasures of working at home are something Geoff could not resist. His most indispensable items are a mechanical pencil and his AT&T fax machine. Without the one, he couldn't be an architect, and without the other, he couldn't freelance.

AT&T Networked Computing

When he started his mail-order nursery business, all Rob had was a green thumb. But thanks to AT&T's networked computing solution, linking up suppliers with his inventory control system has helped improve customer service. And his business just keeps on growing.



The few optical images of bursters that do exist appeared by chance on long-duration exposures and do not tell much about these objects. The RMT may at last determine the source and nature of gamma-ray bursts.

Other elusive celestial objects might reveal themselves with more subtle blips of light. Distant comets, for instance, may occasionally eclipse a star. As many as 200 trillion such comets are thought to orbit the sun in a huge cloud (the Oort cloud) at from 100 to 20,000 times the distance between the sun and the earth. Proving the existence of these comets would greatly illuminate the still murky origins of the sun and planets.

To detect the comets, Tim S. Axelrod and Charles R. Alcock of the Lawrence Livermore National Laboratory and two colleagues are developing an automated telescope that will monitor the intensities of about 1,000 faint background stars 10 times per second. Eclipses of stars by comets would last only a few tenths of a second and would be distributed all over the sky, but they should be fairly frequent. Axelrod and Alcock expect that their telescope should detect roughly 100 eclipses per year; the actual number will constitute an important test of the conditions in the unexplored outer reaches of the solar system.

In an even more ambitious project, Axelrod and Alcock plan to hunt for distant planets and brown dwarfs, which are intermediate in mass between planets and stars, by exploiting the phenomenon called gravitational lensing. Einstein's theory of general relativity holds that gravity can bend light; if a planet outside the solar system passes in front of an even more distant star, its gravity can act as a lens, amplifying the brightness of the star. Depending on the geometry of the event, the lensing could last anywhere from minutes to weeks. To improve the odds of observing a lensing event, Alcock and Axelrod propose to build a dedicated telescope that would monitor brightness fluctuations of about a million stars.

The first gravitational-lensing event from a small object may in fact already have been observed. Michael J. Irwin of the University of Cambridge and his colleagues have been monitoring "Huchra's lens," a quasar whose image is split into four components by the gravity of an intervening galaxy aligned nearly perfectly between the earth and the quasar. Over the course of a year one of the components of the image brightened about 60 percent and then returned to its original brightness. If there had been any major change in either the quasar or the galaxy, all four components should have been affected. Irwin therefore suspects that a small object within the galaxy passed exactly along the line of sight between the earth and one of the image components, briefly acting as a secondary, amplifying lens.

Daniel Nadeau of the University of Montreal and his collaborators calculated the mass of Irwin's suspected interloper and found it to be about eight times the mass of Jupiter, far too lightweight to be a star. Like other reported discoveries of planets and brown dwarfs around other stars, this one is based on maddeningly vague information that is subject to other interpretations. But it does suggest that much can be learned by taking a closer look at some of the subtle tricks the cosmos plays. *—Corey S. Powell*

BIOLOGICAL SCIENCES

All about Eve

Did the mother of us all live less than 200,000 years ago?

A controversial theory tracing part of the genetic heritage of every human being to a woman who lived in Africa less than 200,000 years ago has drawn new support—and new opposition. The support comes from an ambitious worldwide genetic survey; the opposition is based on fossil evidence challenging the accuracy of the molecular clock on which the theory is based.

Molecular clocks are derived from mutations accumulated in several lines known to have diverged from a common ancestral stock. The molecular evolutionist calibrates the clock by estimating the long-term mutation rate in cases where the time of divergence is known from the fossil record; the clock can then be applied to hitherto undated divergences.

In the mid-1980's a group led by Allan C. Wilson of the University of California at Berkeley based such a clock on the DNA encoding the genes of mitochondria, the organelles in the cell's cytoplasm that transduce energy. Because the mitochondrial genes are inherited only from the mother and therefore are not scrambled by sex, as the more numerous genes of the nucleus are, each occurrence of an all-male generation in a family puts an end to a mitochondrial lineage. Such extinctions eventually winnowed out the heritage of all but one woman, the so-called Mitochondrial Eve.

Wilson's group estimates the mitochondrial DNA (mtDNA) mutation rate at between 2 and 4 percent per million years, from which they date Eve to between 100,000 and 200,000 years ago. The workers constructed a tree of genetic differences among populations that represents non-Africans as twigs on an African branch, indicating that Eve lived in Africa. This determination has now been confirmed in a much wider genetic survey done by Linda Vigilant and other students of Wilson. "We think these results clinch the African origin of mtDNA," Wilson says.

Some paleontologists take issue with the assumptions behind the molecular clock, in part because it suggests that Eve's tribe spread recently and quickly throughout the world without interbreeding with indigenous hominids. In a paper presented at the February meeting of the American Association for the Advancement of Science, Milford H. Wolpoff of the University of Michigan contends such total replacement has no parallel in historical examples of conquest, no matter how great the technological superiority of the conquerors. It would imply, he says, that the African upstarts constituted a new species—modern *Homo sapiens*. That is impossible, he maintains, because fossil evidence shows modern human racial distinctions had already become evident hundreds of thousands of years earlier.

Wolpoff and two colleagues in Australia and China came to their conclusions by studying dental and facial traits of human fossils in Europe and Asia. Wolpoff says they analyzed only traits that could not have evolved in adaptation to local environments and that must therefore have been produced by genetic drift. Wilson dismisses this argument, suggesting that obscure mechanisms of natural selection might have caused the newcomers' resemblance to their unrelated predecessors. He is also critical of Wolpoff's data. "He is dealing with a handful of traits, and we have nearly 100," Wilson says.

Many paleontologists have resented the encroachment of molecular biologists on their bailiwick. Wilson and his co-worker Vincent M. Sarich crossed that border in 1967, when their original molecular clock—based on the proteins albumin and hemoglobin—implied that humans, chimpanzees and gorillas had diverged from a common ancestor only about three to five million years ago, more than 10 mil-

lion years later than had previously been supposed. The bold finding forced paleontologists to reexamine their bones and revise their date of divergence considerably.

Wolpoff discounts that success. "Even a broken clock is right twice a day," he says. "The molecular biologists were going to straighten out our fossils for us, but this time it's been turned on its head: our fossils show that the Eve hypothesis does not hold up." *—Philip E. Ross*

Green Thumbs

Doing agricultural genetics in the marketplace

One of the great hopes for biotechnology is that it may provide a quick way of introducing new genes into plants to increase their nutritional value or tolerance of stress and disease. Some, such as tomato and soybean plants, can already be manipulated genetically by such means. Tinkering with cereal crops has been much more difficult, but now BioTechnica International, Inc., of Cambridge, Mass., claims to have introduced a foreign gene into maize that the plants express as protein and pass on to their offspring.

If the claim can be substantiated, it could represent an important advance, opening the way to strains of maize having novel properties. Investigators note, however, that the plant-engineering field is replete with claims that turn out to be false or unrepeatable, and some criticize BioTechnica for premature brouhaha. "I'm distressed, and I don't think that sort of thing should be encouraged," says Susan R. Wessler of the University of Georgia. Many workers say they will not take BioTechnica's claim seriously until the company publishes full details of its work. For now, BioTechnica refuses to provide more information.

Wessler and her colleagues have also transformed maize, but the gene they introduced was not passed on to the plant's offspring. By now other groups may have managed to achieve fertility. Ray Shillito of Ciba-Geigy, which has so far reported only transforming maize, says coyly, "I don't think there should be any problem" in producing fertile offspring from transformed plants.

The trick in engineering a foreign gene into a plant is to introduce the gene into meristematic cells, which are capable of dividing. If these cells can then be stimulated to grow into a new plant, then most or all of the adult plant's cells will incorporate the new gene. Investigators familiar with BioTechnica's work think the company adopted a technique that has been widely used in recent years: a "gun" that literally shoots microscopic particles of tungsten coated with DNA through the cellulose wall of plant cells. Some of the cells take up the DNA and express it as if it were one of the plant's own genes.

Researchers say if BioTechnica's results, said to have been achieved by its agricultural-research subsidiary in Minnesota, are reproducible and are achievable in various strains of maize, the work will certainly represent a breakthrough. Seed companies might then engineer cereals that are resistant to herbicides, for example, or that contain "built-in" pesticide. Similar techniques applied to rice, a staple in many developing countries, might increase yields where food is scarce.

Even if BioTechnica's claim holds up, it may be a while before useful strains appear. Wessler points out that workers have been able to introduce new genes into dicotyledons (most plants other than cereals) for several years, but the technique has yet to yield commercial products such as improved crops. The reason, she says, is that few genes for significant traits have yet been identified and isolated. Progress will be no faster with cereals, Wessler predicts.

BioTechnica promises to publish complete details within a few months, once it has exploited its commercial lead. *—T.M.B.*

Net Result: Folded Protein

A neural network deciphers the structure of protein

In molecular biology, function follows form. The biological activity of proteins—how they assemble into cellular structures, bind to receptors, catalyze metabolic reactions—depends largely on their three-dimensional structure. It is now a straightforward task to determine the sequence of the amino acids that form a protein chain. But how this chain folds up into a functioning molecule remains a cipher. Recent work by several groups suggests one way to solve the problem: neural networks—a computational strategy based on many simple elements interlinked in a specific arrangement. The links are modified according to rules that enable the network eventually to generate solutions to certain problems. (The network can be built in hardware or implemented entirely in software.)

The structure of a protein arises from electric forces that amino acids exert on one another, causing the chain to twist and fold on itself like a tangled telephone cord. To predict this process, one would ideally have to compute the interaction of each part of the chain with every other part as the protein goes through its contortions. As the length of the chain increases, such computation rapidly becomes intractable. Molecular biologists have therefore depended on empirical methods, particularly X-ray crystallography, to determine protein structures. Perfect protein crystals are difficult to grow, however, and X-ray diffraction patterns sometimes cannot be interpreted. A more recent technique, nuclear magnetic resonance (NMR), is limited to relatively small proteins and requires enormous amounts of time on a supercomputer. Hence, progress has been slow: although several thousand proteins have been sequenced, only about 300 structures have been determined.

In 1988 Terrence J. Sejnowski and Ning Qian of Johns Hopkins University decided to tackle the protein-folding problem with a neural network. Earlier Sejnowski, who is now at the Salk Institute and the University of California at San Diego, and graduate student Charles R. Rosenberg had acquired a measure of celebrity with NETtalk, a neural network that "learned" to pronounce written English. The pronunciation of a letter depends on surrounding letters, and so the network is designed to associate a letter with a phoneme by looking at several letters preceding and following it. The network's efforts are compared with the correct phonemes, and then a learning rule modifies the network so that eventually the network produces the correct phoneme most of the time.

Sejnowski and Qian thought deciphering protein folding might not be unlike learning to read: a sequence of amino acids is like a string of letters, and its folded structure is the associated "phoneme." Predicting the tertiary, or complete, structure seemed like a big leap, and so Sejnowski and Qian decided first to develop a network that could predict elements of the "secondary structure" (from which the tertiary structure is assembled). Working with three of the four known secondary structures—called the alpha helix, the beta sheet and the random coil—they trained their network to predict correctly 64.3 percent of the test sequences, which consisted of a structurally diverse group of proteins. This was better than the 50 percent prediction rate achieved by traditional means and had the advantage of being an automated process.

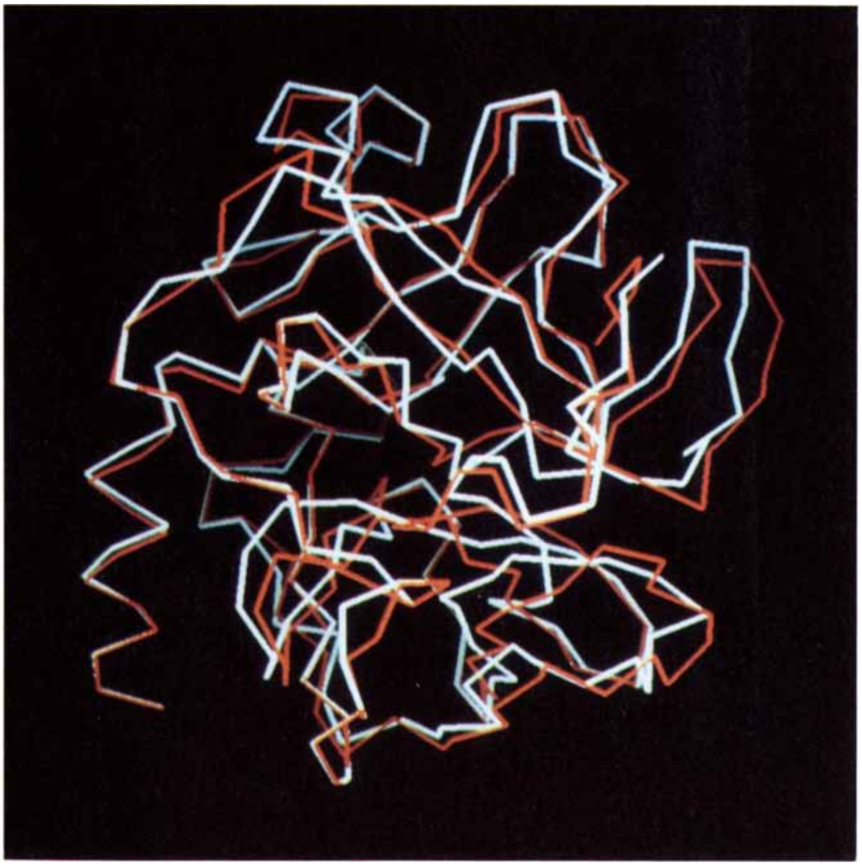
The effort prompted Søren Brunak and Rodney M. J. Cotterill of the Technical University of Denmark, Steffen Petersen of Novo-Nordisk and their colleagues at three other Danish institutions to try training a network using NMR data on folded protein structures. NMR produces information about the distance between amino acids in a folded protein, from which one can calculate the tertiary structure. The network was trained on NMR data for 13 proteases (enzymes that digest proteins). The group then tried to predict the structure of yet another protease, trypsin, and derived a result that closely matched the structure obtained by X-ray diffraction.

The result is impressive, Sejnowski says, although he cautions that the Danish network is trained on only a small family of highly similar proteins. "Its real contribution," he says, "is to show that a network can take a group of homologous proteins and pull out motifs that constrain the tertiary structures." Sejnowski thinks neural networks may one day work in concert with other methods to automate the prediction of protein structures.

In the meantime the Danish group hopes their network can be trained to predict a novel protein structure. But because the number of known protein structures on which the network must be trained is so small, its predictive power may be limited. Cotterill is now also at Proteus Biotechnology, a British company that plans to market a commercial version of the network.

Molecular biologists might complain such networks are black boxes that do not reveal the principles underlying protein folding. But Cotterill thinks these principles may in fact be encoded in the rich interconnections of the network. This possibility is now being examined by his group. "The door of the black box stands ajar," he says. *—June Kinoshita*

Neural networks prove adept at picking out structural motifs in folded proteins



STRUCTURE of trypsin 2TRM* (white) *predicted by a neural network resembles the homologous trypsin 4PTP structure* (red) *from X-ray crystallography.

MEDICINE

Space Invaders

Extra! Extra! Flu linked to sunspots!

Have the flu? Too bad. But try cheering yourself with this notion: the virus that caused your achy muscles, clogged nose and scratchy throat came not from Ted, your sneezing co-worker, but from outer space.

The story comes not from the *National Enquirer* but from *Nature.* That highly respected journal recently published a letter by the iconoclastic British scientists Fred Hoyle and Nalin C. Wickramasinghe. They are the authors of the panspermia hypothesis, which asserts that bacteria or viruses floating through the cosmos served as seeds for life on the earth. The theory has been dismissed by virtually the entire scientific community, but its authors remain unfazed.

In *Nature* they point to the "remarkable coincidence" between flu pandemics and peaks in the occurrence of sunspots over the past 200 years or so. The claim is timely: this year the sunspot cycle (normally 11 years long) is peaking, and the incidence of flu has reached epidemic proportions in Europe and the U.S.

What explains the correlation? Hoyle and Wickramasinghe speculate that the solar wind, which picks up during a sunspot peak, sweeps space-borne viruses down through the earth's atmosphere with greater than normal force. This process, rather than person-to-person infection, fuels flu epidemics, Hoyle maintained to SCIENTIFIC AMERICAN in a telephone interview. The reason the disease tends to strike in the winter, he added, is because cooler weather generates stronger downdrafts.

Unfortunately, the historical data presented by Hoyle and Wickramasinghe in *Nature* are somewhat disappointing; most of the pre-1900 pandemics actually missed the nearest sunspot peak by several years. Their hypothesis has a more fundamental problem. Most scientists think cosmic radiation would almost certainly destroy germs in space. If not, why haven't they been detected?

Hoyle has an answer. He is "99.9 percent sure" the U.S. has detected

Send in t

There's a good reason why more cars are looking more and more alike. It's called the Accord.

The Honda Accord has become the benchmark for other car makers. They take it apart. Study it. Take it apart again.

Now while they're busy doing that, we would like to move on to something bigger and better. The new Accord.

The new Accord was redesigned to create a new car. A car attuned to the driver. One that improves the quality of driving.

panels. So it's stiffer and more solid. The redesigned 4-wheel double wishbone suspension system works even better.

Power steering now operates at all speeds. And it progressively adjusts the amount of assistance for better road feel for the driver.

A new, powerful engine rests beneath

There is more room inside, the ride is more comfortable and the car is more responsive under all driving conditions.

You can see the smooth, flowing lines and the substantial increase in glass area. But the greater refinements are inside the car, where they matter most.

The body is built with thicker steel

the low, sloping hood. It's fuel injected. Unique balance shafts cancel inertial forces that cause other engines to twist and shake. You'll be amazed at just how smoothly and quietly it runs.

There is a new 4-speed automatic transmission available that's compact, smooth shifting and even more efficient.

germs in space or the upper atmosphere but has classified the results because they are "related to biological warfare."

Hmmm. I'd blame Ted. —*J.H.*

Tone of Relief

Ultrasound speeds the release of drugs from medical implants

Ultrasound has already proved itself as a tool for medical diagnostics; it is now carving a niche in the realm of therapeutics. Robert Langer of the Massachusetts Institute of Technology, Joseph Kost of Ben-Gurion University and Kam Leong of Johns Hopkins University have reported that ultrasound can dramatically and controllably boost the release of drugs from implanted biodegradable materials. Speculating on possible applications of this effect in the *Proceedings of the National Academy of Sciences,* they write, "One might envision that a patient could someday wear a portable triggering device (e.g., like a wristwatch) that could be used to augment release on demand."

Today's implants are plastic materials laden with drugs, which gradually diffuse out of the implant and into the bloodstream. A variety of implants, including some for contraception and chemotherapy, are now being developed or used clinically. Implants appeal to patients, because they are convenient, and to physicians, because they ensure that patients receive their medicine as intended.

Furthermore, oral or injected drugs sometimes take patients on a pharmaceutical roller-coaster ride, as blood concentrations of the drug quickly peak, then plummet. For long-term or chronic medical conditions, it is usually better to infuse small doses of drugs steadily into a patient's system. Outside of a hospital, this can often be achieved only with implants or small mechanical pumps.

Some medical problems, such as diabetes and asthma, can benefit from a combined form of treatment—steady release plus higher delivery on demand. A diabetic, for example, might wish to increase the release of insulin to his or her system temporarily after a meal. A pump with external controls can be adjusted accordingly, but a passive implant cannot.

To overcome this limitation, Langer and his colleagues have been observing the effects of various stimuli on polymer implants. According to their studies, implants made of biodegradable materials release drugs up to 20 times faster when exposed to ultrasound; the implants also degrade up to five times faster. Ultrasound can enhance the release of drugs from nonbiodegradable implants, too, by a factor of 10. The rate of release can be controlled by changing the intensity of the emissions from an external ultrasound source. Langer estimates that drugs in an implant, because they are in solid form, could be 100 times more compact than an equivalent pumped dose.

Sound is only one of many stimuli that have been investigated for controlling the release of drugs from implants; others include magnetism, heat and enzyme triggers. Langer has been involved in developing several of these systems and is optimistic about their prospects. Nearly all these approaches, however, involve potentially troublesome chemicals or nonbiodegradable materials. The obvious advantage of biodegradable implants is that they do not have to be removed when their contents are exhausted.

Before a biodegradable implant can be used on a patient, it must be proved that the implanted polymer, its breakdown products and the high concentration of the contained drug are safe; before ultrasound can be applied to improve drug delivery, it must also be shown that prolonged exposure to ultrasonic waves is not harmful. John H. Abeles, president of MedVest, Inc., a medical consulting company in Danville, Calif., suspects that in many instances the fine control over drug release possible with pumps could keep them safer and more practical than even regulatable implants. —*John Rennie*

OVERVIEW

Mandelbrot Set-To

Did the father of fractals "discover" his namesake set?

Who discovered the Mandelbrot set? This is not a trick question—or a trivial one. The set has been called (in this magazine) "the most complex object in mathematics." That is debatable, yet it is almost certainly the most famous such object. The infinitely intricate computer-generated image of the set serves as an icon for the burgeoning field of chaos theory and has attracted enormous public attention.

The set is named after Benoit B. Mandelbrot, a mathematician at the IBM Thomas J. Watson Research Center. He is best known for coining the term fractal to describe phenomena (such as coastlines, snowflakes, mountains and trees) whose patterns repeat themselves at smaller and smaller scales. Mandelbrot claims that he and he alone discovered the Mandelbrot set—which has fractal properties—about a decade ago. He refers to its image as his "signature."

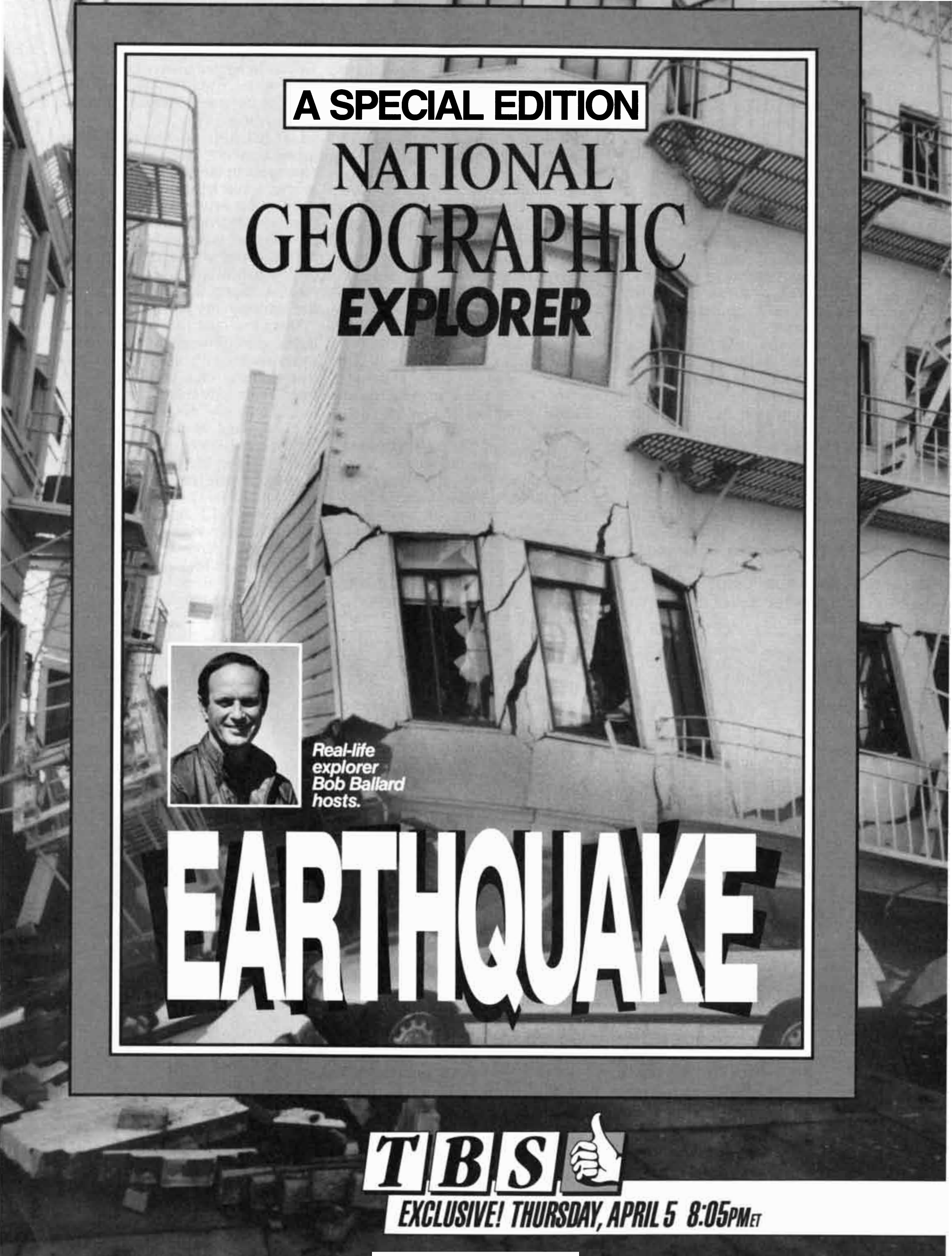
Three other mathematicians have challenged his claim. Two maintain that they independently discovered and described the set at about the same time as Mandelbrot did. A third asserts that his work on the set not only predated Mandelbrot's efforts but also helped to guide them. These assertions have long circulated in the mathematics community but have only recently surfaced in print.

Mathematicians are not known for priority battles, but Mandelbrot—a self-described "black sheep"—has often bumped heads with colleagues. "Were it not for his personality," remarks Robert L. Devaney of Boston University, who says he admires Mandelbrot's work, "there would be no controversy."

The scientific stakes are also high. Even those who scorn the set's popularity acknowledge its mathematical significance. Dennis P. Sullivan of the City University of New York calls it a "crucible" for testing ideas about the behavior of dynamical (or nonlinear, or complex, or chaotic) systems. "It is really quite fundamental," he says.

Part of the charm of the set is that it springs from such a simple equation: $z^2 + c$. The terms z and c are complex numbers, which consist of an imaginary number (a multiple of the square root of -1) combined with a real number. One begins by assigning a fixed value to c, letting $z = 0$ and calculating the output. One then repeatedly recalculates, or iterates, the equation, substituting each new output for z. Some values of c, when plugged into this iterative function, produce outputs that swiftly soar toward infinity. Other values of c produce outputs that eternally skitter about within a certain boundary. This latter group of c's, or complex numbers, constitutes the Mandelbrot set.

When plotted on a graph consisting of all complex numbers, the members of the set cluster into a distinctive shape. From afar, it is not much to look at: it has been likened to a tumor-ridden heart, a beetle, a badly burned chicken and a warty figure eight on its side.


A SPECIAL EDITION
NATIONAL GEOGRAPHIC
EXPLORER
Real-life explorer Bob Ballard hosts.
EARTHQUAKE
TBS
EXCLUSIVE! THURSDAY, APRIL 5 8:05PM ET

A closer look reveals that the borders of the set do not form crisp lines but seem to shimmer like flames. Repeated magnification of the borders plunges one into a bottomless phantasmagoria of baroque imagery. Some forms, such as the basic heartlike shape, keep recurring but always with subtle differences.

Today virtually anyone with a personal computer can "discover" the set [see "Computer Recreations," SCIENTIFIC AMERICAN, August, 1985]. But 11 years ago computers were much less powerful, and few mathematicians associated computers with serious mathematics.

Even Mandelbrot has described his first tentative steps toward the set in 1979 as "mindless fun." He began using a computer to map out Julia sets, which are generated by plugging complex numbers into iterative functions. The sets' peculiar properties had been described as early as 1906 by the French mathematician Pierre Fatou. They were named later for Gaston Julia, who successfully claimed that his work on the sets some dozen years later had greater significance than Fatou's. Mandelbrot, who was born 65 years ago in Poland, had read the work of both men and studied under Julia in the 1940's.

Mandelbrot's early computer images served to confirm his suspicion that Julia sets have fractal properties. He says he began producing recognizable pictures of the Mandelbrot set—which in a sense is a generalized version of all Julia sets—in late 1979. Mandelbrot subsequently displayed images of the set and elaborated on its significance in speeches, papers and books. This discovery and his other work in fractals were also celebrated in the media, in numerous books (notably the best-seller *Chaos,* by former *New York Times* reporter James Gleick) and in IBM advertisements.

No one denies that Mandelbrot's pictures and descriptions spurred other mathematicians to study the set. Two prominent examples are John H. Hubbard of Cornell University and Adrien Douady of the University of Paris. In the early 1980's—in the course of proving that tiny "islands" surrounding the main body of the set are linked to it by infinitesimal filaments—they named the set after Mandelbrot. "Mandelbrot was the first one to produce pictures of it, using a computer, and to start giving a description of it," Douady wrote in 1986.

Douady now says, however, that he and other mathematicians began to think that Mandelbrot took too much credit for work done by others on the set and in related areas of chaos. "He loves to quote himself," Douady says, "and he is very reluctant to quote others who aren't dead."

Last fall Steven G. Krantz of Washington University aired some of these grievances in the *Mathematical Intelligencer,* a quarterly journal. The main point of his article was that fractals, computer-generated graphics and other "popular" mathematical phenomena associated with Mandelbrot have contributed little of substance to mathematics, especially in comparison to the publicity they have garnered.

This view—and its opposite, which holds that Mandelbrot's "popular" work has been a stimulating force in mathematics—had been voiced before. Krantz introduced a new element into the debate, however, by stating that the Mandelbrot set "was *not* invented by Mandelbrot but occurs explicitly in the literature a couple of years before the term 'Mandelbrot set' was coined." He cited a paper by Robert Brooks and J. Peter Matelski published in the proceedings of a 1978 conference at Stony Brook, N.Y.

Sure enough, the paper contains the famous $z^2 + c$ formula and a crude but unmistakable computer printout of the set's basic image. Brooks and Ma-

telski say they did not actually present the paper at the 1978 conference, but they did circulate it as a preprint in early 1979. Brooks, who is now at the University of California at Los Angeles, also presented the paper at Harvard University in the spring of that year. (Mandelbrot, who held an appointment at Harvard then, says he did not hear Brooks speak and only saw the paper years later.) The paper was not published, however, until early 1981.

In a rebuttal to Krantz's article, called "Some 'Facts' that Evaporate upon Examination," Mandelbrot noted that he "fully published" on the Mandelbrot set before Brooks and Matelski did. (Mandelbrot's paper, published in the December 26, 1980, *Annals of the New York Academy of Sciences*, features a function and image that are variants of those now associated with the Mandelbrot set, which Mandelbrot did not publish until 1982.)

Mandelbrot also suggested that even if Brooks and Matelski's publication had preceded his, they still could not be considered discoverers of the set, because they did not appreciate its significance. "[They were] close to something that was to prove special, but they gave no thought to the picture," he wrote.

Brooks retorted in the following issue of the *Intelligencer:* "I don't know how he can be so sure of what we gave thought to and what we didn't." Brooks says he respects Mandelbrot's achievements as a popularizer and does not object to the set's being named after him. "It makes more sense than 'the thing with the big cardioid,'" he says, recalling how he and Matelski had referred to the set. "I just wish Mandelbrot were a little more gentlemanly."

Matelski, who works at the Hartford Graduate Center in Connecticut, notes that neither he nor Brooks asked Krantz to credit them with having discovered the Mandelbrot set. (Krantz confirms that another mathematician drew his attention to their paper.) But now that the issue has become public, Matelski says he and Brooks should be acknowledged as co-discoverers with Mandelbrot.

"You don't have to fully exploit the mineral resources of a continent to discover it," Matelski was quoted as saying in the *Hartford Courant,* a newspaper that reported on the dispute in December. "All you have to do is kneel down and kiss the beach."

A subtly different claim of precedence has been made by Hubbard, who is now considered one of the world's experts on the Mandelbrot set. In 1976, he explains, he began using a computer to map out sets of complex numbers generated by an iterative process known as Newton's method. Hubbard says he did not realize it then, but he had found a different way of generating the Mandelbrot set.

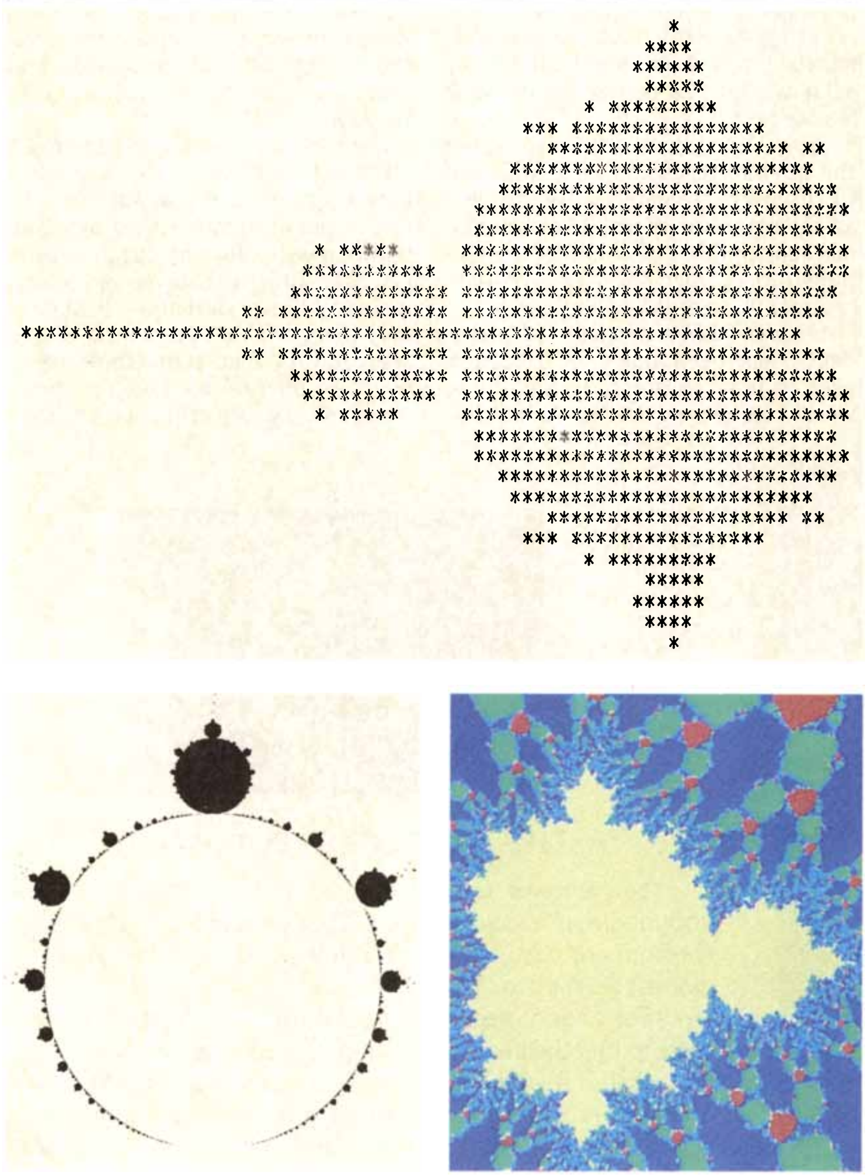
In late 1978 one of Hubbard's graduate students, Frederick Kochman, approached Mandelbrot at a conference and showed him Hubbard's pictures. Mandelbrot "didn't seem very interested," Kochman recalls. Yet shortly thereafter Mandelbrot wrote a letter to Hubbard inviting him to IBM to discuss his work. In the letter, which Hubbard kept, Mandelbrot wrote: "When sampling the works of Fatou and Julia, I had thought of doing these things myself, but had not mustered the courage. Nevertheless I can claim that I was awaiting your pictures for a long long time."

Hubbard says he went to IBM early in 1979 and, while there, told Mandelbrot how to program a computer to plot the output of iterative functions. Hubbard concedes that he did not appreciate the full significance of his own images then and that they showed only pieces of the Mandelbrot set. He also admits that Mandelbrot developed a superior method for generating images. Nevertheless, Hubbard says he was and continues to be "outraged" that Mandelbrot did not give him credit in the 1980 paper and later writings. "It was a breach of mathematical ethics," he asserts.

Mandelbrot recalls seeing "one impressively early drawing of a Julia set" by Hubbard but denies that it contributed to his own discovery. In response to Hubbard and Douady's charge that he is stingy in granting credit, Mandelbrot says he has also been accused of overcitation. He adds that his failure to cite the early finding of Brooks and Matelski might have spared them "derision" for "their failure to do anything with it."

What about the suggestion of Hubbard, Matelski and Brooks that the true discoverer of the Mandelbrot set is Fatou, who was the first to define the set and speculate about its properties? Brooks even proposes that "if Fatou had had access to modern computing facilities, he could have and would have drawn pretty much the same pictures that Matelski, Mandelbrot and I did." Mandelbrot calls such speculation pointless and insists that Fatou's definition of the Mandelbrot set does not constitute discovery. "Definition counts for nothing," he says. "You have to say why something is important."

The Mandelbrot set's history may, like its image, be more complicated than it first appears



MANDELBROT SET can be generated in different ways and take various forms. The image published by Robert Brooks and J. Peter Matelski in 1981 **(top)** ***was derived from the standard formula*** $z^2 + c$. ***The 1980 paper by Benoit B. Mandelbrot shows an image produced by a slightly different function*** **(bottom left).** ***John H. Hubbard determined some years later that an iterative process called Newton's method can also produce the distinctive image of the Mandelbrot set*** **(bottom right).**

Other mathematicians familiar with the case are somewhat bemused. "It seems strange to me that there should be such a fuss," remarks John Milnor of Princeton University. He maintains that neither Brooks and Matelski nor Mandelbrot did anything mathematically important. "Hubbard and Douady are the first ones to really obtain some sharp results," he says, "and let us know something about the set."

The dispute over precedence, Milnor suggests, may spring from a clash of different mathematical cultures. "In pure mathematics," he explains, "there is a tradition of letting others praise your work." Mandelbrot, he notes, is in applied mathematics.

"Mathematical developments don't take place single-handedly," William P. Thurston of Princeton points out, "and it's pretty common that things are not named after the first person to develop them. The Mandelbrot set follows that pattern." He suggests, however, that no one would begrudge recognizing Mandelbrot's achievements if he would reciprocate more himself. "He could be a bit more magnanimous," Thurston says.

Sullivan, who has also been acclaimed for his studies of the Mandelbrot set, calls himself "sort of a defender of Mandelbrot." Mandelbrot deserves to have the set named after him, Sullivan says, because his efforts brought the set to the attention of both the public and of the pure-mathematics community.

The fact that it was only "by coincidence" that the set proved later to be mathematically significant, Sullivan says, in no way diminishes Mandelbrot's achievement. "That's the wonderful thing about mathematics," he adds. "Even amateurs can make important contributions."

So who did discover the Mandelbrot set? Sullivan calls the question meaningless. Perhaps. Sheldon Axler, editor of the *Intelligencer,* plans to publish a letter pointing out that the Hungarian mathematician F. Riesz reported on work related to the set in 1952.

The final answer, if pursued, seems likely to recede in a blur of ever finer detail. *—John Horgan*

SCIENTIFIC AMERICAN April 1990 Volume 262 Number 4

Deforestation in the Tropics

Government policies that encourage exploitation—in particular excessive logging and clearing for ranches and farms—are largely to blame for the accelerating destruction of tropical forests

by Robert Repetto

Tropical forests are disappearing at the rate of tens of thousands of square miles per year. The deforestation is laying waste a valuable natural resource throughout much of the developing world and is driving countless plant and animal species to extinction, and it may well have significant effects on world climate.

Among the agents of the devastation are inefficient commercial logging operations and the conversion of forested areas to cattle ranching and agriculture. Data collected by numerous investigators and evaluated by my colleagues and me at the World Resources Institute indicate that both the logging and the conversion are largely the result of government policies. Many of those policies are driven by the severe economic pressures afflicting debt-burdened underdeveloped countries. Those pressures in turn are exacerbated by certain practices of developed countries and their national and international financial institutions. Hence the causes as well as the effects of tropical-forest degradation should elicit worldwide concern.

The destruction of tropical forests is a more serious problem than it was thought to be only a decade ago, judging by recent estimates based on remote sensing from satellites and on careful field surveys [*see bottom illustration on page 40*]. In India, for example, studies by the National Remote Sensing Centre lead to an estimated deforestation rate for the early 1980's of 1.5 million hectares (3.7 million acres) per year, some 10 times an earlier estimate by the Food and Agriculture Organization; the satellite imagery showed that large areas legally designated as forestland were already in fact virtually treeless.

Deforestation at this rate poses extreme risks to natural systems. The consequent release of carbon dioxide to the atmosphere is estimated to account for from 15 to 30 percent of annual global carbon dioxide emissions, and so it contributes substantially to the buildup of greenhouse gases. Moreover, the loss of tropical forests is rapidly eliminating the habitat of large numbers of plant and animal species. About half of the world's species inhabit tropical forests; in 10 biologically rich and severely threatened regions that account for 3.5 percent of the remaining tropical-forest area, 7 percent of all plant species will probably go extinct by the end of the century, if current trends continue.

A sense of crisis is emerging in the tropics as governments recognize that the rapid deforestation represents a

ROBERT REPETTO is director of the program in economic policies and institutions at the World Resources Institute in Washington, D.C. A 1959 graduate of Harvard College, he went on to get an M.Sc. from the London School of Economics and a Ph.D. in economics from Harvard University. Repetto was an economist in India, Pakistan and Indonesia, associate professor of economics and population at the Harvard School of Public Health and a consultant to a number of U.S. and multinational development-assistance agencies before joining the institute in 1983.

RAIN-FOREST LUMBER MILL is in the Brazilian state of Pará, where, as in many other developing countries, excessive logging is a major cause of deforestation. Often inefficient milling operations have been subsidized by governments in an effort to provide local employment.

waste of valuable resources and a severe economic loss. In some cases governments have taken action. In Thailand commercial logging was recently banned, over the protests of influential concession holders, when surveys showed that forest cover there had declined from 29 to 19 percent of the land area between 1985 and 1988 and after landslides from deforested hillsides cost 40,000 people their homes. In the Philippines, where undisturbed forests containing one valuable family of tall trees, the dipterocarps, have shrunk from 16 million hectares in 1960 to less than a million hectares left standing in remote hill regions, logging has been suspended in most provinces; as a result, mills in the Philippines are closing or are importing logs from Sabah and Sarawak, two states in Malaysia.

Mills in the once rich Indonesian production centers of Sumatra and Kalimantan are also experiencing shortages of accessible high-quality timber and are importing logs from Sabah, Sarawak and the Indonesian province of Irian Jaya. Government officials have begun to realize that Indonesia's ambitious plans for developing its timber industry could be thwarted by a lack of timber. Even Sabah and Sarawak, currently the major sources of logs in Asia, are harvesting almost twice the sustained yield of their forests, which are being depleted rapidly.

In the Ivory Coast, where forest cover has decreased by 75 percent since 1960, an estimated 200 million cubic meters of commercial timber has simply been burned to clear the land for agriculture, incurring a loss of perhaps $5 billion. In Ghana, where 80 percent of the forests have disappeared, the forest department estimates that only 15 percent of the timber was harvested before the land was cleared. In Brazil, where little timber is extracted before forestland is cleared by burning, the resulting loss in commercial timber approximates $2.5 billion annually.

Burning valuable timber in the course of clearing forests is only one obvious kind of wastage. The loggers themselves destroy enormous quantities of timber through careless use of equipment and inefficient logging practices. If loggers extract 10 percent of the timber in an area, selecting mature trees of the most valuable species, they typically destroy at least half of the remaining stock, including immature trees of the valued species as well as harvestable stocks of somewhat less desirable varieties. Loggers often keep reentering partially harvested areas to extract more timber before stands have recovered, inflicting heavy damage on residual trees each time and making regeneration impossible. In Ghana and the Ivory Coast, some stands have been reentered as often as three times in 10 or 15 years as concessionaires obtained sales contracts for logs of lesser-known species.

According to a recent study commissioned by the International Tropical Timber Organization, not even one tenth of 1 percent of remaining tropical forests are being actively managed for sustained productivity. Moreover, in most countries, forests designated for logging are left virtually unprotected from encroachment by settlers and shifting cultivators after a timber harvest and are thus exposed to burning and clearing. Surveys in the Amazon make it clear that deforestation is particularly rapid where roads for logging or other purposes have opened up a region.

The biological degradation of tropical forests carries a high and escalating price tag. The timber cost alone

has been unexpectedly high, because tropical-timber prices have bucked the general downward trend of commodity prices and many previously uncommercial species now find ready markets. Countries where loggers have been allowed to extract as few as two or three trees per hectare, destroying the rest as uncommercial, now regret their shortsightedness. The upward trend in tropical-timber prices is likely to continue as supplies are depleted in Asia, Central America and West Africa over the next decade; the timber in the Amazon Basin, which is now being recklessly burned, will become increasingly valuable.

Potential revenues from timber sales are by no means the only economic losses in deforested countries. Probably 70 percent of the wood harvested in tropical countries is used locally, mainly for fuel. As forests recede, severe fuelwood shortages loom. Other forest resources become unavailable to local residents, including animals killed for meat, fruits, oils, nuts, sweeteners, resins, tannins, fibers, construction materials and a wide range of medicinal compounds. In Indonesia the value of exported nontimber forest products had climbed to $123 million by 1986.

Recent studies have shown that the capitalized value of the income derived from such nontimber forest products—readily renewable resources that can be extracted sustainably—may greatly exceed that of the timber harvest. Moreover, the incomes so derived are the livelihoods of local residents, whereas the profits from timber operations are typically captured by distant elites or foreign corporations. Logging operations have sparked some violent protests by villagers in Sarawak and in the Philippines and other countries.

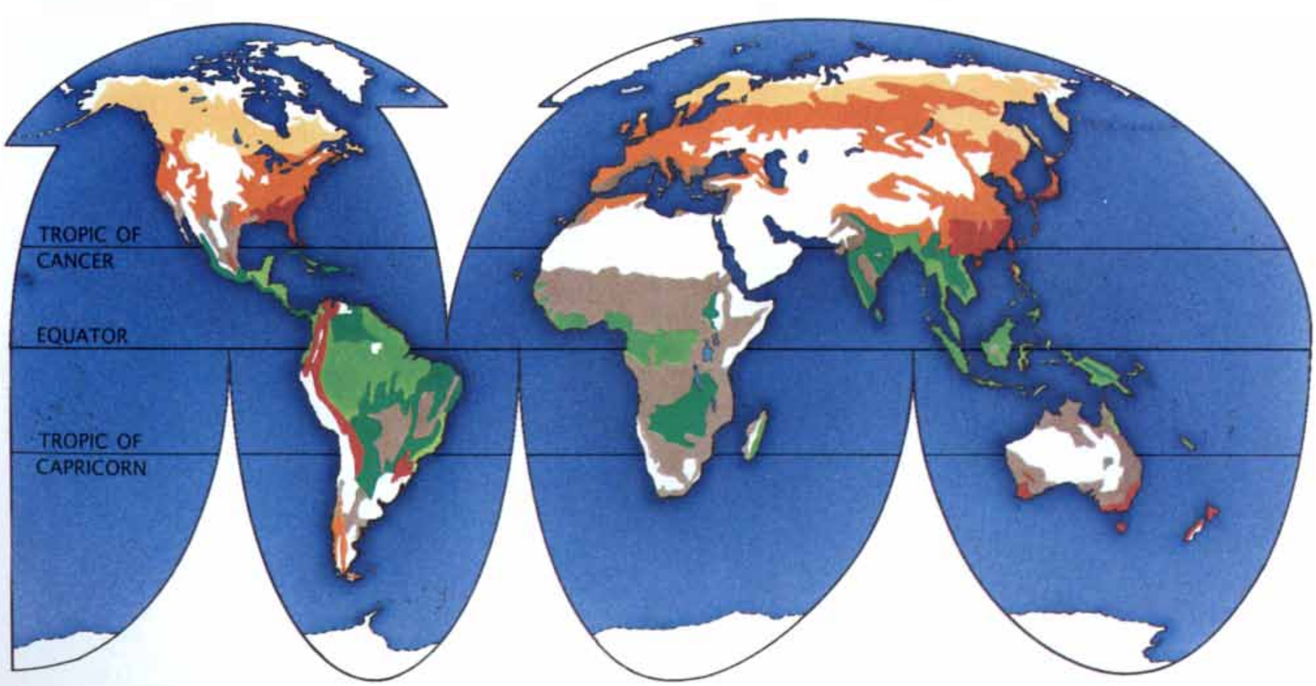
Quite aside from the loss of timber, deforestation often has a severe environmental impact on soil, water quality and even local climate. Shallow, easily leached soils are damaged by heavy equipment, and when they are exposed to heavy tropical rains, they can quickly erode or at least lose any remaining nutrients. Studies in Ghana showed that elimination of savanna forest raised soil-erosion rates from less than a ton to more than 100 tons per hectare per year, with a consequent nutrient loss 40 percent higher than what is being supplied by the application of chemical fertilizer. Riverine fisheries have been damaged by the increased sedimentation that results from erosion or by deforestation in floodplains that provide critical seasonal habitats for fish. Large-scale deforestation interrupts moisture recycling, thus reducing rainfall, raising soil temperatures and perhaps promoting long-term ecological changes.

Logging is often the first step in deforestation: it may be followed by complete clearing of trees and a deliberate shift to land uses—typically cattle ranching and inappropriate modes of agriculture—that not only are unsound environmentally but also result in direct economic loss. For example, studies in the rapidly deforesting Brazilian state of Acre show that because pastures quickly lose productivity and can carry few cattle, the present per-hectare revenue from the collection of wild rubber and Brazil nuts is four times as high as the revenue from cattle ranching.

All in all, both experience and analysis reinforce the argument that deforestation has not been a path to economic development; in most tropical countries it has instead been a costly drain on increasingly valuable resources. Moreover, deforestation is not inevitable. It is largely the consequence of poor stewardship, inappropriate policies and inattention to significant social and economic problems whose true locus is outside the forest sector.

To begin with, in the developing countries most governments—which are the proprietors of at least 80 percent of the mature closed-canopy tropical forest—have not put an adequate value on that resource. As proprietors, they could capture the entire resource value of the forests' timber, except for the cost of the labor and capital committed to managing and harvesting it, by charging high enough royalties and taxes or by selling harvesting rights to the highest bidders. Instead, with very few exceptions, governments have allowed most of these resource rents to flow to timber concessionaires and speculators, who are often linked to foreign enterprises.

In the Philippines, for example, if the government had been able to collect the full resource value of the roughly three million cubic meters of timber



harvested in 1987, its timber revenues would have exceeded $250 million—more than six times what was actually collected. Low royalties and taxes, combined with widespread log smuggling and tax evasion, left much of the excess profits in the hands of timber-concession holders, mill owners and timber traders. The Asian Development Bank has estimated that total profits averaged at least $4,500 per hectare harvested.

Governments have created such windfalls as these by keeping royalties and fees charged to timber-concession holders low, reducing export taxes on processed timber to stimulate domestic industry and granting income-tax holidays to logging companies. Moreover, governments have failed even to enforce the modest official charges effectively. (Between 1979 and 1984 in Indonesia, 125 million hectares were harvested, but taxes and royalties were collected on only 86 million.) As a consequence, few tropical countries have limited timber exploiters to a normal rate of profit and thereby captured the value of the forest resource for the public treasury [*see illustration on page 41*].

The resulting bonanza atmosphere has sparked timber booms throughout the tropics, drawing both domestic and foreign entrepreneurs—many with little forestry experience—into the search for quick fortunes. Under their pressure, governments have awarded timber concessions that cover areas far greater than they can effectively supervise or manage and that sometimes extend beyond designated production forests into protected areas and national parks. In the Ivory Coast concessions were let for two thirds of the nation's production forests in just seven years. Of 755 politically favored concessionaires, only 51 actually log their holdings themselves; most of them merely sell their cutting rights, profiting as middlemen. In Indonesia, Thailand and the Philippines, the areas under concession exceed the total area of production forest.

COOL CONIFEROUS FOREST
TEMPERATE MIXED FOREST
WARM TEMPERATE MOIST FOREST
TROPICAL MOIST EVERGREEN FOREST
TROPICAL MOIST DECIDUOUS FOREST
DRY FOREST

WORLD'S FORESTS can be grouped into six main vegetation zones. The tropical forests dealt with in this article are moist evergreen forests, moist deciduous forests and some dry forests.

The opportunity for private gain has attracted politicians as well as businessmen. In Thailand, Sarawak, Sabah, the Philippines and other places, cabinet ministers, senators and other senior politicians are involved in the timber industry. In the Philippines, for example, Senator Juan Ponce Enrile, the principal opposition leader, holds extensive timber concessions he acquired under the Marcos regime. In Indonesia most of the 544 concession holders are retired military or government officials who can bring pressure in Djakarta to halt investigations into violations of forestry regulations. Effective supervision by forestry-department personnel, often low-ranking officials, is virtually impossible.

While they sacrifice enormous sums in potential forest revenues, governments in the tropics are failing to invest enough in stewardship and management of the forest. In Indonesia nearly half of all trained foresters work in Djakarta, hundreds of miles by sea from the forests; those who do get out into the field find themselves dependent on concession holders for shelter and transportation. A study in Ghana found that 66 percent of all government posts for professional foresters, 54 percent of the posts for junior professionals and 43 percent of the technical-grade positions were vacant. Gabon has enough forestry personnel, but there is no way for them to do their job in the field: the departmental budget was reduced by 75 percent between 1984 and 1988. All of this means that although in many countries forestry codes and concession agreements are worded to ensure sustained productivity over at least several cycles, almost nowhere are forests being managed to achieve that goal.

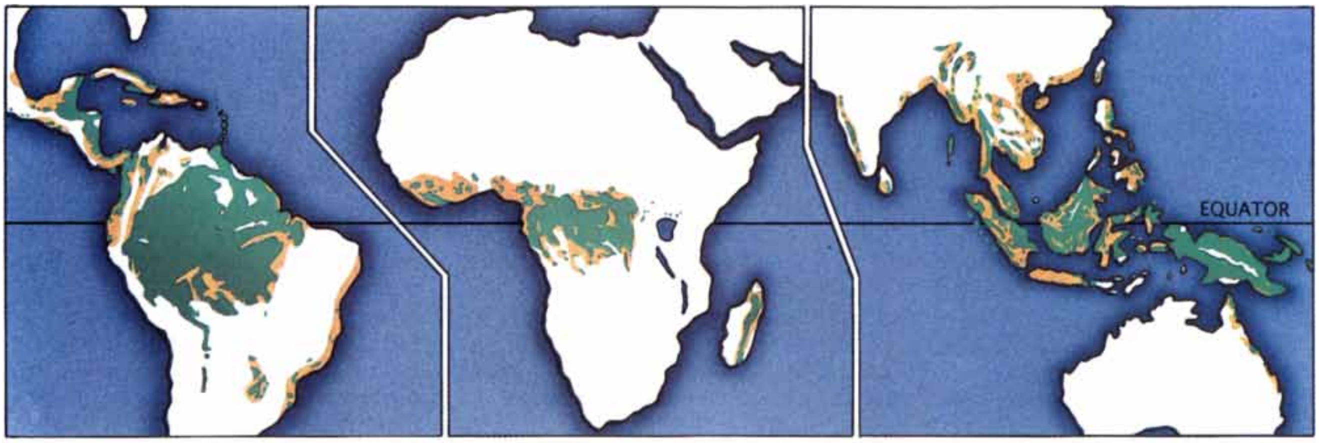
Ineffective government supervision is compounded by the perverse incentives established for timber companies by the terms of concession agreements, which actually discourage any possible interest loggers might have in management for sustained yield. Even though intervals of 25 to 35 years are prescribed between successive harvests in selective-cutting systems—and longer intervals in monocyclic systems (when all salable timber is extracted at once)—most agreements run for 20 years or less, some for less than five years. Concession holders are given little reason to care whether or not productivity is maintained for future harvests.

Again, relatively undifferentiated fees are often levied, based simply on the volume of wood extracted. This encourages "high-grading," a practice in which loggers take out only logs having the highest value and do so over large areas and at minimum cost. Because trees whose standing value is less than the royalty rate can be destroyed with impunity, extensive damage is often inflicted on residual stands. In Sabah, Indonesia and the Philippines, from 45 to 75 percent of residual trees are destroyed or seriously damaged during harvesting operations. Royalties based not on what is extracted but on the size of the concession and on the total salable timber it contains would encourage more complete utilization of the timber within a smaller harvesting area; ad valorem royalties (based on the value of extracted logs) would also encourage fuller utilization.

Distorted incentives also reduce the efficiency of wood-processing industries. Many countries seek to increase both employment and the value added to forest products domestically by encouraging processing rather than the exporting of logs. They must provide strong incentives to local mills to overcome high rates of protection against the importation of processed wood in Japan and Europe. Extreme measures, such as bans on log exports or export quotas based on the volume of logs processed domestically, have created inefficient local industries, which are sometimes set up only to preserve valuable log-export rights.

In the Ivory Coast such quotas have created a large processing industry requiring 30 percent more logs than efficient mills would consume to produce the same output. This inefficiency is supported by the sale of rights—worth as much as $15 per cubic meter on the open market—to export high-value logs. In Zaire concessionaires must process 70 percent of their harvest domestically. The requirement has increased timber cutting, because profitable export of prized species supports inefficient sawmilling, which dumps low-value output domestically at prices about 30 percent below production costs.

Such extreme protection can create powerful local industries able to resist regulation. Indonesia has successfully captured between 70 and 80 percent of the world's hardwood plywood market by banning log exports and providing generous industrial incen-



TROPICAL RAIN FORESTS once extended over the entire area shown in color. Now the rain forests have disappeared from the yellow areas; their current extent is shown in green. The map is based on one published by the Smithsonian Institution.

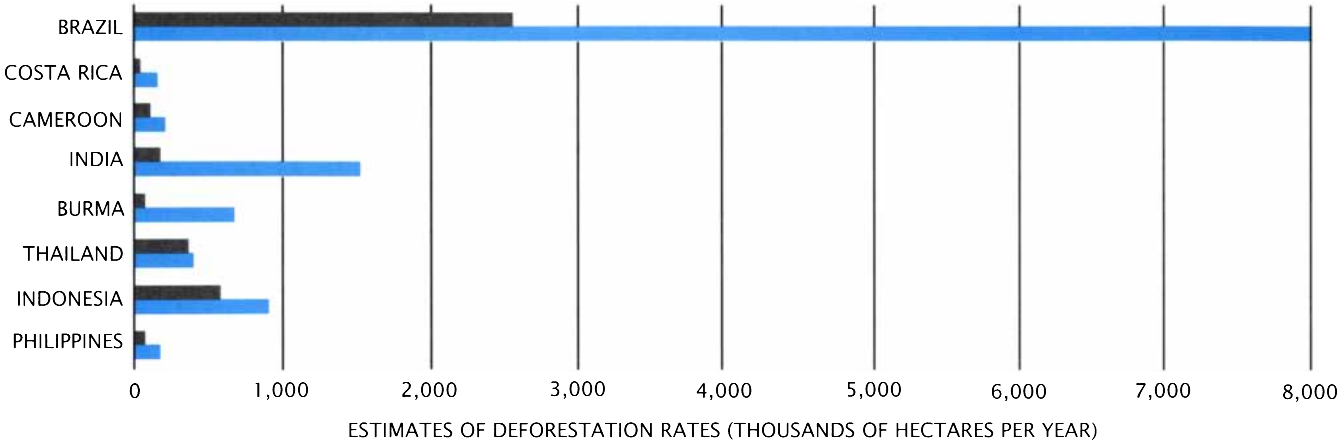
tives. The rapidly expanding but inefficient processing industry now consumes 35 million cubic meters of logs annually, more than previous peak exports, and current plans call for doubling capacity during the 1990's.

Countries sheltering inefficient processing industries can incur heavy economic and fiscal losses. In the Philippines each log exported as plywood is worth from $100 to $110 less per cubic meter than it would be if exported without processing or as sawed timber; the government sacrifices more than $20 million annually in forgone export taxes to encourage these plywood exports.

Industrialized countries have contributed to—and profited from—these forest-policy problems in the tropics. European and U.S. companies have held interests in logging and processing enterprises, especially in tropical Africa and Latin America, but Japanese business now heavily outweighs its rivals in the tropical-timber trade. Japan is the largest importer, accounting for 29 percent of the tropical-timber trade in 1986, roughly the same share as the European Economic Community. Imports (which, unlike those of the EEC, are mostly of unprocessed logs) were 30 percent higher in 1987, mainly because of a construction boom in Japan, where most tropical-hardwood imports are processed into construction plywood, primarily as disposable forms for molding concrete.

Large Japanese trading companies are involved in all stages of exploitation, as partners and financiers of logging concessionaires, as exporters and importers, and as processors and distributors. As log supplies were successively depleted, Japanese firms have shifted their attention from the Philippines to Indonesia, then to Sabah and Sarawak, and now they are interested in Amazonian forests. François Nectoux and Yoichi Kuroda find that the Japanese have shown little interest in sustained management of their holdings; their highly leveraged operations have harvested as much as possible as fast as possible in order to pay off financing charges. Moreover, Japanese firms have participated in the bribery, smuggling and tax evasion that make tropical timber cheap to import and at the same time deprive exporting countries of much of the value of their resource.

Inadequate forest policy and management are often abetted by misguided agricultural policy. Many countries actively encourage the conversion of tropical forests to other uses. Rules of land tenure in many states, such as Sabah, allow private parties to obtain title to forested land by showing evidence of "improving" it—by clearing away the trees, for example. In the Philippines, Brazil and elsewhere, recognized rights of occupancy or possession are awarded on the basis of the area of land cleared.



RATES OF DEFORESTATION appear to be increasing. Here estimates made by the Food and Agriculture Organization in the early 1980's (*gray*) are compared with more recent estimates (*color*) based on satellite imagery and field surveys.

Such provisions often become a mechanism for privatizing land from the public forest estate. Those who obtain ownership soon sell out to larger capitalists, who consolidate the land to establish private ranches and accumulate speculative holdings.

In many cases such activities would be uneconomic without heavy government subsidies. In the Brazilian Amazon, road-building projects financed by the federal government and multinational development banks have fueled land speculation. More than 600 cattle ranches, averaging more than 20,000 hectares each, have been supported by subsidized long-term loans, tax credits covering most of the investment costs, tax holidays and write-offs. The ranches proved to be uneconomic, typically losing more than half of their invested capital within 15 years.

Indeed, surveys showed that meat output averaged only 9 percent of what was projected and that many ranches were reorganized and resold repeatedly, having served only as tax shelters. (In that respect, they were unquestionably productive, generating returns of up to 250 percent of their owners' actual equity input.) Although the Brazilian government has suspended incentives for new cattle ranches in Amazonian forests, supports continue for existing ranches, covering 12 million hectares, that have already cost the treasury more than $2.5 billion in lost revenue.

More general agricultural policies contribute indirectly to deforestation. In Latin America and the Philippines the aggregation of the better agricultural land into large, generally underutilized estates pushes the growing rural population into forested frontiers and upper watersheds. The extreme concentration of landholdings is supported by very low agricultural taxes that make farms and ranches attractive investments for people in upper-income brackets, for whom it costs almost nothing to keep extensive holdings that generate relatively little income.

Subsidized rural-credit programs also promote land concentration: ceilings on interest rates inevitably lead banks to ration credit in favor of large landholders who have ample collateral and secure titles. Particularly in inflationary settings in which land provides security, large landholders with access to virtually free credit can easily buy out small farmers who cannot finance investments to raise agricultural productivity. Many of the recent migrants into Rondônia and Acre in Brazilian Amazonia are small farmers and farm laborers who have been displaced from Paraná by large-scale mechanized cultivation.

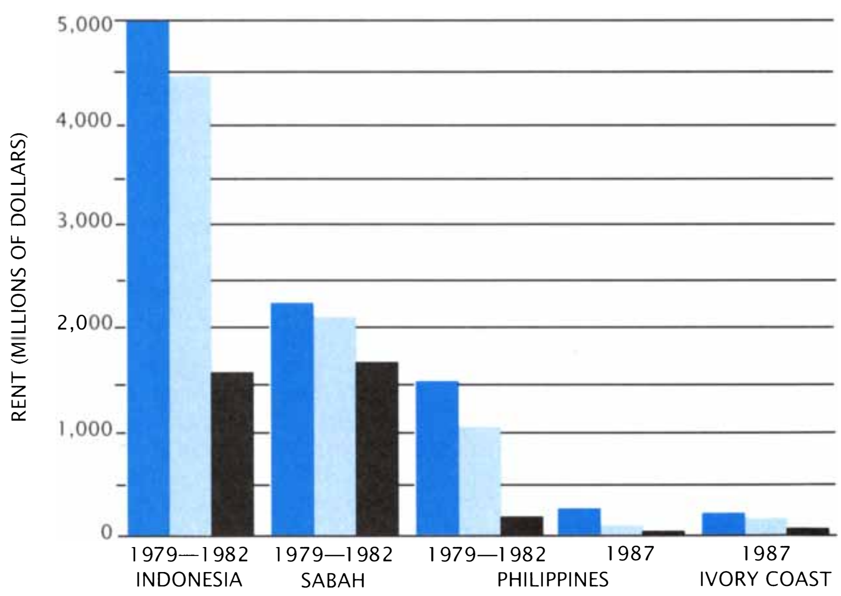
In many countries, deforestation has provided a temporary escape valve—a respite from development pressures that can be dealt with effectively only at a more fundamental level. In the Philippines, population growth rates in the forested uplands are even higher than the high national average of 2.5 percent per year, resulting in high rates of deforestation and soil erosion. Yet the government has been reluctant to address population control directly or to attack highly skewed patterns of landholding in the lowlands.

The Indonesian government's ambitious "transmigration" program, which has so far resettled about a million families from crowded Java to the outer islands—80 percent of them to sites cleared in primary or secondary forest—was largely an attempt to provide employment and livelihoods. At a cost of $10,000 per household (in a country that invests only $125 per capita annually), transmigration could obviously not compensate for slow employment growth on Java itself, and it has been sharply curtailed in recent budget cuts necessitated by lower petroleum prices.

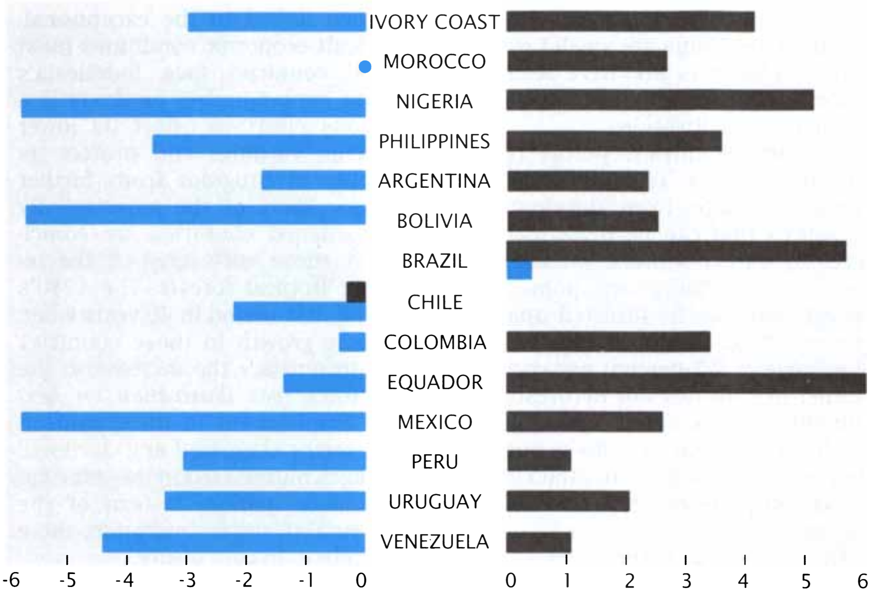
Indeed, rapid deforestation in the tropics during the 1980's has generally been linked to the exceptionally difficult economic conditions most tropical countries face. Indonesia's drive to export timber products is a conscious effort to offset its lower petroleum earnings and protect its development program from further cutbacks. Many of the most heavily debt-burdened countries are coincidentally those with most of the remaining tropical forests. The 1980's were the first period in 40 years when economic growth in those countries failed to outpace the increase in the labor force [*see illustration on next page*]. Employment in the organized urban sector stagnated and declined; real wages plummeted in the informal urban labor market. Instead of the usual rural-to-urban migration, there was a pileup in agriculture.

In Brazil, for example, the agricultural labor force grew by 4 percent a year between 1981 and 1984, compared with a growth rate of only .6 percent between 1971 and 1976; agricultural wages fell almost 40 percent in real terms between 1981 and 1985. With no alternative, given the concentration of agricultural land in large holdings and the absence of jobs, rural households migrated to the frontier in increasing numbers. A more favorable economic climate could reduce the pressures of unemployment, poverty and population growth on the remaining tropical forests.

Is there hope for improvement?



TROPICAL COUNTRIES have generally failed to collect (in royalties, export taxes and other fees) the full resource value of timber taken from their forests. The dark-colored bars show, for four-year periods or for single years, the "potential rent": what might have been collected if harvested logs had been disposed of (exported as logs, sawed or further processed) to yield the highest possible net return. The light-colored bars show the "actual rent" that could have been collected given the actual disposal of the harvest. What the governments did in fact collect is shown in black.



GOVERNMENT POLICIES that promote deforestation are often motivated by economic pressures. One such pressure is illustrated. From 1965 to 1980 (*black bars*), economic growth (measured by gross domestic product) generally outpaced the growth of labor forces, but from 1980 to 1986 (*colored bars*), labor forces grew faster than most economies, generating unemployment and reducing government revenues.

There are, in fact, signs of a new approach to forest policy reflecting increasing awareness of the national and global significance of tropical forests.

Many countries are taking steps to capture resource rents at full value. The government of the Philippines has imposed partial bans on logging, cracked down on illegal logging and raised timber royalties. It plans to increase timber taxes further and is considering assigning harvesting rights on the basis of competitive bids. The Ivory Coast also expects to move to competitive bidding. Indonesia has raised timber taxes substantially. Ghana, with World Bank assistance, has doubled timber royalties to an average of 12 percent of export value and plans a further 50 percent increase by 1992.

A number of governments are now strengthening their forest-management capability with the help of development-assistance agencies. The World Bank and the Asian Development Bank now have loans for forest-management improvement in the pipeline for a dozen countries. Most of these loans support forest-policy reform as well as institutional strengthening. Under the aegis of the Tropical Forestry Action Plan, sponsored by the World Resources Institute, the World Bank, the U.N. Development Program and the Food and Agriculture Organization, more than 50 countries are preparing national action plans to conserve and manage their forests.

International interest in tropical forests has bloomed, accompanied for the first time by a willingness to contribute to their maintenance. Several voluntary organizations in developed countries have raised money for debt-for-nature swaps: a bit of the external debt of a tropical country is bought up at a discount and then exchanged for a local-currency fund (usually to be managed by a local voluntary agency) that will finance forest-conservation programs. Some business groups have also taken an active interest. Associations of tropical-timber traders in the Netherlands and the U.K. have proposed that all importing countries levy a surcharge on tropical-timber imports to create a fund for forest conservation.

There remains a great deal more that the world outside the tropics might do. Inappropriate consumption of tropical hardwoods (for disposable concrete molds, for example) contributes to deforestation. Some businesses in industrialized countries are still taking part in forest destruction. Barclays Bank was recently found to be the majority owner, through its Brazilian subsidiary, of two huge Amazonian cattle ranches that have burned half a million acres of forest to create pasture. (The *Sunday Times* of London reported that the bank's chairman, on learning of this involvement, declared, "Being personally an extremely keen gardener and botanist...I was extremely cross.")

Development-assistance agencies are still financing activities that are destructive to tropical forests. The African Development Bank has recently agreed to a project that will run a road through one of the Ivory Coast's few remaining tracts of rain-forest and mangrove habitat. Another of its projects would develop sawmilling capacity affecting more than 800,000 hectares of virgin forest in the Congo, even though the country has no forest-management capability and the area in question is the home of Pygmy communities. Such projects should be supplanted by others designed to improve forest management and step up the pace of reforestation.

The potential scope of international cooperation to halt the destruction of tropical forests is large. The Tropical Forest Action Plan provides one useful framework. The Montreal Protocol for protection of the ozone layer and a proposed convention to mitigate global climate change could also be powerful mechanisms for international cooperation. Fees levied on chlorofluorocarbons and taxes on fossil-fuel and other greenhouse gases in industrialized countries would help to reduce emissions and also provide funds needed to implement national programs formulated under the Tropical Forest Action Plan. Programs in which debt reduction is linked to improved resource management and conservation could be expanded substantially.

New forms of international cooperation would reflect the world's growing awareness that the disappearing tropical forests are not only national treasures but also essential elements of the biosphere on which everyone, everywhere, depends.

FURTHER READING

THE PRIMARY SOURCE: TROPICAL FORESTS AND OUR FUTURE. Norman Myers. W. W. Norton & Co., Inc., 1984.

THE FOREST FOR THE TREES? GOVERNMENT POLICIES AND THE MISUSE OF FOREST RESOURCES. Robert Repetto. World Resources Institute, 1988.

PUBLIC POLICIES AND THE MISUSE OF FOREST RESOURCES. Robert Repetto and Malcolm Gillis. World Resources Institute, 1988.

TIMBER FROM THE SOUTH SEAS: AN ANALYSIS OF JAPAN'S TROPICAL ENVIRONMENTAL IMPACT. François Nectoux and Yoichi Kuroda. Gland, Switzerland, World Wildlife Fund International, 1989.

THE POWER OF BELIEF: #2 in a series

Who will nourish the seeds of tomorrow's technologies?

The handmaiden of every new technology is risk. There are no guidebooks to the New. ■ Success or failure often hinges on instinct, and the willingness to invest in a belief. ■ Decades ago, Motorola committed millions to an emerging technology that many dismissed as a gimmick. Today, cellular phones are used in more than 40 countries.


Motorola's EMX family of switches are now the most widely used cellular telephone switching systems in the world.

Winner 1988 Malcolm Baldrige National Quality Award

And, we produced the first hand-held cellular portable phone. In 1989, we introduced the Motorola Micro T·A·C(TM) Personal Telephone, shirt-pocket small, with a fraction of the parts of the original cellular phone. ■ These developments, along with others, are the product of an annual R&D investment more than twice the national average. Such is the measure of our belief.

The new, body-friendly Motorola Micro T•A•C™ Personal Telephone is the smallest cellular telephone ever made.

Building On Beliefs

MOTOROLA

Ⓜ, Motorola and Micro T•A•C are trademarks of Motorola, Inc., © 1989 Motorola, Inc
For more information call 1-800-457-0400.

What is Happening at the Center of Our Galaxy?

This mysterious region contains antimatter, radioactive clouds, veiled sources of intense radiation, swirling gas and dust and an unseen object with a tremendous gravitational pull—perhaps a massive black hole

by Charles H. Townes and Reinhard Genzel

On a clear, moonless night the shimmering light of the Milky Way glows especially bright toward the constellation Sagittarius. For years astronomers have been aware, from the distribution of groups of stars and from measurements of stellar motions, that objects in our galaxy must travel in orbits around a center located in that direction. Astronomers also have found that in most cases other galaxies are especially bright toward the center because the density of stars increases markedly there. In many instances the central regions also seem to be the sites of intriguing behavior, including the generation of enormous quantities of energy, peculiar radiations and other unusual effects. More and more, it seems that massive, unimaginably dense objects—black holes—lie at the heart of some of these galaxies. Could our own galaxy also harbor such an exotic object at its center?

CHARLES H. TOWNES and REINHARD GENZEL collaborate on studies of the structure and dynamics of the Milky Way. Townes is professor emeritus of astronomy at the University of California, Berkeley. He received a Ph.D. in physics from the California Institute of Technology in 1939. His long career has involved forays into radio and infrared astronomy, microwave spectroscopy, nuclear and molecular structure and quantum electronics; he was awarded a Nobel prize in 1964 for his contributions to the last of these fields. He continues to conduct research on the nature of the galactic center and has recently completed development of novel-design infrared telescopes for more precise observations of this and other intriguing regions. Genzel is director at the Max Planck Institute for Physics and Astrophysics and a visiting professor of physics at the University of California, Berkeley. He earned a Ph.D. in physics and astronomy from the University of Bonn; his current work also focuses on observations of infrared emission from galactic and extragalactic sources.

The central region of the Milky Way has fascinated astronomers for many decades. After all, our galactic center is only about 25,000 light-years away as opposed to millions of light-years for centers of the nearest other galaxies, and so it is the one astronomers might reasonably hope to see and to understand best. Yet for a long time there was no direct way to see the center of our galaxy or to learn much about it, because it is cloaked in large and dense clouds of gas and dust.

Recent discoveries and new technologies have made it possible at last to study the center of our own galaxy in some detail. These developments include improved techniques for collecting and analyzing astronomical radio waves and infrared radiation as well as space flights above the earth's atmosphere, which have made possible the detection of energetic X-ray and gamma-ray radiation emanating from the center of the galaxy.

All these types of waves—radio, infrared, X-ray and gamma-ray—are similar to visible light in that they are all forms of electromagnetic radiation and differ only in their wavelengths and energy levels. Unlike light, however, they can penetrate interstellar clouds of dust with some ease and hence provide a window into the structure and dynamics of the galactic center. Each wavelength region probes different aspects of the physical properties there; for example, X rays are emitted by very hot gas, whereas much infrared radiation is emitted by interstellar dust grains and cooler gases. Studies at these various wavelengths have produced a wealth of information about and new understanding of the central part of our galaxy. They have also raised a host of new puzzles.

Specific emissions from the center of the galaxy were first discovered by Karl Jansky in the early 1930's, although his work was neither well appreciated nor exploited until substantially after World War II. Jansky, who worked at the Bell Telephone Laboratories, was examining radio noise that might affect long-range radio communication. Exploring radio frequencies of about 20 megahertz, Jansky found the radio noise to be especially strong in a particular direction. With the persistence of a thorough engineer, he found this direction varied, with respect to day, night and season of the year, as if it were fixed among the stars. In time, he recognized this fixed position was the direction in which the center of the galaxy was thought to lie.

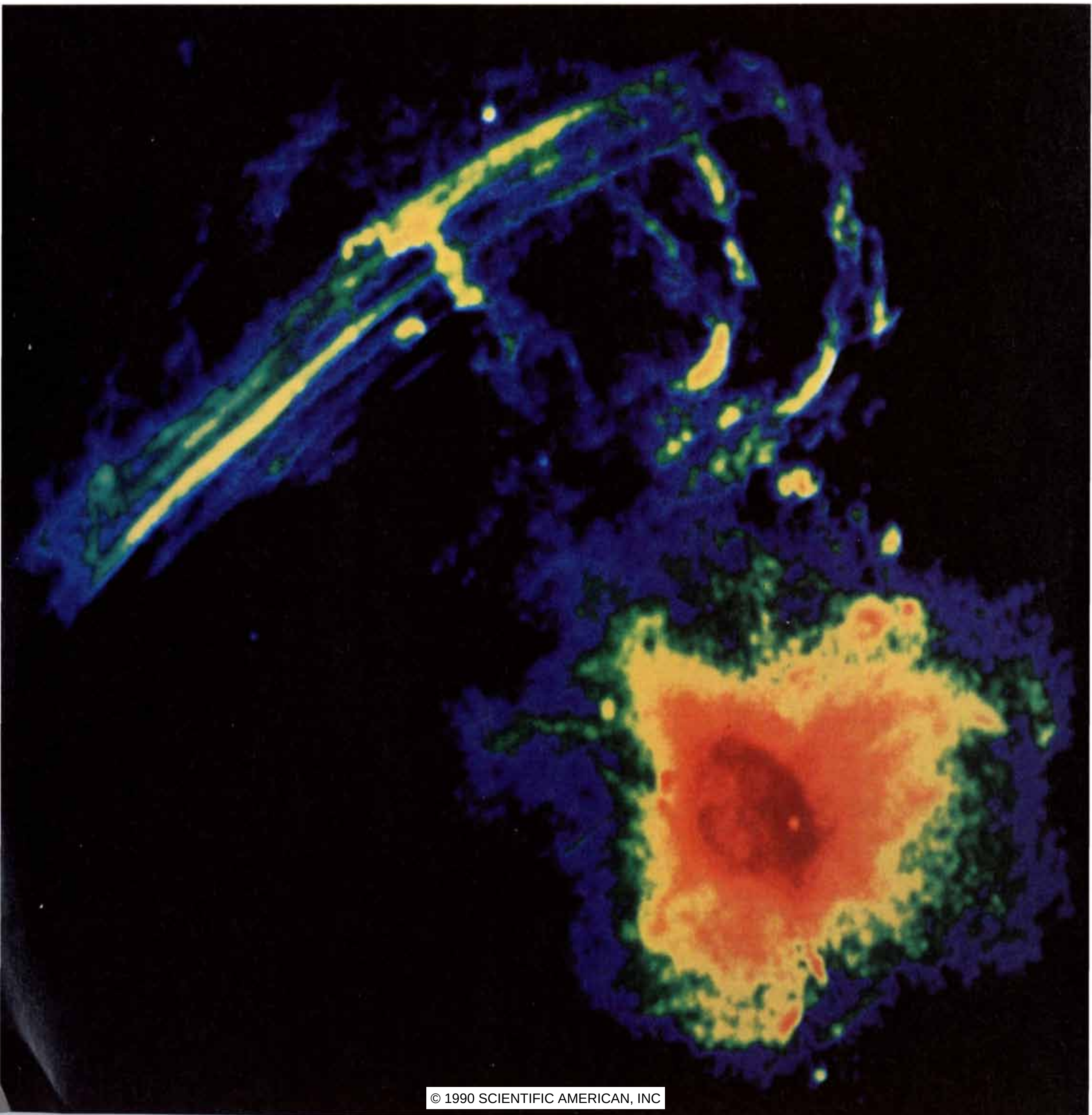
Jansky commented that the sound of the "cosmic noise" was similar to that generated by electrons moving in an electric resistor. This radio noise is now known to be radiation produced by electrons colliding with protons (hydrogen nuclei) in hot gas surrounding the galactic center and by electrons

THE GALACTIC CENTER is the site of disrupted structures and energetic radiation. Narrow arcs of gas roughly 200 light-years long radiate 20-centimeter radio waves that have been mapped by the Very Large Array (VLA) antennas in Socorro, N.M. Long, thin streamers (*blue*) emit radiation generated by electrons spiraling around magnetic field lines; the more prominent arcs close to the center (*red*) are thought to be illuminated by ultraviolet radiation from an as yet unidentified source near the center of the galaxy.

spiraling through a magnetic field. At the time, the origin of these radio waves was unknown, and astronomers were not impressed with the possibility of learning much from them.

The development of radar and microwave electronics during World War II fueled interest in radio frequencies and encouraged investigation of the cosmic noise. Enormous radio antennas now have been erected in many countries to study radio waves from astronomical sources. These instruments have revealed a great deal about the structures of the sources of radio emissions and the mechanisms that produce these emissions. A particularly fruitful technique has been radio interferometry, in which multiple antennas are linked together to improve sensitivity and resolution. Signals received by two or more antennas separated by thousands of kilometers can be combined to yield the effective resolving power of a single receiver as wide as the distance between the antennas, a technique known as very-long-baseline interferometry. The Very Large Array (VLA), a collection of 27 movable antennas located in Socorro, N.M., has been especially valuable in studying the galactic center.

A recent and crucially important development has been the ability to detect and identify radio emissions created by specific resonances of molecules. This technique has permitted astronomers not only to identify the composition of interstellar gas clouds but also, by exploring detailed characteristics of the emissions, to determine their density, motions, temperature



and source of energy. Nevertheless, many radio studies of the galactic center still focus on the radio noise that Jansky first detected, which is associated not with discrete molecular emissions but rather with a continuous spectrum of frequencies.

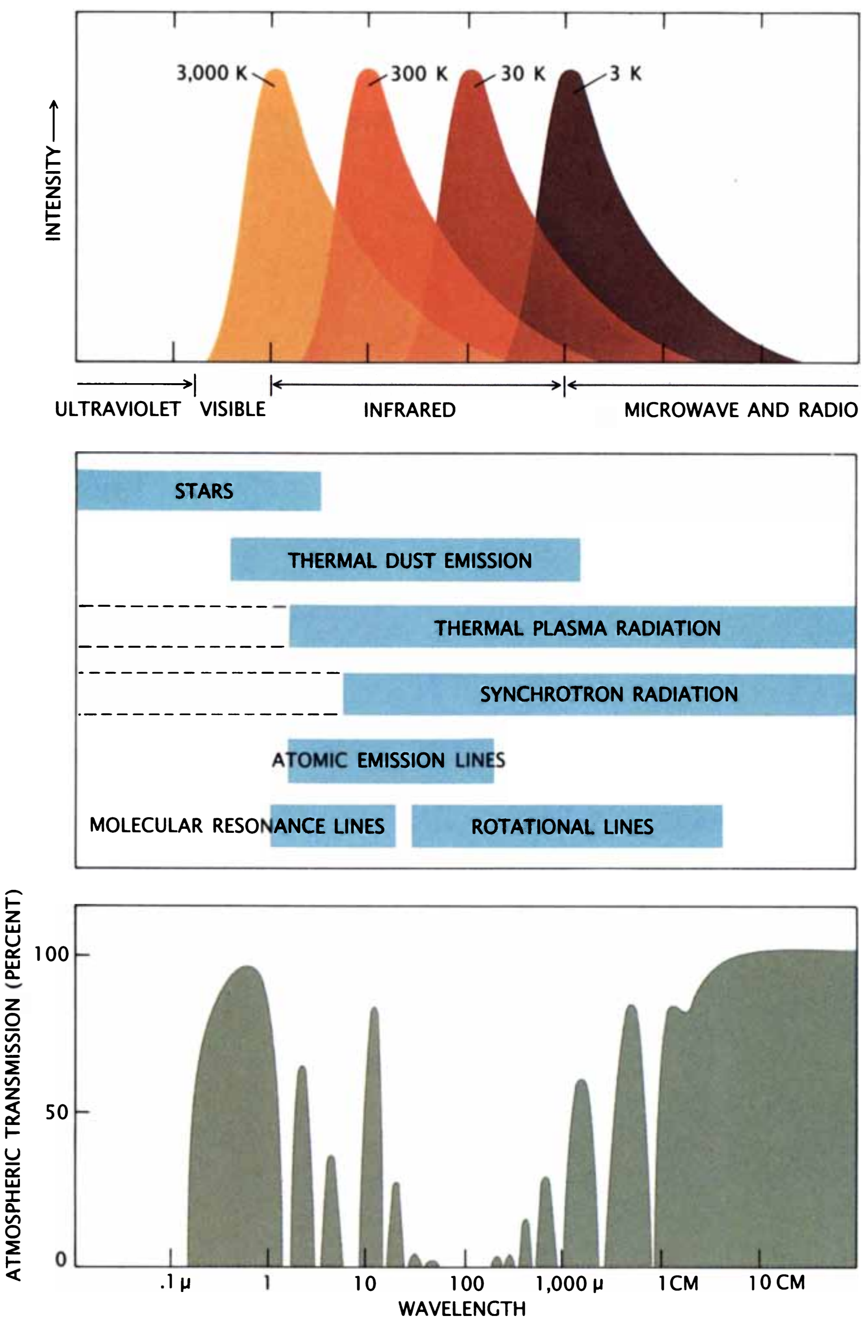
Such continuum radiation from the Milky Way is emitted primarily from regions of ionized gas—that is, gas composed of atoms that have been stripped of some or all of their electrons. Ultraviolet radiation ionizes the gas; the ionized gas in turn emits microwave radiation from collisions between electrons and protons (hydrogen nuclei). Some of the microwave radiation is caused by an entirely different mechanism called synchrotron radiation. In this case, high-energy electrons spiral around magnetic fields and radiate as they move. Synchrotron radiation is emitted from a globular region near the galactic center; this structure is probably a relic of a recent explosion that may have exceeded the power of a supernova.

Detailed measurements of radio-wave intensity in the innermost region of the galaxy (within 10 light-years of the center) reveal that much of the ionized material there is arranged in streamers or arcs, which have been interpreted as material either being ejected from the center or falling inward from an outlying orbit. The streamers are rather patchy; there are also somewhat more separated, blob-like clouds of ionized gas [*see illustration on opposite page*].

A much larger radio view of the galactic center shows extensive narrow arcs about 200 light-years long, discovered by Mark Morris, Farhad Yusef-Zadeh and Donald R. Chance working with the VLA antennas. The long, thin streamers must glow owing to synchrotron radiation, and the magnetic fields associated with the radiation must be oriented along the direction of the streamers, approximately perpendicular to the plane of the galaxy. Investigators have not yet been able to measure accurately the strength of the magnetic fields, but their intensity could be as high as one milligauss, about 1/1,000 of the strength of the earth's magnetic field at the surface. High-energy particles moving in magnetic fields near the galactic center, inferred from the synchrotron radiation, should also produce X rays. Indeed, fairly intense X rays are observed in this region.

Additional streamers and arcs nearby may also be caused by magnetic fields, but their shape is not as clearly indicative of such an origin. We, along with Gordon J. Stacey and Andrew I. Harris, have studied the ionization of these streamers from the airborne *Kuiper Astronomical Observatory* and have found that some of the more prominent arcs, closer to the galactic center than the longest streamers, must be the result of gas ionized by ultraviolet radiation.

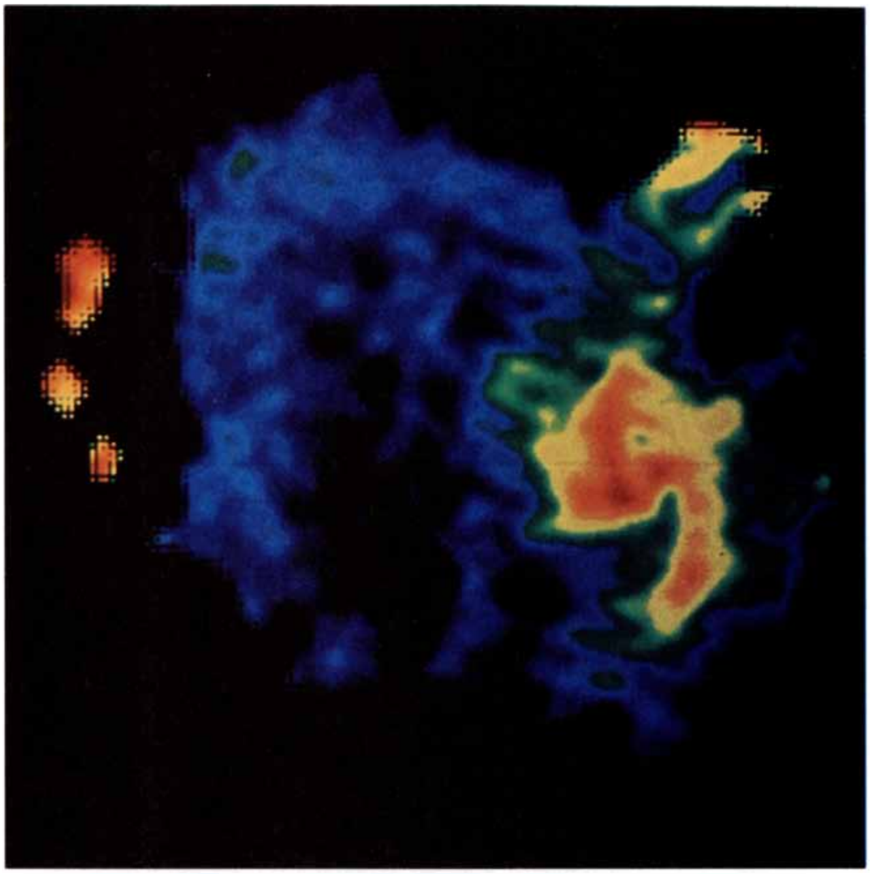
Examining the central region on a still larger scale reveals further striking features, most notably an elongated mass of ionized gas about 600 light-years long and only a few tens of light-years across that projects out of the galactic plane and is now commonly referred to as the galactic lobe. It is surrounded by even longer, although somewhat less distinct, features. These lobes project roughly perpendicularly



OBJECTS EMIT ELECTROMAGNETIC RADIATION at wavelengths characteristic of their temperature or energy state. In general, the higher the temperature, the shorter the wavelength at which a body radiates (*top*). Various molecular, atomic and nuclear processes produce radiation at distinctive wavelengths, which reveal the composition and physical conditions of the radiating source (*middle*). Unfortunately, the earth's atmosphere is transparent only to some of these wavelengths; to view the others, instruments must be lofted by balloons, airplanes or rockets (*bottom*).

from the plane of the galaxy, a hint that they might have been ejected from the central region or might even be falling into it from some unknown high-latitude source.

A very small object almost precisely at the galactic center, called Sagittarius A* (or Sgr A*), is the source of particularly intense radio emission. Its size, intensity and relatively constant radiation make it unique among the objects known in the Milky Way. Very-long-baseline radio interferometry by Kwok-Yung Lo and Donald C. Backer of the University of California at Berkeley and other co-workers has shown this object extends no more than about 1/1,000 of one arc second (one arc second is 1/3,600 of a degree). At the distance of the galactic center—approximately 25,000 light-years—this angle corresponds to a diameter of about 1.2 billion kilometers, roughly the size of Jupiter's orbit around the sun. Sgr A* may be even smaller than this, because gas and dust clouds in the galaxy tend to distort and smear distant images before they reach the earth; the apparent size may be the result of such smearing.



RADIO EMISSION from the immediate vicinity of the galactic center hints at a violent past. Energetic ultraviolet radiation from the source near the central region ionizes hydrogen atoms, splitting them into their electron and proton components. Red and yellow depict regions emitting microwaves as a result of collisions between electrons and protons; the deeper the red, the more intense the radiation. Blue colors depict weaker microwave radiation generated by energetic electrons, probably accelerated by a supernova explosion, as they spiral around magnetic fields. The globular shape of the blue region probably is a relic of an enormous explosion near the galactic center, perhaps exceeding even the most powerful supernova.

Infrared radiation from the galactic center was discovered in 1967 as a result of a more planned search than the one that led to Jansky's discovery of galactic radio emission. Gerry Neugebauer and co-workers at the California Institute of Technology were using infrared detectors sensitive to wavelengths near 2.2 microns (some three to four times longer than visible light waves) to examine infrared astronomical sources such as cool giant stars and dust clouds around them. Eric E. Becklin, then a graduate student at Caltech, thought it would be interesting to turn the detectors toward the region of the galactic center to see if any infrared signals would show up. He discovered the region is indeed an intense source of short-wavelength infrared rays. Becklin and Neugebauer mapped the infrared radiation and found it peaked strongly just about where the center was thought to be.

In more recent years astronomers have observed the galactic center at many infrared wavelengths in order to study the tightly packed stars, dust warmed by the stars, and the surrounding gaseous material. Infrared spectra of the gas contain the distinctive emission lines of various molecules, ions and atoms; consequently, it is possible to determine the motions, states of excitation and abundances of these substances.

Detailed information about the galactic center has been obtained from a diverse assortment of recently developed instruments. Ground-based instruments examine wavelengths transmitted by the earth's atmosphere, such as the 2.2-micron radiation observed by Becklin and Neugebauer or the wavelength range near 10 microns, which we have studied extensively. Many infrared wavelengths are absorbed by the atmosphere and can be recorded only by instruments lifted to high altitudes. Such instruments can be placed on earth-orbiting satellites (the *Infrared Astronomical Satellite,* for example), suborbital rockets and balloons. A particularly valuable instrument has been the National Aeronautics and Space Administration's *Kuiper Astronomical Observatory*—a high-flying C-141 airplane that houses a gyroscopically stabilized 36-inch telescope. From this observatory Becklin, Ian Gatley and Michael W. Werner observed infrared radiation in the 30- to 100-micron range and detected a doughnut-shaped ring of warm dust roughly 10 light-years wide surrounding the center of the galaxy and thought to be rotating around it.

Observations of energetic gamma rays and X rays have turned up some surprises regarding the behavior of the galactic center. Gamma rays from the direction of the galactic center were first detected in 1970 by means of instruments lofted into the stratosphere by large helium balloons. In 1977 Marvin Leventhal of Bell Labs and his associates measured these gamma rays and found them to have an energy of 511,000 electron volts (for comparison, visible light has an energy of about two electron volts), precisely the energy produced by the mutual annihilation of an electron and a positron, its antimatter twin. The intensity of this radiation emanating from the galactic center corresponded to the annihilation of vast quantities of antimatter—

about 10 billion tons of positrons each second. Balloon flights since 1977 have shown that this intensity has changed dramatically; over eight years it decreased by a factor of about four, dipping below the detectable limit, and then returned to its full intensity.

The amount and extreme variability of the annihilation radiation indicate a source unlike any other known in the galaxy. The source appears to be somewhere near the very center, but its exact location remains uncertain. A more definitive determination of the origin of this puzzling radiation will have to await the launching of a space-based instrument capable of locating energetic gamma-ray sources quite precisely.

Annihilation radiation can be produced in the vicinity of collapsed, extremely dense objects known as neutron stars. A neutron star orbiting a normal star can tear off material from its companion, creating outbursts of energy, and perhaps of antimatter and annihilation radiation, as the material is pulled down to the surface. This kind of radiation can also be produced by an even more collapsed object known as a black hole, about which we shall have more to say later.

More energetic gamma rays, with an energy of 1.8 million electron volts, have also been detected emanating from the galactic center by William A. Mahoney, Allen S. Jacobson and their colleagues at the Jet Propulsion Laboratory in Pasadena, Calif. This energy is characteristic of radiation emitted during the decay of aluminum 26, a radioactive variant of normal aluminum. As with the annihilation radiation, the total amount of matter involved is startlingly large: the mass of radioactive aluminum in the region seems to equal the mass of several stars. Current understanding has it that aluminum 26 is produced only in violent stellar explosions such as novas and supernovas or in massive, hot and fiercely active objects known as Wolf-Rayet stars. Even in these stars, aluminum 26 is produced in rather small quantities; perhaps a large number of supernovas in the vicinity of the galactic center have produced the stunning amount of aluminum 26.

DENSE SWARMS OF STARS near the galactic center are obscured by thick clouds of gas and dust. Infrared rays with a wavelength of about two microns partially penetrate the clouds and reveal the otherwise hidden distribution of stars. This view was obtained with a newly designed infrared detector array. The calculated mass of the observed stars is inadequate to explain their rapid motions near the center.

In addition to the radiations created by specific, identifiable processes, the galactic central region emits substantial X-ray continuum radiation. This year two important space-borne X-ray and gamma-ray observatories, the *Gamma Ray Observatory* (GRO) and the *Roentgen Satellite* (ROSAT), will be launched and will help pinpoint the sources of X rays and radiation from antimatter and aluminum 26 in the active region around the galactic center.

Although the discoveries made in the X-ray and gamma-ray bands are significant, the most extensive and revealing spectroscopic work on the galactic center has been performed at infrared and radio wavelengths, where the center was first observed. One extensively studied part of the radio spectrum has been the 21-centimeter emission from atomic hydrogen. Hydrogen is the most abundant atom in the universe, which compensates for the inherent weakness of its radiation. In the regions of the Milky Way where clouds of interstellar gas are not too dense and where ultraviolet radiation is not too intense, the hydrogen exists mostly as isolated, electrically neutral atoms; the distinctive radio signal of this atomic hydrogen has been carefully mapped to reveal the broad outlines of our galaxy.

More than about 1,000 light-years from the galactic center, atomic hydrogen emission provides a good picture of the rotation of the galaxy and the structure of its spiral arms. It does not provide much information about conditions close to the center, because the hydrogen there tends to be either bound up in molecules or ionized (hydrogen splits into a proton and an electron).

The thick clouds of molecular hydrogen obscure the galactic center and indeed most distant objects in the plane of the galaxy, as seen from the earth. Fortunately, microwave and infrared telescopes can peer through the clouds and reveal both what they contain and what lies behind them in the center. In addition to hydrogen molecules, the clouds contain large quantities of the stable molecule carbon monoxide (CO), for which the longest characteristic wavelength is about three millimeters. This radiation penetrates the earth's atmosphere and can be measured by ground-based radio antennas; CO is plentiful in essentially all dark dust clouds and so is useful for mapping their size and densities. By measuring Doppler shifts—changes in the frequency and wavelength of a signal caused by the motion of an object emitting the signal toward or away from the viewer—it is also possible to measure the velocities of the clouds.

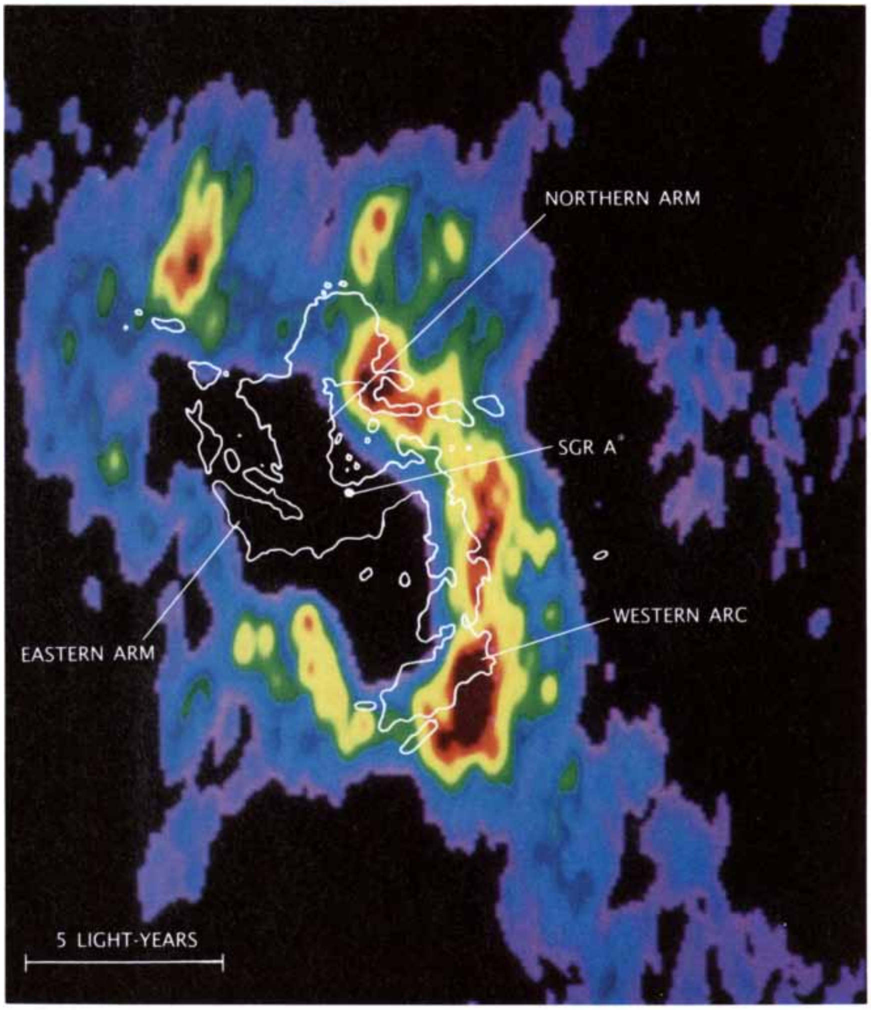
Dark clouds tend to be quite cool, about 15 kelvins, or -260 degrees Celsius; much of the CO in these clouds therefore lies in low energy states and emits at relatively low frequencies in the millimeter range. Some material near the galactic center is clearly much warmer. From the *Kuiper Astronomical Observatory,* we, along with Dan M. Watson, John W. V. Storey, John B. Lugten and other colleagues at the University of California at Berkeley, have detected more energetic CO emissions in the far infrared range, which indicate temperatures near 400 kelvins, approximately the boiling point of water. This gas is heated by ultraviolet radiation from the galactic center and perhaps also by shock waves generated when clouds moving about the galactic center collide. Ultraviolet radiation penetrates the clumpy clouds and warms the gas in the vicinity. Therefore, temperatures decrease with increasing distance from the center and are much higher than the 10 to 15 kelvins typical of clouds not associated with the center.

Spectra of oxygen atoms and ionized carbon provide further information on temperatures and densities of the gas. Other molecules that have been heavily studied include hydrogen cyanide (HCN), hydroxyl (OH), carbon monosulfide (CS) and ammonia (NH_3). A high-resolution map of HCN emission has been obtained with the University of California's radio interferometer by our group working at Berkeley, including Melvyn C. H. Wright and Rolf Güsten. It provides a clear picture of a rather clumpy and irregular disk of warm molecular clouds surrounding an evacuated hole, or cavity, about 10 light-years wide in the center of the galaxy. Since the disk is tilted with respect to the line of sight to the earth, the circular hole appears elliptical [*see illustration at right*].

Atoms of carbon and oxygen, some of them ionized by ultraviolet light, are mixed with the molecular gas in the disk. Maps of infrared and radio emission lines from these ions and atoms, and from various molecules, show that the gas disk is circulating around the galactic center at a velocity of about 110 kilometers per second and that the gas is warm and clumpy. The measurements also reveal the presence of some clouds whose motions do not at all fit this general pattern of circulation, perhaps material that has fallen in from a distance. Ultraviolet radiation from the central region strikes the inner rim of the disk of clouds, creating almost a full circle of ionized material; there are also ionized streamers and clumps of gas inside the central cavity. Lo, Robert L. Brown and Ronald D. Ekers of the National Radio Astronomy Observatory and others have used the VLA antennas to map these structures in beautiful detail.

Some fairly abundant ionized elements—including neon with one electron removed, argon with two electrons removed and sulfur with three electrons removed—emit prominent spectral lines at wavelengths of near 10 microns, a portion of the infrared band that penetrates the atmosphere and can be observed from the earth's surface. Beginning as graduate students at the University of California at Berkeley, John H. Lacy, Eugene Serabyn, Eric R. Wollman and Thomas R. Geballe have for some years been probing these wavelengths with sensitive spectrometers. They have found that singly ionized neon is by far the most prevalent of these elements and that triply ionized sulfur is almost nonexistent in the material near the center. Removing three electrons from sulfur requires far more energy than stripping a single electron from neon; the composition of the material indicates that ultraviolet radiation is plentiful in the central region but that the radiation is only moderately energetic. The radiation was therefore probably produced by hot stars with temperatures of from 30,000 to 35,000 kelvins, and stars substantially hotter than this cannot be present.

Spectroscopic examination of radiation from ions has also provided de-

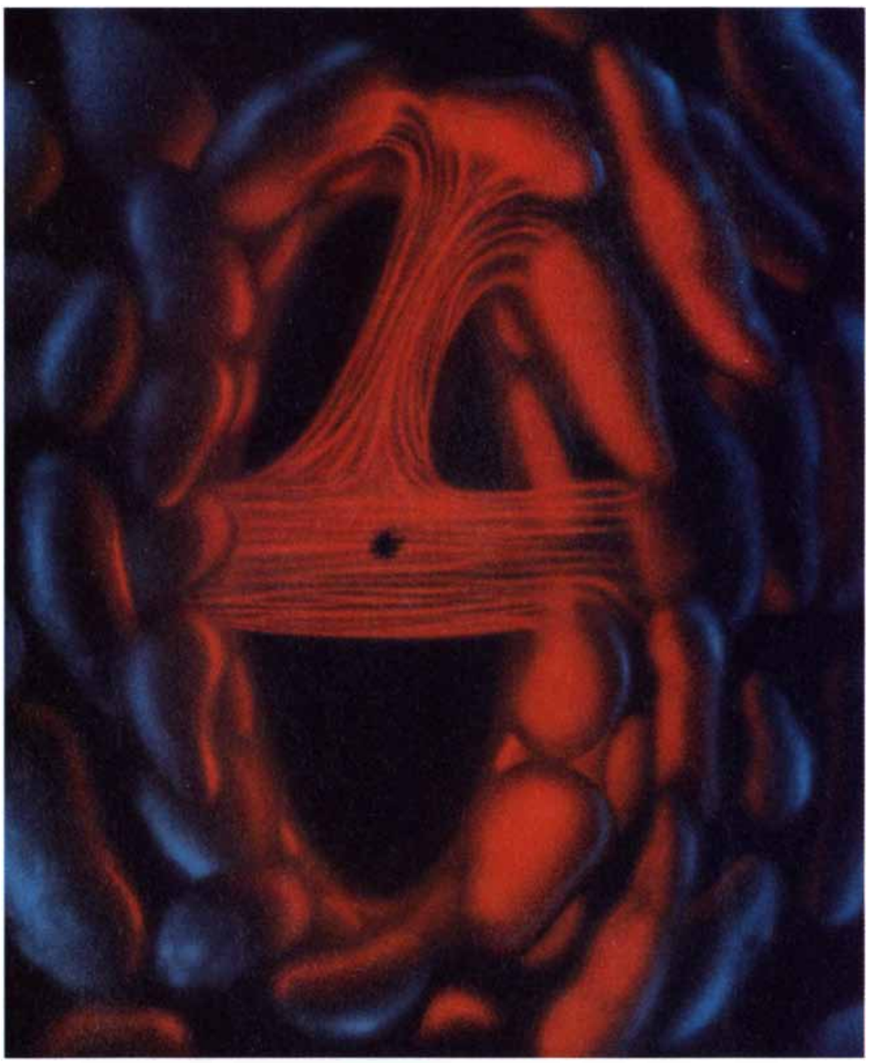


HYDROGEN CYANIDE MOLECULES in warm, dense gas clouds near the galactic center emit short-wavelength radio emissions, shown on this map made with the University of California's millimeter-wavelength telescope. Colder gas farther out is not visible at these wavelengths. An elliptical distribution of gas with a hollowed-out center lies at the heart of the galaxy. The gas is clumpy and varies in its intensity of radiation, indicating that it has been disturbed recently, most likely within the past 100,000 years. Collisions between clumps will eventually smooth out the gas.

tailed information about the velocity of the tenuous material within the 10 light-year-wide cavity surrounding the center. In some parts of the cavity, the velocities are similar to those of the circulating ring of molecular gas, about 110 kilometers per second. A number of clouds in the inner part of the region are moving significantly faster, roughly 250 kilometers per second, and some have velocities as high as 400 kilometers per second.

At the very center, Donald N. Hall of the University of Hawaii, Geballe and their colleagues have detected ionized material with velocities of up to 1,000 kilometers per second. This material is associated with an interesting set of objects near the center of the cavity known as IRS16, discovered by Becklin and Neugebauer during a search for sources of short-wavelength infrared waves. Most of the tiny sources they found appear to be single massive stars, but IRS16 (their 16th infrared source) is different: it has since been resolved into approximately five bright and unusual components.

This whole central region, both the warm disk of gas and the inner cavity, appears to have been the scene of some violent and relatively recent disturbance. A ring or disk of gas rotating about the galactic center should eventually settle down into a smooth structure as a result of collisions between any faster- and slower-moving clumps of material within the disk. Doppler shifts reveal velocity differences of some tens of kilometers per second between clumps in the ring of molecular gas. These clumps should collide and smooth each other on a time scale of about 100,000 years, or within one or two rotations about the center. Evidently the gas has been severely disturbed within that time—possibly by an outburst of energy from the center or by an infall of material from some distance—and collisions between clumps should still be violent enough to produce shock waves in the gas. The accuracy of these conclusions can be tested by searching for signs of shock waves.



STREAMERS, FILAMENTS AND RINGS of gas and dust surround the center of the Milky Way. Atoms in the inner regions of the ring are ionized and energized, primarily by ultraviolet radiation from the inner central region that strikes the near edges of the many thick clouds (*red areas*). The outer clouds are shadowed and so contain relatively undisturbed atoms and molecules. The ring orbits about the center; faster- and slower-moving clumps will eventually collide with each other. Some gas appears to be pulled off the ring and into the center, but the origin and orbits of the horizontal gas filaments are unclear. A massive black hole at the center may have triggered an explosive outburst in the past that helped to form this structure.

Shock waves can be identified by the resulting spectral lines from hot, highly excited molecules. Such molecules have been observed from the *Kuiper Astronomical Observatory;* they include hydroxyl radicals—fragments of water molecules that have been energetically torn apart. Short-wavelength infrared radiation from hot hydrogen molecules has also been found and shows that in some locations the temperatures in the molecular gas clouds rise as high as 2,000 kelvins, a temperature that would in fact be created by shocks.

What is the source of the dense molecular dust clouds near the center? The material contains heavy elements, indicating that it has been processed in the interiors of stars, where nuclear reactions fuse hydrogen and helium into heavier elements such as carbon, oxygen and nitrogen. Old stars expand and shed large amounts of material or, in some cases, explode as supernovas; either way, heavy elements are injected into interstellar space. The clouds near the center seem to have been more heavily processed inside stars than the material farther away from the center, because certain rare isotopes that form only within stars are quite abundant in this region.

Not all of this material was necessarily produced by earlier stars in the immediate vicinity of the center; possibly some of the clouds have fallen inward. Friction and the galactic magnetic field gradually drag material toward the center, and one might expect that a large amount of material would accumulate there. We shall consider later what may become of this material.

Even though stars near the galactic center cannot be observed optically, they can be detected at infrared wavelengths of two to three microns, slightly longer than visible light. These wavelengths, characteristic of objects at stellar temperatures, are scattered and absorbed by the dust between the earth and the center in the same way that light is. Still, enough infrared can pass through to make it possible to measure the concentration of the stars in the region; this is the radiation that Becklin and Neugebauer discovered when they pointed their infrared instruments toward the galactic center.

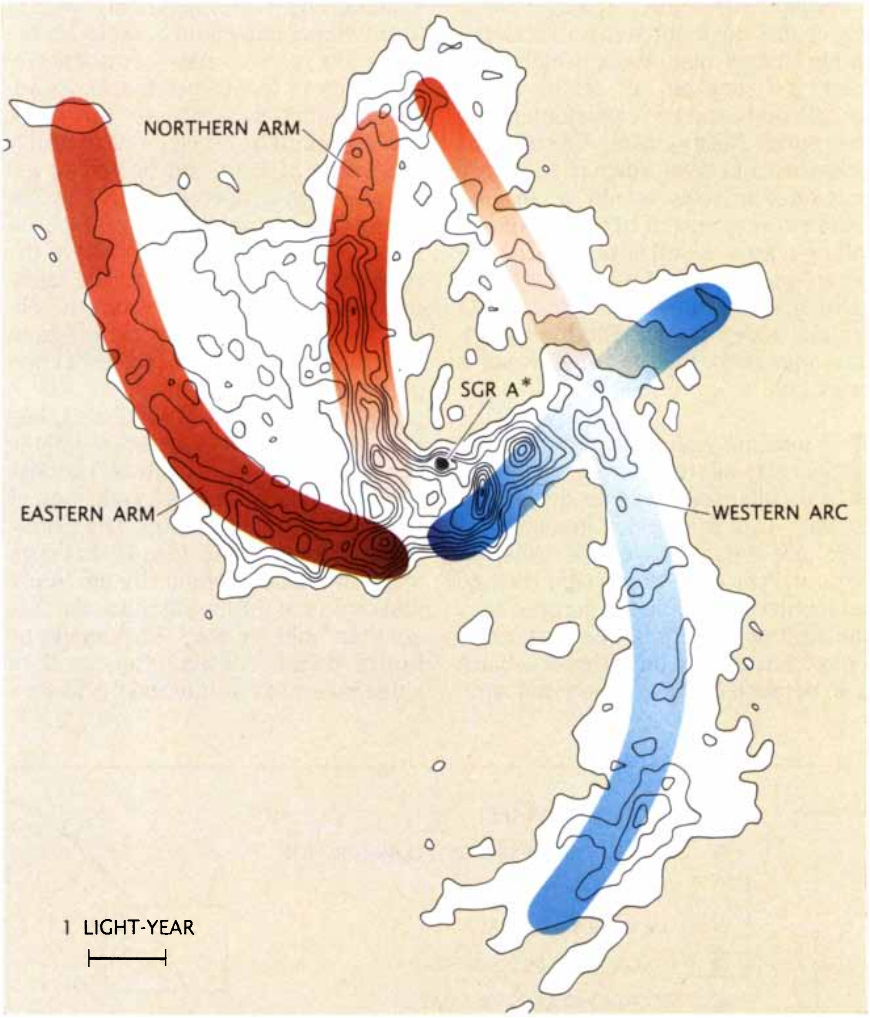
Such measurements show that in the inner region of the galaxy the density of stars is quite high and increases toward the center. The average distance between stars is perhaps 1/300 of the distance between the sun and its nearest neighbor; under these conditions, every 100 million years or so a star will approach its neighbors closely enough so that they will deflect each other's paths and share their kinetic energy, or energy of motion. As a result, the probable velocity of any star in the central cluster must be the same as that of any other star.

For all the stars to have the same probable velocity and to be held together by their mutual gravitational field, they must follow a particular distribution. In the ideal case, stellar density is proportional to $1/R^2$ where R is the radius (the distance from the center). The total mass of stars inside a given radius is then simply proportional to R, and the gravitational field is inversely proportional to the distance from the center.

From measurements of their short-wavelength infrared emission, it is clear that the dense cluster of stars in the center of the galaxy roughly follows the idealized theoretical distribution; hence, one would expect stars and other objects in this region to orbit around the center, all at rather similar velocities. In fact, the velocities of the gas motions, particularly those of the ionized gases inside the central cavity, increase strikingly toward the very center of the galaxy. It seems that the stars detectable in the infrared do not account for all of the mass responsible for the gravitational field. This finding provided the first reasonably clear evidence for the existence of a massive but undetected object near the center. Various investigators have suggested nongravitational forces that might explain the high velocities, such as magnetic fields or interstellar winds. Increasingly detailed knowledge of the conditions near the center indicates, however, that none of these proposed forces would be adequate.

Astronomers have recently been able to measure stellar velocities in this region. Such measurements are possible because many of the stars near the center are red giants, intrinsically bright stars that radiate strongly in short-wavelength infrared and that have a good deal of CO in their atmospheres. The distinctive CO spectra from these stars can be examined to determine their motions. Such measurements have been carried out by George H. Rieke and Marcia J. Rieke of the University of Arizona. Particularly clear, recent measurements by Kristen Sellgren, Martina T. McGinn and their colleagues working on Mauna Kea in Hawaii show that stellar velocities, like gas velocities, increase toward the center of the galaxy. The motions of stars and gas behave as if the region within one or two light-years from the center encompasses a mass of three to four million suns—considerably more than what could be expected from stars alone.

From measurements of the material now falling into the galactic center, it is possible to calculate how much mass should have collected there. Matter seems to fall into the center by diffusion and by intercloud collisions at a rate of about one solar mass every 1,000 years. If the galaxy has been in something like its present condition for five billion years, which is perhaps a reasonable estimate, and if material has been accumulating throughout that time, the total infall should have amounted to about five million solar masses—close to the three to four million solar masses estimated from observed gravitational pulls.

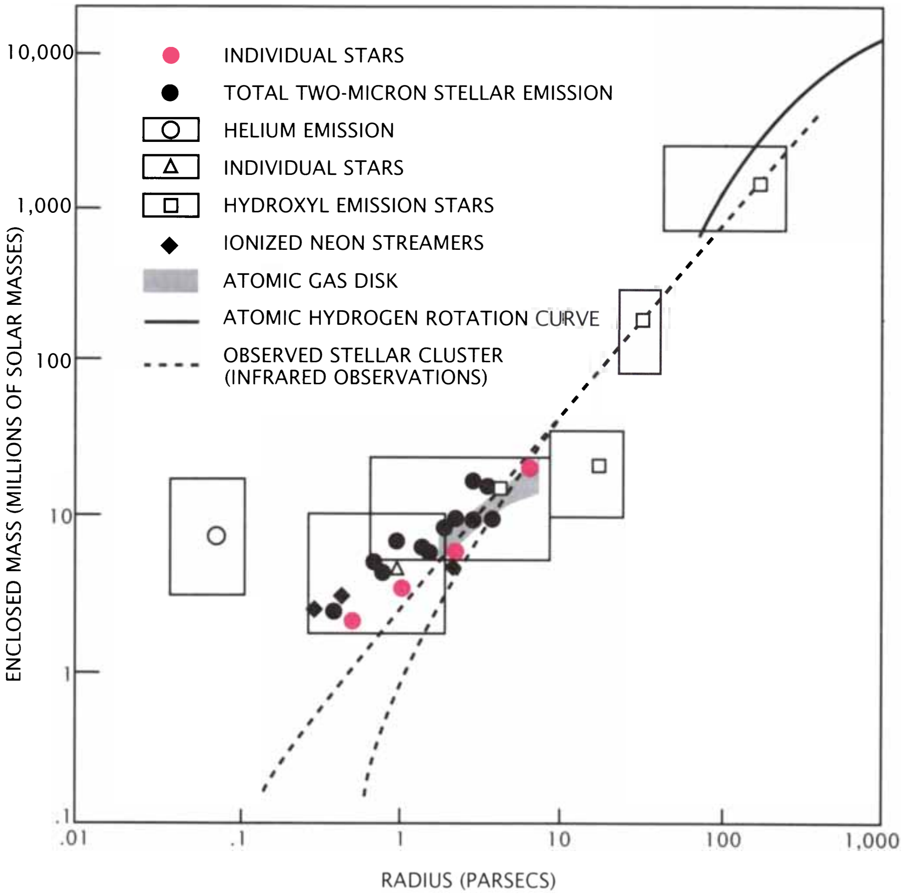


VELOCITIES IN THE GAS RING near the center are determined by measuring shifts in the radiation emitted there. Motion toward the earth makes the wavelength shorter (*bluer*); motion away makes it longer (*redder*). The clouds rotate at about 110 kilometers per second; some clumps move from 10 to 20 kilometers per second faster or slower. The map of radio intensity shows streamers of gas either being ejected from the region or falling into it from outside. The intense radio source very near the center is the mysterious Sagittarius A*. Its small size, prodigous output and generally constant radiation make it unique among the objects in the galaxy.

Perhaps the matter falling toward the center does not remain there but rather forms new stars, which then might be knocked out of the center by collisions and be redistributed into the central stellar cluster. Calculations alone cannot reveal which of these two outcomes actually would occur. But there is every reason to believe the infalling matter is still sitting at the center of the galaxy. How could such a great mass accumulate there and yet remain undetectable? The most obvious answer is that it has formed a black hole.

Einstein's equations of general relativity allow for the existence of a collapsed mass so dense that nothing can escape its gravitational field, not even light. John A. Wheeler, then at Princeton University, dubbed such objects black holes because they can emit no radiation. The center of a galaxy is a natural location for a black hole because it is a gravitational sink; surrounding matter inevitably collects at the center and could conceivably acquire the critical mass and density necessary to form a black hole. As additional matter falls into the black hole, it would be accelerated to nearly the speed of light and heated to extremely high temperatures by friction. A black hole could thus produce intense light and radio emissions of the type observed from Sgr A*. The radiation would emanate not from the object itself, which remains cut off from the rest of the universe, but from energetic material falling into it.

Could other kinds of objects explain the observed characteristics of the center? Some investigators propose that the unseen mass is a collection of dark stars near the center. The primary problem with this idea is that dark stars are characteristically no more massive (and often significantly less so) than brighter stars. They might be brown dwarfs, objects too small to shine by means of internal nuclear reactions, or perhaps neutron stars, the collapsed remains of old stars, that are no longer radiating.



GRAVITATIONAL TUGS near the galactic center indicate the amount of material present; the strength of gravity is inferred from the orbital motions of stars and clouds. Measurements of these motions are plotted along with those that would be expected (*broken lines*) if all of the mass were accounted for by the stars observed in the infrared. At large distances, theory and observation agree, but within a few light-years of the center, velocities appear to be anomalously high. Objects there are being pulled by some massive yet invisible object—possibly a black hole.

The gravitational sink at the center of the galaxy pulls heavier objects more strongly than light ones. Near collisions between lighter and heavier objects will eventually eject the former from this region; low-mass dark stars therefore should have been evicted by brighter, more massive stars. Even neutron stars, which can have masses comparable to those of normal stars, should be as widely dispersed as the visible stars clustered around the center. In short, the only kinds of stars currently known that would not produce enough infrared radiation to betray their presence could not remain in the center for much longer than the characteristic time between stellar encounters there, about 100 million years; this is a short time compared with the age of the galaxy. According to current understanding, the only object that could exert the observed gravitational pull and yet be otherwise undetected is a black hole of about three million solar masses.

The black hole at the center could be quite small despite its great mass. A black hole with a mass of about three million normal stars would have a "radius" (the largest distance from which nothing can escape) merely the same as that of an ordinary star like the sun. It would be an unspectacular minor object in size, and in the swarm of stars near the galactic center, it might be considered a needle in a haystack. One might expect that at least some material would always be trickling in toward the central mass, radiating strongly as it falls; one could then hope to locate the black hole by searching for the particular emissions expected from this infalling matter.

Occasionally a star might pass close to the black hole; the star would be torn apart by tidal forces, and some of its matter would form an inwardly spiraling disk of gas. According to theoretical calculations, antimatter can form as material falls into a black hole, and this antimatter may be the source of the annihilation radiation emanating from the central region. The infalling material can also produce intense radio radiation, and so the compact but powerful radio source Sgr A* seems to be a likely candidate for the black hole. On the other hand, both radio emission and annihilation radiation can also be created around a neutron star. The most readily observable distinction between a neutron star and the inferred black hole is the enormous difference in mass.

Backer and Richard A. Sramek of the National Radio Astronomy Observatory have observed Sgr A* for seven years and have shown that it has not moved more than about a tenth of its diameter during this time; unless it is moving along the line of sight toward the earth, its velocity is substantially less than that of most of the other objects near the galactic center. This finding suggests that Sgr A* is an extremely massive object, barely disturbed by gravitational tugs even in the dense center of the galaxy. Perhaps it is in fact the central mass: a heavy black hole about which other things revolve. It could also be the source of some of the annihilation radiation and the cause of the agitated motions in its surroundings. This supposition still leaves some phenomena unexplained, such as the intense ultraviolet radiation from the galactic center. If material falling into Sgr A* were responsible for this radiation, the object would also emit intense infrared radiation, which is not actually observed.

The putative large black hole at the center of the galaxy probably has been far more active in the past than it is now. The stellar density near the center is such that every few thousand years a star should pass close to the black hole and be torn apart by it. The resulting outburst should be spectacular but brief, lasting perhaps only a few years, according to calculations by Martin J. Rees of the University of Cambridge. A large central black hole could probably spend most of its life in a quiescent state—a state it would have to be in now—but past violent actions should have left some detectable remnants. Perhaps the central cavity is the result of an outburst that occurred some 10,000 to 100,000 years ago, and the clumpy, disturbed molecular gas surrounding it carries the memory of this violent episode.

Like the Milky Way, most galaxies seem to have fairly quiescent inner regions, but some exhibit powerful radio and infrared emissions and long jets of high-velocity material apparently ejected from the center. The quantity of energy released by active galaxies seems too great to have been generated by nuclear power alone, and so many astronomers think the energy is produced by material falling into a black hole.

At least two nearby galaxies, the well-known Andromeda galaxy and a smaller one in the constellation Triangulum, are now thought to contain black holes even though they are relatively quiet galaxies not unlike the Milky Way. Alan M. Dressler and Douglas O. Richstone, working at the Mt. Palomar Observatory in California, found that stellar velocities toward the center of these galaxies appear to increase more rapidly than would be expected, as if they also harbor large unseen masses. The distances to even these nearby galaxies are too great to permit the study of fine details of the state and motion of matter in the immediate vicinity of the center, as has been possible for the Milky Way.

Some nagging problems remain for scientists who seek a thorough understanding of what is going on in the center of our galaxy. Is the radio point source Sgr A* a black hole? Or is the massive black hole associated with IRS16 a collection of very hot objects close to Sgr A* that radiate intensely in the infrared? Or is something else happening that so far has given no outward signs other than the presence of a huge gravitational pull?

The extremely slow measured velocity of Sgr A* hints it may indeed be the massive black hole at the galactic center. If that is true, then what is IRS16? The five closely spaced components of this object seem to have temperatures typical of those of very hot stars, but they radiate with an intensity millions of times greater than that of the sun, far more than expected from normal hot stars. Furthermore, high-velocity gas in the vicinity may have been emitted by one of these objects. These characteristics are reminiscent of energetic but extremely rare Wolf-Rayet stars.

During an eclipse of the galactic center by the moon, Becklin and his associates watched the five components of IRS16 pass behind the rim of the moon and found they blinked out rapidly, revealing they are quite small and implying they must each be single objects and not clusters of stars. Either they are objects of a character never before seen or else perhaps they truly are Wolf-Rayet stars. But why would five Wolf-Rayet stars appear in this particular place, where relatively little gas exists to form new stars? Perhaps an appropriate collection of infalling gas accumulated there in the past; perhaps these objects represent the result of recent stellar collisions. Or perhaps astronomers simply do not understand these objects and must study them further to comprehend the activity in this intriguing and energy-intense region.

Recent astronomical discoveries and sophisticated new techniques for observing radio and infrared radiations have yielded striking and detailed information about the galactic center—the materials there, their motions and densities and temperatures, the amount and type of energy involved, and the numbers of stars and their patterns of motion. As a result, astronomers now have a good overall understanding of the heart of our galaxy and strong evidence of a massive black hole lurking there. Yet much remains to be learned.

A more definitive understanding will demand continued improvements in astronomical technologies and more advanced instruments, many already in construction or planning stages. Precise optical and infrared observations will be made using large new telescopes such as the 10-meter Keck telescope being built in Hawaii and the four eight-meter telescopes being built in Chile by the European Southern Observatory. Sophisticated new infrared interferometers will provide additional high-resolution detail about the distribution of stars, dust and warm gas in the center of our galaxy.

Powerful new instruments for space or high altitudes also are being either initiated or planned by NASA and by the European Space Agency. These include the *Stratospheric Observatory for Infrared Astronomy* (a more advanced version of the *Kuiper Astronomical Observatory*) and helium-cooled orbiting infrared telescopes that will far exceed the capabilities of previous detectors. This year will see the launch of GRO and ROSAT, two important new instruments that will measure X rays and gamma rays sensitively and with great precision.

In the half-century since Jansky's pioneering discovery, understanding of the galactic center has advanced remarkably. And each new discovery has brought with it new questions that have whet interest and appetites for more powerful instruments that will delve ever deeper into the mysteries at the heart of our galaxy.

FURTHER READING

THE GALACTIC CENTER. American Institute of Physics Conference Proceedings 155. Edited by Donald C. Backer. American Institute of Physics, 1987.

PHYSICAL CONDITIONS, DYNAMICS, AND MASS DISTRIBUTION IN THE CENTER OF THE GALAXY. R. Genzel and C. H. Townes in *Annual Review of Astronomy and Astrophysics*, Vol. 25, pages 377–423; 1987.

THE CENTER OF THE GALAXY. International Astronomical Union Symposium, No. 136. Edited by Mark Morris. Kluwer Academic Publishers, 1989.

The Unusual Origin of the Polymerase Chain Reaction

A surprisingly simple method for making unlimited copies of DNA fragments was conceived under unlikely circumstances—during a moonlit drive through the mountains of California

by Kary B. Mullis

Sometimes a good idea comes to you when you are not looking for it. Through an improbable combination of coincidences, naiveté and lucky mistakes, such a revelation came to me one Friday night in April, 1983, as I gripped the steering wheel of my car and snaked along a moonlit mountain road into northern California's redwood country. That was how I stumbled across a process that could make unlimited numbers of copies of genes, a process now known as the polymerase chain reaction (PCR).

Beginning with a single molecule of the genetic material DNA, the PCR can generate 100 billion similar molecules in an afternoon. The reaction is easy to execute: it requires no more than a test tube, a few simple reagents and a source of heat. The DNA sample that one wishes to copy can be pure, or it can be a minute part of an extremely complex mixture of biological materials. The DNA may come from a hospital tissue specimen, from a single human hair, from a drop of dried blood at the scene of a crime, from the tissues of a mummified brain or from a 40,000-year-old woolly mammoth frozen in a glacier.

In the seven years since that night, applications for the PCR have spread throughout the biological sciences: more than 1,000 reports of its use have been published. Given the impact of the PCR on biological research and its conceptual simplicity, the fact that it lay unrecognized for more than 15 years after all the elements for its implementation were available strikes many observers as uncanny.

The polymerase chain reaction makes life much easier for molecular biologists: it gives them as much of a particular DNA as they want. Casual discussions of DNA molecules sometimes make them sound like easily obtained objects. The truth is that in practice it is difficult to get a well-defined molecule of natural DNA from any organism except extremely simple viruses.

The difficulty resides in the nature of the molecule. DNA is a delicate chain made of four deoxynucleotides: deoxyadenylate (*A*), deoxythymidylate (*T*), deoxyguanylate (*G*) and deoxycytidylate (*C*); the sequence of these bases encodes the genetic information. Rarely does one find a single strand of DNA; usually pairs of strands with complementary sequences form double helixes in which the *A*'s in one strand bind with the *T*'s in the other, and the *G*'s bind with the *C*'s [*see illustration on opposite page*]. Inside a cell this DNA helix is surrounded and further coiled by various proteins. When biologists try to isolate a naked DNA chain, the DNA is so long and thin that even mild shearing forces break it at random points along its length. Consequently, if the DNA is removed from 1,000 identical cells, there will be 1,000 copies of any given gene, but each copy will be on a DNA fragment of differing length.

For years this problem made it difficult to study genes. Then in the 1970's enzymes known as restriction endonucleases were discovered: these enzymes snipped strands of DNA at specific points. The endonucleases made it possible to cut DNA into smaller, sturdier, more identifiable pieces and thereby made it easier to isolate the pieces containing a gene of interest.

By the late 1970's, therefore, molecular biologists were busily studying DNA with endonucleases and with other molecules called oligonucleotide probes. An oligonucleotide is a short chain of specifically ordered nucleotide bases. Under the right conditions, an oligonucleotide will bind specifically with a complementary sequence of nucleotides in single-strand DNA. Therefore, radioactively labeled, man-made oligonucleotides can serve as probes for determining whether a sample of DNA contains a specific nucleotide sequence or gene. In 1979 the Cetus Corporation in Emeryville, Calif., hired me to synthesize oligonucleotide probes.

By 1983 the charm of synthesizing oligonucleotides for a living had entered a decline—a decline that most of us so employed were happy to witness. The laborious but very quaint chemical art form for making oligonucleotides manually, to which we had grown comfortably numb, had given way to a much less charming but reliable automated technique. It was an immense improvement.

In the aftermath of this minor industrial revolution, we nucleotide chemists found ourselves success-

KARY B. MULLIS describes himself as "a generalist with a chemical prejudice." In addition to the polymerase chain reaction, he is also known for having invented a plastic that changes color rapidly when exposed to ultraviolet light. While working as a biochemistry graduate student at the University of California, Berkeley, he published a paper in *Nature* entitled "The Cosmological Significance of Time Reversal." Mullis received his Ph.D. in biochemistry in 1972. After working as a postdoctoral fellow at the University of Kansas Medical School and the University of California, San Francisco, Mullis joined the Cetus Corporation, where he discovered the polymerase chain reaction. In 1986 he became the director of molecular biology at Xytronyx, Inc. Today Mullis works in La Jolla, Calif., as a private consultant on polymerase-chain-reaction technology and nucleic acid chemistry.

fully underemployed. Laboratory machines, which we loaded and watched, were making almost more oligonucleotides than we had room for in the freezer and certainly more than the molecular biologists—who seemed to be working even more slowly and tediously than we had previously suspected—could use in their experiments. Consequently, in my laboratory at Cetus, there was a fair amount of time available to think and to putter.

I found myself puttering around with oligonucleotides.

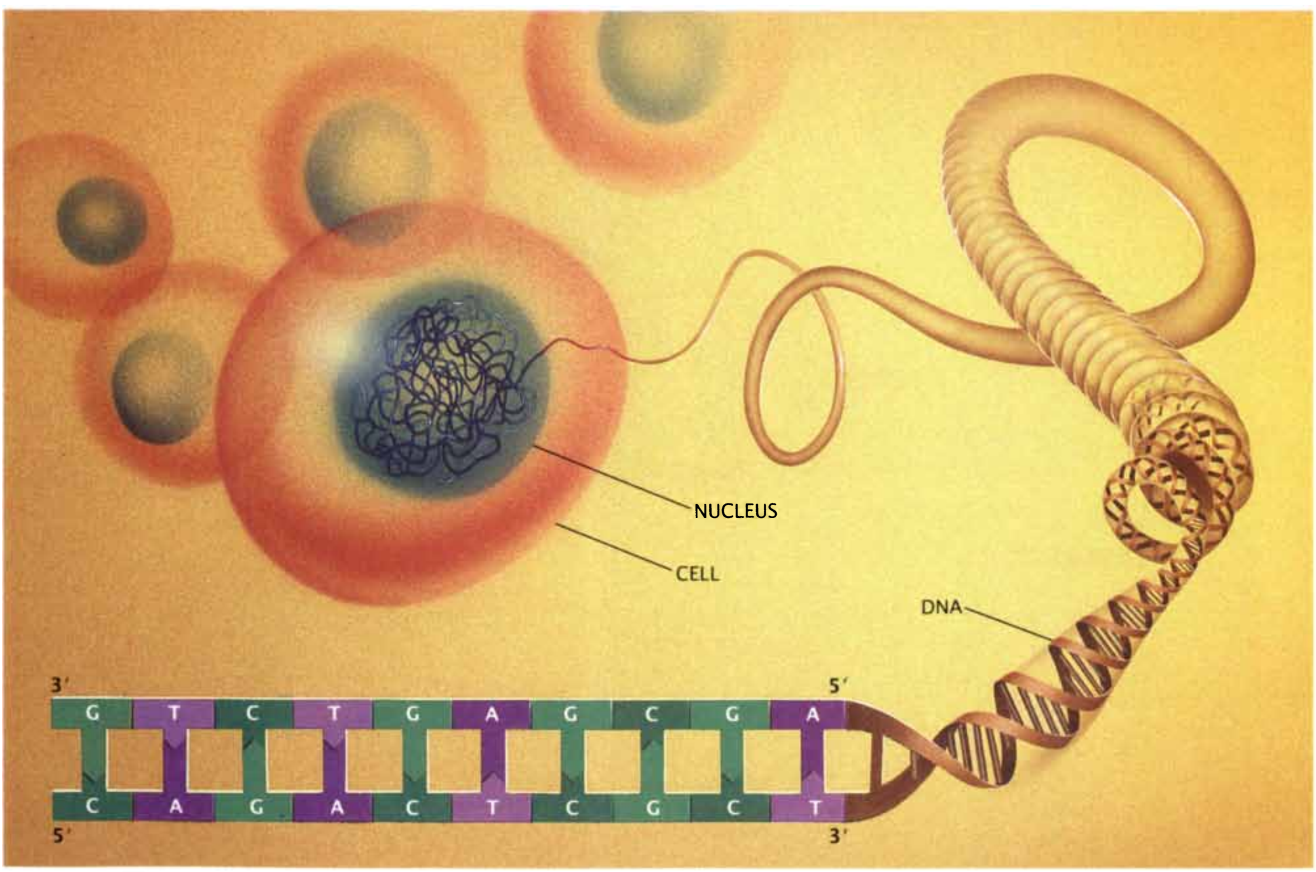
I knew that a technique for easily determining the identity of the nucleotide at a given position in a DNA molecule would be useful, especially if it would work when the complexity of the DNA was high (as it is in human DNA) and when the available quantity of the DNA was small. I did not see why one could not use the enzyme DNA polymerase and a variation of a technique called dideoxy sequencing, and so I designed a simple-minded experiment to test the idea.

To understand the approach I had in mind, it is worth reviewing certain facts about DNA. A strand of the molecule has one end that is known, by chemical convention, as three-prime and one end that is five-prime. In a double helix of DNA, the complementary strands are said to be antiparallel, because the three-prime end of one strand pairs with the five-prime of the other strand, and vice versa.

In 1955 Arthur Kornberg of Stanford University and his associates discovered a cellular enzyme called a DNA polymerase. DNA polymerases serve several natural functions, including the repair and replication of DNA. These enzymes can lengthen a short oligonucleotide "primer" by attaching an additional nucleotide to its three-prime end, but only if the primer is hybridized, or bound, to a complementary strand called the template. The surrounding solution must also contain nucleotide triphosphate molecules as building blocks.

The nucleotide that the polymerase attaches will be complementary to the base in the corresponding position on the template strand. For example, if the adjacent template nucleotide is an *A*, the polymerase attaches a *T* base; if the template nucleotide is a *G*, the enzyme attaches a *C*. By repeating this process, the polymerase can extend the primer's three-prime end all the way to the template's five-prime terminus [*see illustration on page 59*]. In a double helix of DNA, each strand serves as a template for the other during replication and repair.

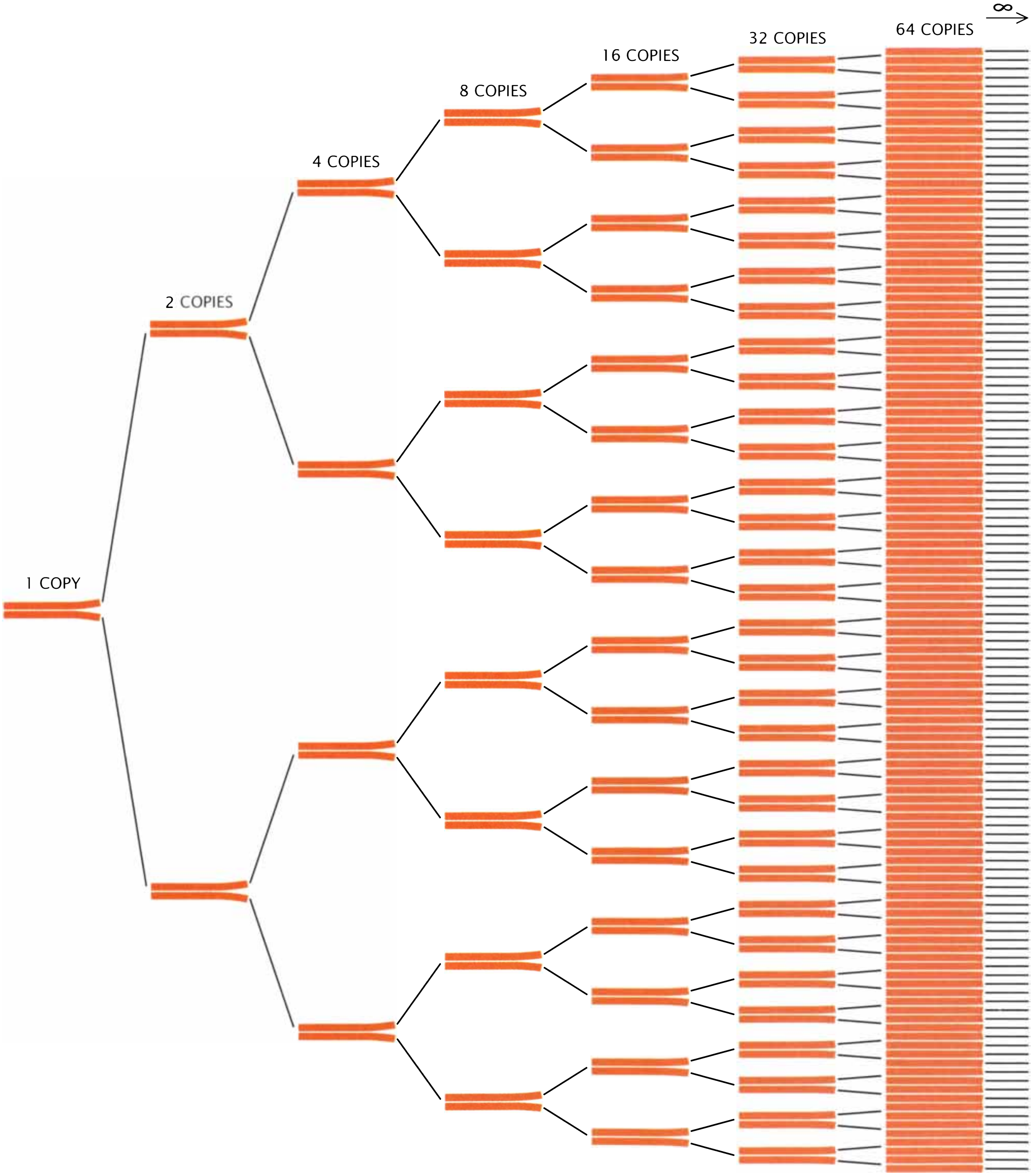
Now for dideoxy sequencing, which is also commonly called the Sanger technique after one of its inventors, Frederick Sanger of the British Medical Research Council Laboratory of Molecular Biology. This technique uses a DNA polymerase, template strands, primers, nucleotide triphosphates and special dideoxynucleotide triphosphates (ddNTP's) to determine DNA sequences. Like ordinary nucleotides, ddNTP's can be attached to growing primers by polymerases; however, a ddNTP will "cap" the three-prime end of a primer and prevent the addition of any more bases. The Sanger technique produces primers that have been lengthened to varying extents and then capped by a ddNTP. By arranging these fragments according to length and by knowing which ddNTP's



DNA consists of two strands of linked nucleotides: deoxyadenylates (*A*'s), deoxythymidylates (*T*'s), deoxyguanylates (*G*'s) and deoxycytidylates (*C*'s). The sequence of nucleotides in one strand is complementary to that in the other strand—the *A*'s are always opposite *T*'s, and the *G*'s are opposite *C*'s—and this complementarity binds the strands together. Each strand has a three-prime and a five-prime end. Because their orientations oppose one another, the strands are said to be antiparallel.

have been added, an investigator can determine the sequence of bases in the template strand. For example, if a dideoxyadenine (ddA) base were added at a given position, the corresponding complementary base in the template would be a *T;* the addition of a dideoxyguanine (ddG) implies the presence of a *C* in the template.

In the modified version of this technique that I was contemplating, I would use only polymerases, templates, ddNTP's and primer molecules—that is, I would omit the ordinary nucleotide triphosphates from the mixture. Extension of the primers would therefore terminate immediately after the addition of one base from a ddNTP to the chain. If I knew which ddNTP had been added to the primers, I would also know the identity of the corresponding base in the template strand. In this way, I could deduce the identity of a base in the template



POLYMERASE CHAIN REACTION is a simple technique for copying a piece of DNA in the laboratory with readily available reagents. Because the number of copies increases exponentially, more than 100 billion can be made in only a few hours.

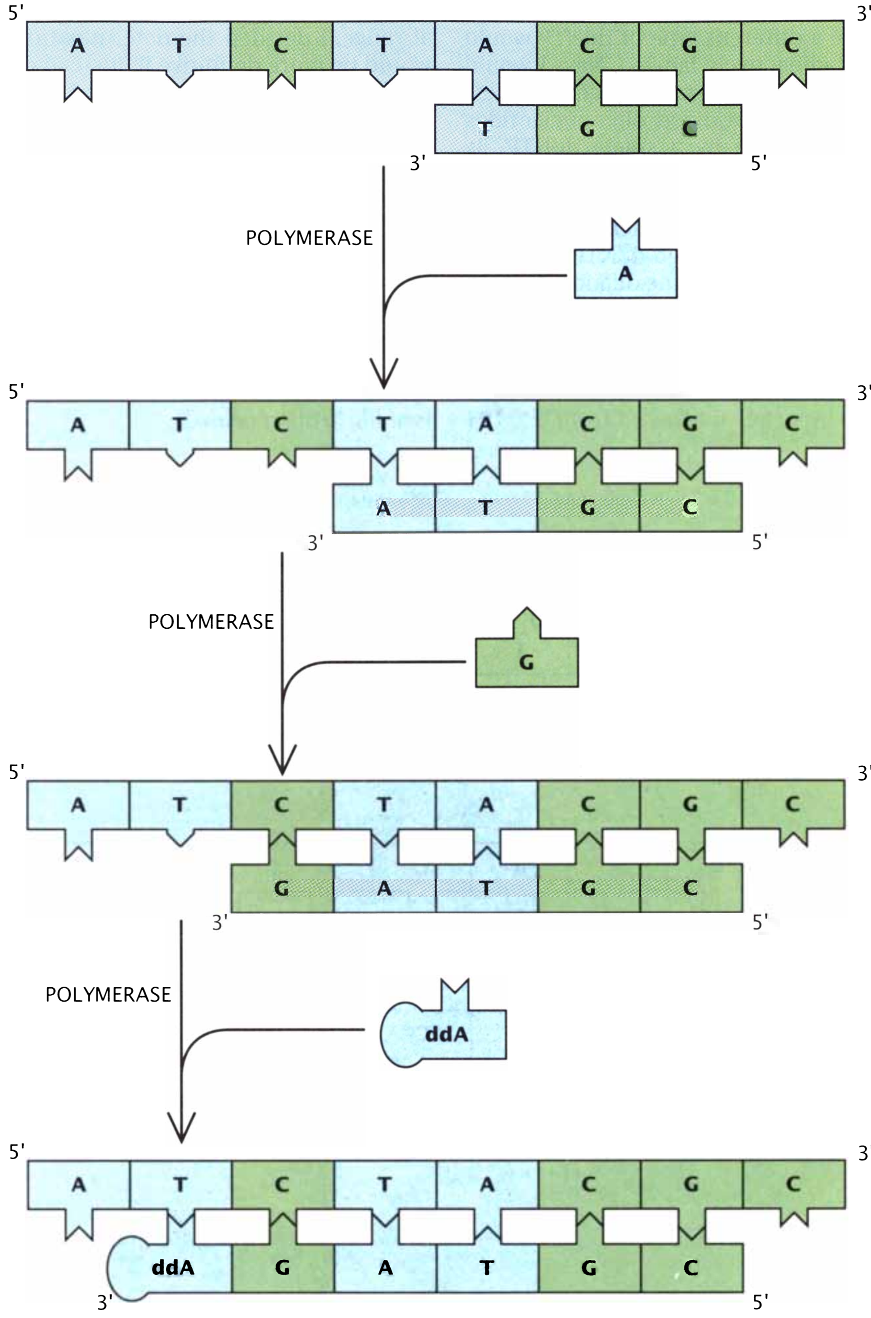
strand adjacent to the site where the primer binds.

What I did not realize at the time was that there were many good reasons why my sequencing idea could not work. The problem was that oligonucleotides sometimes hybridize with DNA sequences other than those intended; these unavoidable pairings would have made my results ambiguous. Even in the hands of those skilled in the art of careful hybridization, it was impossible to bind oligonucleotides to whole human DNA with sufficient specificity to get anything even approaching a meaningful result.

It was because of this limitation that researchers had resorted to more difficult procedures for looking at human DNA. For instance, restriction enzymes could be employed to cleave the DNA sample into various fragments that could be separated from each other by electrophoresis; in this way, the sample could be "purified," to some extent, of all DNA except the target fragment before the hybridization of oligonucleotide probes. This approach reduced erroneous hybridizations sufficiently to provide meaningful data, but just barely. Moreover, this procedure was lengthy and would not work on degraded or denatured samples of DNA.

Another technique that was much too lengthy for routine DNA analysis involved cloning. A human DNA sequence of interest could be cloned, or copied, into a small ring of DNA called a plasmid. Copies of this plasmid and the targeted sequence could then be produced in bacteria, and sequence information could be obtained by oligonucleotide hybridization and dideoxy sequencing. In the early 1980's dideoxy sequencing of cloned DNA was the method by which most human DNA sequence information had been obtained.

In proposing my simple-minded experiment, I was implicitly assuming that no such cloning or other step would be necessary to detect specific human DNA sequences by a single oligonucleotide hybridization. In token defense of my misguided puttering, I can point out that a group down the hall led by Henry A. Erlich, one of Cetus's senior scientists, was trying another method based on the hybridization of a single oligonucleotide to a human DNA target. No one laughed out loud at Henry, and we were all being paid regularly. In fact, we were being paid enough to lead some of us to assume, perhaps brashly, that we



DNA POLYMERASE, an enzyme, can lengthen a short strand of DNA, called an oligonucleotide primer, if the strand is bound to a longer "template" strand of DNA. The polymerase does this by adding the appropriate complementary nucleotide to the three-prime end of the bound primer. If a dideoxynucleotide triphosphate (ddNTP) such as dideoxyadenine (ddA) is added, however, no further extension is possible, because the three-prime end of the ddA will not link to other nucleotides.

were somewhere near the cutting edge of DNA technology.

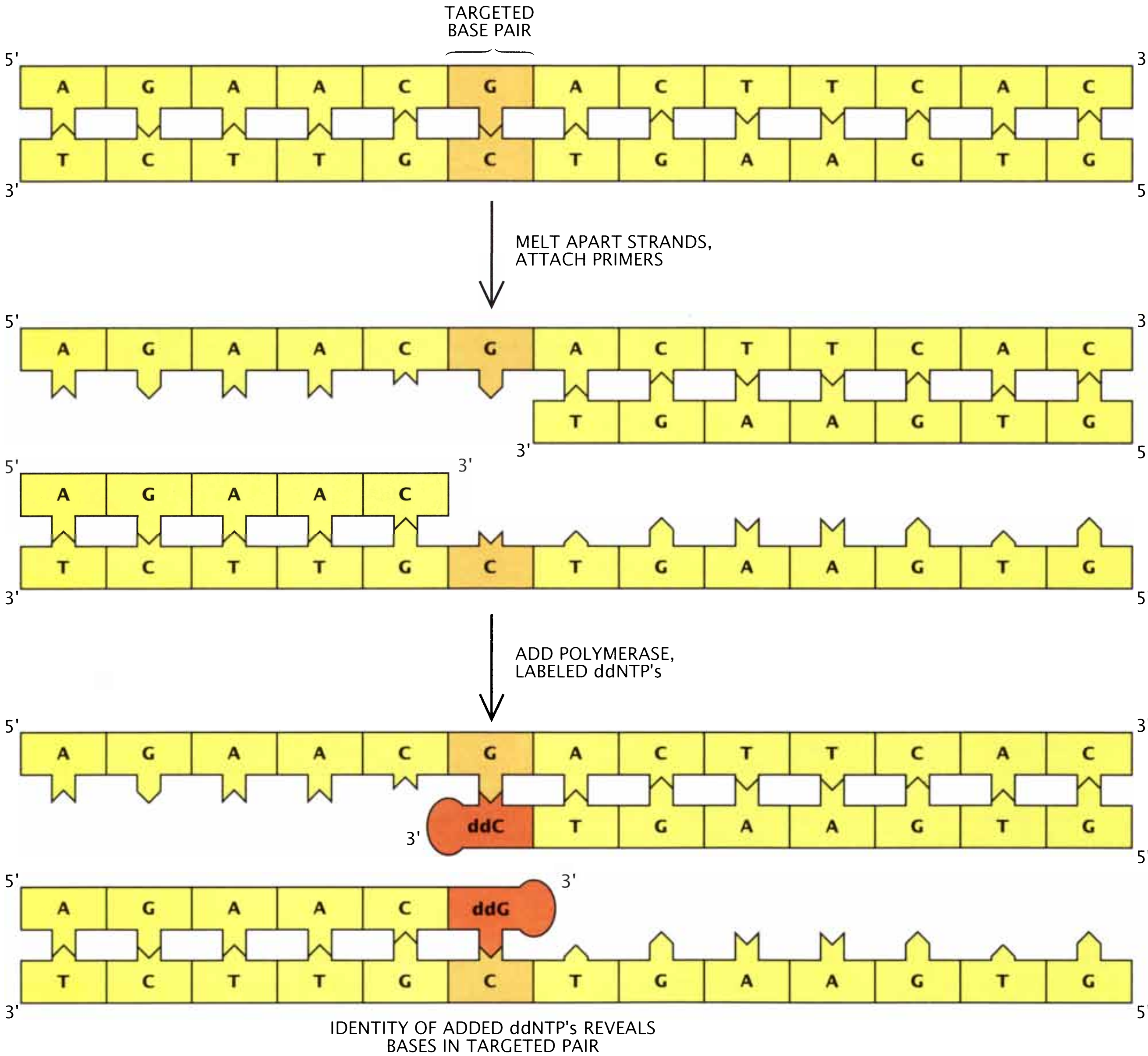
One Friday evening late in the spring I was driving to Mendocino County with a chemist friend. She was asleep. U.S. 101 was undemanding. I liked night driving; every weekend I went north to my cabin and sat still for three hours in the car, my hands occupied, my mind free. On that particular night I was thinking about my proposed DNA-sequencing experiment.

My plans were straightforward. First I would separate a DNA target into single strands by heating it. Then I would hybridize an oligonucleotide to a complementary sequence on one of the strands. I would place portions of this DNA mixture into four different tubes. Each tube would contain all four types of ddNTP's, but in each tube a different type of ddNTP would be radioactively labeled. Next I would add DNA polymerase, which would extend the hybridized oligonucleotides in each tube by a single ddNTP. By electrophoresis I could separate the extended oligonucleotides from the residual ddNTP's; by identifying which radioactively labeled ddNTP had been incorporated into the oligonucleotide, I could determine the corresponding complementary base in the target strand. Simple.

Around Cloverdale, where California 128 branches northwest from U.S. 101 and winds upward through the coastal range, I decided the determination would be more definitive if, instead of just one oligonucleotide, I used two. The two primers would bracket the targeted base pair I hoped to identify. By making the oligonucleotides of different sizes, I would be able to distinguish them from each other. By directing one oligonucleotide to each strand of the sample DNA target, I could get complementary sequencing information about both strands. The experiment would thereby contain an internal control at no extra inconvenience [*see illustration below*].

Although I did not realize it at that moment, with the two oligonucleotides poised in my mind, their three-prime ends pointing at each other on opposite strands of the gene target, I was on the edge of discovering the polymerase chain reaction. Yet what I most felt on the edge of was the mountain road.

That night the air was saturated with moisture and the scent of flowering buckeye. The reckless white stalks poked from the roadside into the glare of my headlights. I was thinking about the new ponds that



TO DETERMINE THE IDENTITY of a targeted base pair in a piece of DNA, the author hoped to apply a variation on a technique called dideoxy sequencing. First two primers would be bound to the opposing strands in the DNA at sites flanking the targeted pair. DNA polymerase and dideoxynucleotide triphosphates (ddNTP's) would then be added to the mixture, which would allow each of the primers to be extended by only one base. The identity of the added ddNTP bases would reveal what the complementary targeted bases were. The technique could work with only one primer, but the use of two would provide a control for checking the results. Planning this experiment led the author to the polymerase chain reaction.

I was digging on my property, while also hypothesizing about things that might go wrong with my base-sequencing experiment.

From my postdoctoral days in Wolfgang Sadee's laboratory at the University of California at San Francisco, where John Maybaum was devising clinical assays for nucleotides, I remembered that my DNA samples might contain stray traces of nucleotide triphosphates. It would complicate the interpretation of the gel, I figured, if stray nucleotides introduced with the sample added themselves to the three-prime end of the primers before the planned addition of the labeled ddNTP's.

One thought I had was to destroy any loose nucleotide triphosphates in the sample with alkaline phosphatase, a bacterial enzyme. This enzyme would chew the reactive phosphate groups off any nucleotide triphosphates, thereby rendering them inert to a polymerase reaction. Yet I would then somehow have to eliminate the phosphatase from the sample, or else it would also destroy the ddNTP's when I added them. Normally one can deactivate unwanted enzymes by heating them and altering their essential shape; I believed, however, bacterial alkaline phosphatase could refold itself into its original form. I therefore rejected alkaline phosphatase as an answer to the problem.

I was, in fact, mistaken. Much later I learned that alkaline phosphatase can be irreversibly denatured by heating if no zinc is present in the solution. As it turned out, my mistake was extraordinarily fortunate: had I known better, I would have stopped searching for alternatives.

Every mile or so another potential solution arose but fell short. Then, as I began the descent into Anderson Valley, I hit on an idea that appealed to my sense of aesthetics and economy: I would apply the same enzyme, DNA polymerase, twice—first to eliminate the extraneous nucleotide triphosphates from the sample, then to incorporate the labeled ddNTP's.

I reasoned that if there were enough nucleotides in the sample to interfere with the experiment, there would also be enough for the DNA polymerase to act on. By running the sample through a kind of preliminary mock reaction with oligonucleotide primers and polymerase but without ddNTP's, I could easily deplete any nucleotides in the mixture by incorporating them into the extending oligonucleotides. Then, by raising the temperature of the sample, I could separate the extended oligonucleotides from the DNA targets. True, the extended oligonucleotides would still be in the sample; but because there would be far more unextended primers than extended ones in the mixture, the DNA targets would probably hybridize with unextended primers when the mixture cooled. I could then add ddNTP's and more polymerase to perform my sequencing experiment.

Yet some questions still nagged at me. Would the oligonucleotides extended by the mock reaction interfere with the subsequent reactions? What if they had been extended by many bases, instead of just one or two? What if they had been extended enough to create a sequence that included a binding site for the other primer molecule? Surely that would cause trouble....

No, far from it! I was suddenly jolted by a realization: the strands of DNA in the target and the extended oligonucleotides would have the same base sequences. In effect, the mock reaction would have doubled the number of DNA targets in the sample!

Suddenly, for me, the fragrance of the flowering buckeye dropped off exponentially.

Under other circumstances, I might not have recognized the importance of this duplication so quickly. Indeed, the idea of repeating a procedure over and over again might have seemed unacceptably dreary. I had been spending a lot of time writing computer programs, however, and had become familiar with reiterative loops—procedures in which a mathematical operation is repeatedly applied to the products of earlier iterations. That experience had taught me how powerful reiterative exponential growth processes are. The DNA replication procedure I had imagined would be just such a process.

Excited, I started running powers of two in my head: two, four, eight, 16, 32.... I remembered vaguely that two to the tenth power was about 1,000 and that therefore two to the twentieth was around a million. I stopped the car at a turnout overlooking Anderson Valley. From the glove compartment I pulled a pencil and paper—I needed to check my calculations. Jennifer, my sleepy passenger, objected groggily to the delay and the light, but I exclaimed that I had discovered something fantastic. Nonplussed, she went back to sleep. I confirmed that two to the twentieth power really was over a million and drove on.

About a mile farther down the road I realized something else about the products of the reaction. After a few rounds of extending the primers, dissociating the extension products, rehybridizing new primers and extending them, the length of the exponentially accumulating DNA strands would be fixed because their ends would be sharply defined by the five-prime ends of the oligonucleotide primers. I could replicate larger fragments of the original DNA sample by designing primers that hybridized farther apart on it. The fragments would always be discrete entities of a specified length.

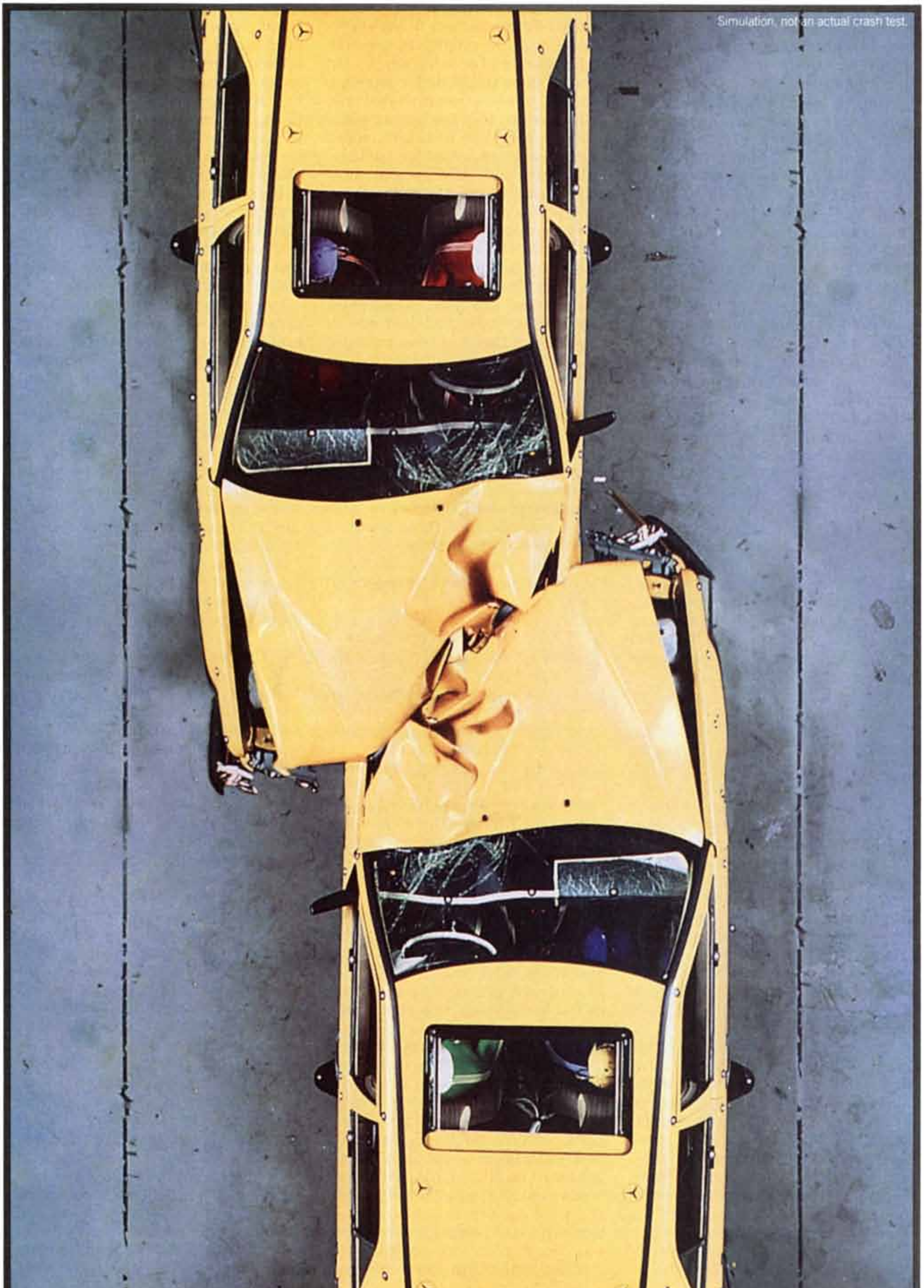
I stopped the car again and started drawing lines of DNA molecules hybridizing and extending, the products of one cycle becoming the templates for the next in a chain reaction.... Jennifer protested again from the edge of sleep. "You're not going to believe this," I crowed. "It's incredible."

She refused to wake up. I proceeded to the cabin without further stops. The deep end of Anderson Valley is where the redwoods start and where the "ne'er-do-wells" have always lived. My discovery made me feel as though I was about to break out of that old valley tradition. It was difficult for me to sleep that night with deoxyribonuclear bombs exploding in my brain.

Yet in the morning I was too tired not to believe that someone, somewhere, must have tried this idea already. Thousands of investigators had, for various reasons, extended single oligonucleotides with polymerases; surely someone would have noticed the possibility of a polymerase chain reaction. But if it had worked, I was sure I would have heard about it: people would have been using it all the time to amplify, or multiply, DNA fragments.

Back at Cetus on Monday I asked one of the librarians, George McGregor, to run a literature search on DNA polymerase. Nothing relevant to amplification turned up. For the next few weeks I described the idea to anyone who would listen. No one had heard of its ever being tried; no one saw any good reason why it would not work; and yet no one was particularly enthusiastic about it. In the past, people had generally thought my ideas about DNA were off the wall, and sometimes after a few days I had agreed with them. But this time I knew I was on to something.

Years ago, before biotechnology—when being a genetic engineer meant that you, your dad and his dad all drove trains—our building at Cetus had been owned by the Shell Development Company. Our laboratory space,

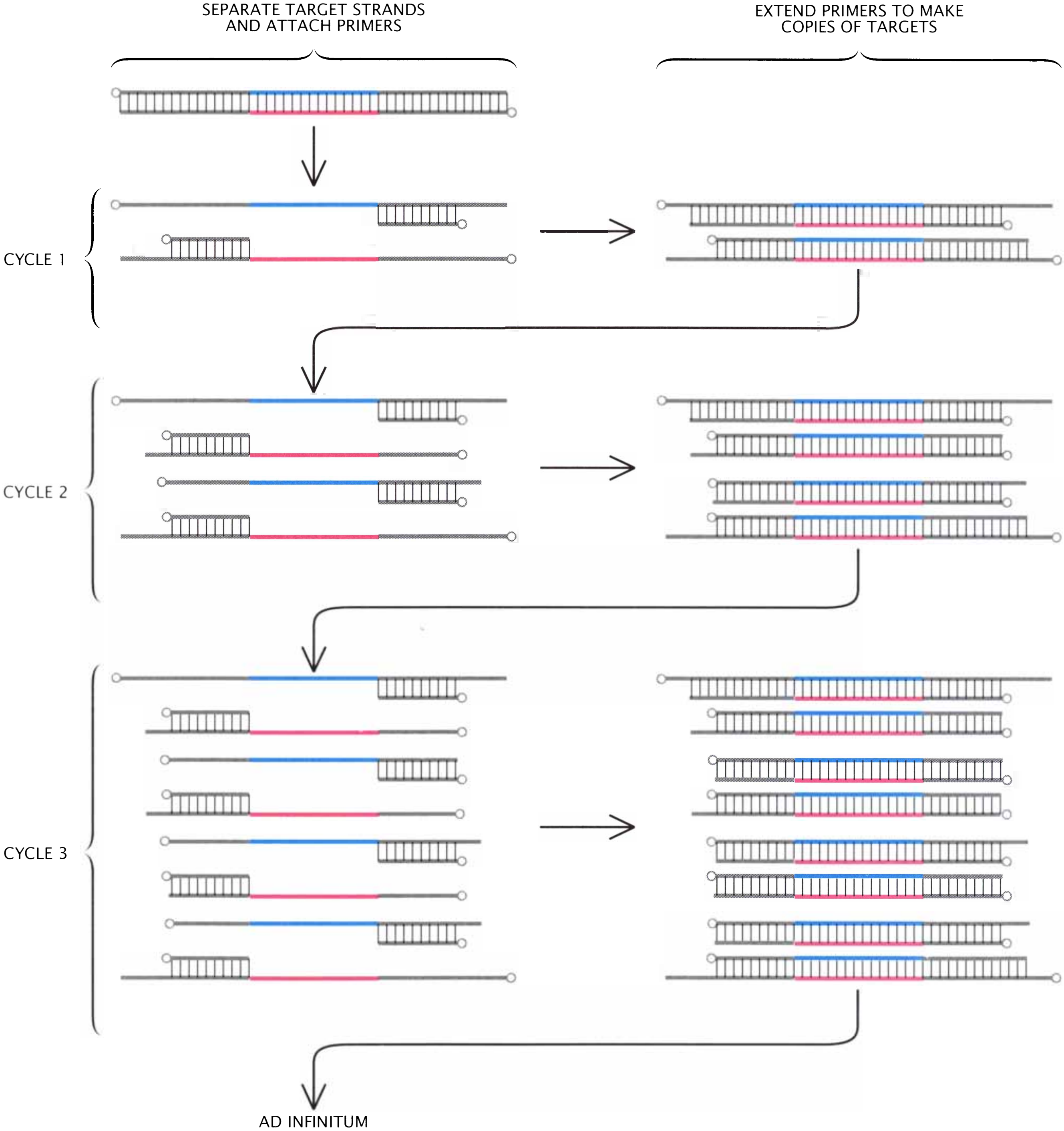

Simulation, not an actual crash test.
© 1990 Mercedes-Benz of N.A., Inc., Montvale, NJ

whose rear windows looked grandly out on the Berkeley hills, had given birth to the "No-Pest Strip." It did not escape my notice that the PCR might someday travel as far as its sibling invention, that distinctively scented piece of yellow plastic.

Months passed as I prepared for my first experiment to verify whether the PCR would work. I had to make many educated guesses about what buffer solutions to use, what the relative and absolute concentrations of the reactants should be, how much to heat and cool the mixtures, how long the mixtures should run and so on. Some of Kornberg's early papers on DNA polymerase helped. To run the experiment, I selected a 25-base-pair target fragment of a plasmid and two oligonucleotide primers that were 11 and 13 bases long, respectively.

When everything was ready, I ran my favorite kind of experiment: one involving a single test tube and producing a yes or no answer. Would the PCR amplify the DNA sequence I had selected? The answer was yes.

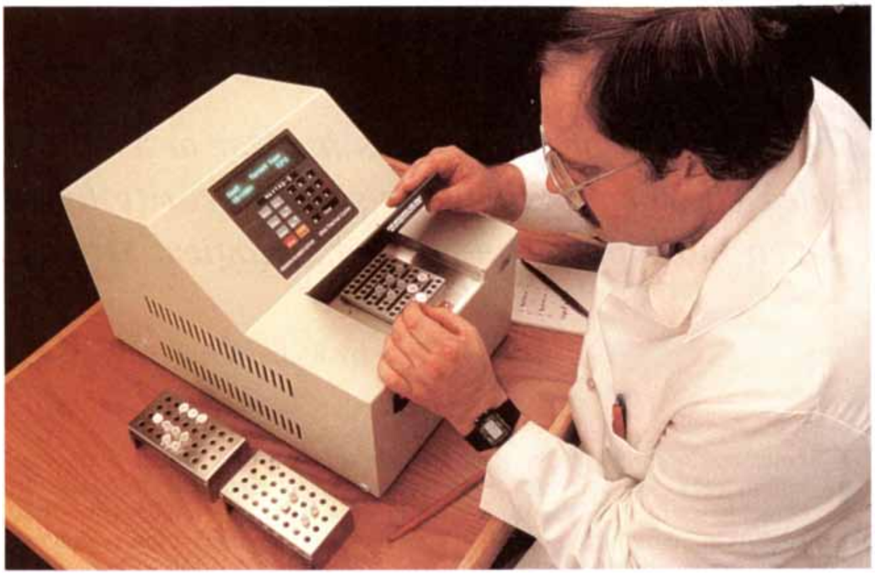
Walking out of the lab fairly late in the evening, I noticed that Albert Halluin, the patent attorney for Cetus, was still in his office. I told him that I had invented something and de-



POLYMERASE CHAIN REACTION is a cyclic process; with each cycle, the number of DNA targets doubles. The strands in each targeted DNA duplex are separated by heating and then cooled to allow primers to bind to them. Next DNA polymerases extend the primers by adding nucleotides to them. In this way, duplicates of the original DNA-strand targets are produced.

scribed the PCR. Al was the first person, out of maybe a hundred to whom I had explained it, who agreed that it was significant. He wanted to see the autoradiogram showing the experimental data right away; it was still wet.

Some people are not impressed by one-tube experiments, but Al was not noticeably skeptical. Patent attorneys, after all, have a vested interest in inventions. He had followed my explanation of the process in his office and agreed that it made sense. Now in the lab he was even a little excited and suggested that I get to work on the experiment and write a patent disclosure. As he left, he congratulated me.



MACHINE that performs the polymerase chain reaction is shown being loaded with samples of DNA. Such devices are rapidly becoming common fixtures in laboratories.

For the next few months I continued to study and refine the PCR with the help of Fred A. Faloona, a young mathematics wizard whom I had met through my daughter. Fred had helped me with the first PCR experiment by cycling the DNA mixture—in fact, that had been his very first biochemistry experiment, and he and I celebrated on the night of its success with a few beers.

In the following months we confirmed that the PCR would work on larger and larger fragments of plasmid DNA. Eventually we obtained some human DNA from Henry Erlich's laboratory and produced evidence for the amplification of a fragment from a single-copy gene.

Today many of the initial hitches or inefficiencies of the PCR have been worked out. Several slightly different protocols are now in use. I usually recommend that the DNA samples be cycled between temperatures of about 98 degrees Celsius, just below boiling, and about 60 degrees C. These cycles can be as short as one or two minutes; during each cycle the number of DNA target molecules doubles. The primers are usually from 20 to 30 bases long. One of the most important improvements in the process is the use of a particular DNA polymerase originally extracted from the bacterium *Thermus aquaticus*, which lives in hot springs. The polymerase we had originally used was easily destroyed by heat; consequently, more had to be added during each cycle of the reaction. The DNA polymerase of *Thermus aquaticus*, however, is stable and active at high temperatures, which means that it only needs to be added at the beginning of the reaction. This high-temperature polymerase is now produced conveniently by genetically engineered bacteria.

The virtually unlimited amplification of DNA by the PCR was too unprecedented to be accepted readily. No one was prepared for a process that provided all the DNA one could want. The reaction seemed self-evident to Fred and to me because it was our toy. For most people, it took some getting used to.

In the spring of 1984, while working on the patent, I presented a poster describing the PCR at the annual Cetus Scientific Meeting. These meetings were always fun, because Cetus had some first-rate scientific advisers, and I was looking forward to talking with them about my invention.

Yet nobody seemed to be interested in my poster, and I felt increasingly anxious. People would glance at it and keep walking. Finally, I noticed Joshua Lederberg, president of the Rockefeller University, nearby, and I snared him into looking at my results. Josh looked the poster over carefully and then turned his enormous head, the Nobel-laureated head, the head that had deduced in 1946 that bacteria could have sexual intercourse. "Does it work?" He seemed amused.

Pleased, I confirmed that it did, and we talked for a long time. At one point he mentioned that about 20 years previously, after Kornberg had discovered DNA polymerase, the two of them had considered the notion that the enzyme could somehow be harnessed to make large quantities of DNA. They had not figured out exactly how to do it, however. I reminded him that oligonucleotides were not readily available at that time and that there was hardly any DNA sequence information either.

But he looked back at my poster with an expression that I have almost come to expect. I think that Josh, after seeing the utter simplicity of the PCR, was perhaps the first person to feel what is now an almost universal first response to it among molecular biologists and other DNA workers: "Why didn't I think of that?" And nobody really knows why; surely I don't. I just ran into it one night.

FURTHER READING

SPECIFIC ENZYMATIC AMPLICATION OF DNA IN VITRO: THE POLYMERASE CHAIN REACTION. Kary Mullis, Fred Faloona, Stephen Scharf, Randall Saiki, Glenn Horn and Henry Erlich in *Cold Spring Harbor Symposia on Quantitative Biology*, Vol. 51, No. 1, pages 263-273; 1986.

SPECIFIC SYNTHESIS OF DNA IN VITRO VIA A POLYMERASE-CATALYZED CHAIN REACTION. Kary B. Mullis and Fred A. Faloona in *Methods in Enzymology*, Vol. 155, Part F, pages 335-350; 1987.

AMPLIFICATION OF HUMAN MINISATELLITES BY THE POLYMERASE CHAIN REACTION: TOWARDS DNA FINGERPRINTING OF SINGLE CELLS. Alec J. Jeffreys, Victoria Wilson, Rita Neumann and John Keyte in *Nucleic Acids Research*, Vol. 16, pages 10953-10971; 1988.

DNA SEQUENCING WITH THERMUS AQUATICUS DNA POLYMERASE AND DIRECT SEQUENCING OF POLYMERASE CHAIN REACTION-AMPLIFIED DNA. M. A. Innis, K. B. Myambo, D. H. Gelfand and Mary Ann D. Brow in *Proceedings of the National Academy of Sciences*, Vol. 85, No. 24, pages 9436-9440; December, 1988.

THE POLYMERASE CHAIN REACTION. T. J. White, Norman Arnheim and H. A. Erlich in *Trends in Genetics*, Vol. 5, No. 6, pages 185-188; June, 1989.

Impact Cratering on the Earth

The explosion of the world's nuclear arsenal could not match the power released when a kilometer-size meteorite hits the earth. Have impacts even more powerful altered biological as well as geologic evolution?

by Richard A. F. Grieve

Thousands of times in the past billion years, an asteroid or a comet has struck the earth, propelling debris at 50 times the speed of sound, vaporizing tons of solid rock and carving a crater many kilometers across. Each event lasted only a few seconds. Yet its effects can reverberate through the course of geologic and biological history.

Although it may seem inconceivable that the earth could conceal such catastrophic events, investigators must search painstakingly for evidence of terrestrial impact cratering. Most of the impact craters on the earth disappeared long ago. The forces that erode mountains, deposit sediments, eject lava and shift continents serve also to obliterate terrestrial craters at a rate faster than that observed on any other rocky planet. Many impact craters that have survived are inaccessible, submerged beneath the ocean. By studying the more than 120 known terrestrial craters, however, workers have discovered techniques to identify other craters, determine the dimensions of the impacting body and chart the stages of crater formation.

This information has been valuable to geologists studying how the impact of very large celestial objects may have affected the evolution of the earth. Did the primordial earth collide with an object the size of Mars and eject enough material to form the moon? Did the remarkably large number of celestial objects that hit the earth up to four billion years ago retard the evolution of life? Has the impact of large asteroids or comets in more recent geologic times affected the environment and biosphere, leading to mass extinctions? The answers are emerging from current studies of terrestrial impact cratering.

RICHARD A. F. GRIEVE is a research scientist with the Geological Survey of Canada. In 1970 he received a Ph.D. from the University of Toronto. He has also earned a master's degree from Brown University and a D.Sc. from Aberdeen University, U.K. While working for the Apollo program, he became interested in terrestrial impact phenomena. He now studies the effects of impacts on crustal formation of the planets.

Before the 1970's most geologists considered impact cratering an insignificant process. Some even claimed that all craters resulted from the explosion of volcanoes. Yet a few workers staunchly defended impact cratering as an important part of the earth's evolution.

Daniel M. Barringer, an American engineer and entrepreneur, first argued in 1905 that a hole in the ground in Arizona was the result of the impact of a large iron meteorite. In making such a claim, Barringer staked more than just his scientific reputation. He believed buried beneath this crater was several million tons of iron meteorite, which he hoped to exploit commercially. He was correct about the origin of the crater, which is now known as Barringer, or Meteor, Crater. Unfortunately for Barringer's commercial ambitions, virtually all remnants of the meteorite had been destroyed under the extreme pressures and temperatures of impact.

Two decades later the geologist Walter H. Bucher studied the rock formations at several craters in the U.S. Although Bucher proposed that the craters had formed in some kind of localized explosion, he assumed incorrectly that volcanoes had provided the explosive power. In 1936, however, John D. Boon and Claude C. Albritton, Jr., reasoned that the features investigated by Bucher were, in fact, impact craters. By 1960 many investigators, including Edward C. T. Chao and Eugene M. Shoemaker, both of the U.S. Geological Survey, and Robert S. Dietz of the Institute of Oceanography, had studied the geology of craters in enough detail to recognize definitive evidence of impact. More specifically, they could identify the effects of shock metamorphism—the irreversible changes produced in rocks under the high pressures and temperatures of a meteorite impact.

The discovery of shock-metamorphic effects inspired Carlyle S. Beals and his colleagues at the Dominion Observatory in Canada and Wolf von Engelhardt of the University of Tübingen in West Germany to search systematically for impact structures. By 1970 they and others had identified more than 50 impact craters on the earth. Despite such findings, some investigators still debated the origin of these craters.

Workers associated with the Apollo missions effectively resolved the debate when they examined lunar samples and determined the rate of cratering on the moon. Lunar impact cratering is closely related to that on the earth because the earth and the moon have been subject to the same flux of asteroidal and cometary bodies throughout the 4.5 billion years of geologic time. In some sense the moon contains a better record of terrestrial impact cratering than the earth does itself. For although the earth's gravitational field captures more asteroids and comets than the moon does, the average crater is preserved 100 times longer on the moon than on the earth. When workers ascertained the rate of lunar cratering, they established beyond doubt that impact cratering has been a common and important process on the moon and, by implication, on the earth.

The study of the lunar craters had its limitations, however. It was difficult to establish a direct link between individual lunar craters and the rocks that were gathered on their surface. The samples had all been transported from their source by one or more impact events. As geologists recognized the limitations of lunar studies and the importance of impact cratering in planetary evolution, they began to appreciate the need to investigate terrestrial impact cratering. Furthermore, the study of terrestrial impact craters provides the essential framework for the results of computational, experimental and planetary studies of impact cratering.

More than 120 terrestrial impact craters have been discovered so far, and several new craters are identified each year. The most exhaustive search for impact craters is currently being conducted by Viktor L. Masaitis of the All-Union Geological Research Institute in Leningrad, Alexander T. Basilevsky of the Vernadsky Institute in Moscow and Vilen I. Feldman of Moscow M. V. Lomonosov State University.

Impact craters on the earth range in age from a few thousand to almost two billion years. Most impact craters are younger than 200 million years, even though evidence from the moon suggests the cratering rate has been roughly constant during the past three billion years. Old craters are less abundant simply because they have been destroyed by erosion, sedimentation and other geologic processes.

In addition to the temporal bias toward younger structures, the destruction of impact craters results in an uneven geographic distribution. About two thirds of the known impact craters are located in regions known as cratons, which are the stable interior cores of continents. Cratons experience low rates of erosion and other destructive processes, and they therefore preserve craters for longer periods. (So far more craters have been

BARRINGER CRATER covers over a square kilometer of the Arizona desert. The crater was formed about 30,000 years ago by the impact of an iron meteorite. Like most simple craters, it consists of a bowl-shaped depression and an elevated rim.

found in the cratons of North America, Europe and Australia than in those of South America and Africa simply because national programs to search for impact craters are better supported on some continents than on others.)

It is also taken for granted that 70 percent of all meteorites land in the ocean. Yet the only underwater impact crater discovered to date lies beneath shallow waters off the coast of Nova Scotia. Named the Montagnais structure, it is 50 million years old and 60 kilometers wide.

Lunar and terrestrial evidence indicates that from one to three craters larger than 20 kilometers are expected to form somewhere on the earth every million years. Smaller craters should populate the earth at an even faster rate. Hence, if one takes into account the age of the earth's surface and the rates of erosion, one can estimate that perhaps only 10 percent of the craters larger than 10 kilometers and less than 100 million years old have been discovered so far.

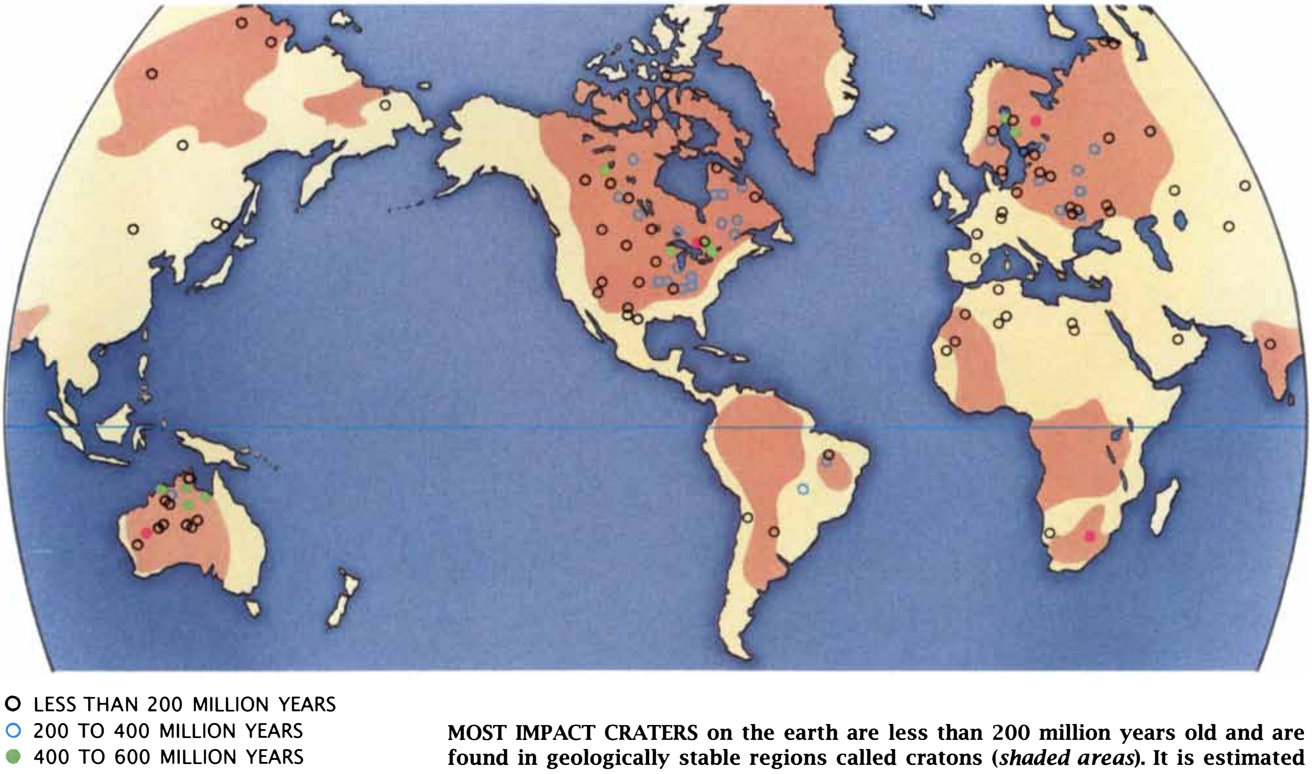
Impact cratering is a unique geologic process in that vast amounts of energy are released in a small area in a very short time. The magnitude of the energy release depends mainly on the speed and size of the impacting body. Asteroids strike the earth at an average speed of about 25 kilometers per second (about twice as fast as the fastest rocket). A body whose mass is more than 1,000 metric tons careens through the atmosphere practically unhindered. A body whose mass is less than 100 tons, however, decelerates through the atmosphere to about 50 percent of its original velocity. Even small meteorites traveling at these velocities transfer considerable kinetic energy to the ground. This energy is converted to pressure and heat. The pressure exerted on the meteorite and the target rocks can exceed 100 gigapascals (one million times atmospheric pressure); temperatures can reach several thousand degrees Celsius.

These extreme conditions vary depending on the impact velocity and the composition of both target surface and meteorite. The smallest and slowest of the meteorites that strike the earth can simply shatter into pieces. Larger meteorites, however, generate enough heat and pressure on impact to melt and even vaporize both the impacting body and some of the surrounding terrestrial rock. Unless a large meteorite breaks apart during its journey through the atmosphere, therefore, fragments of the impacting body are not usually found near large impact craters.

The two basic forms of impact craters are known as simple and complex. The classic example of a simple structure is Barringer Crater. Like other simple craters, it is shaped like a bowl with a raised rim. Barringer Crater is 1,200 meters across and 170 meters deep. Below its apparent floor is a lens of brecciated, or broken and mixed, rock. This "breccia lens" contains some rocks that have been transformed by the high pressures and temperatures of impact—the process of shock metamorphism. The bottom of the lens is some 380 meters below the top of the crater, or about twice the distance from the apparent floor to the top. The rocks surrounding the breccia lens are also broken and exhibit shock-metamorphic effects, but they have not moved much or changed orientation. These rocks make up the true walls of the crater.

The projectile that formed Barringer Crater was an iron meteorite whose diameter was estimated to be 60 meters and whose mass was approximately one million metric tons. The meteorite hit the earth at a speed of about 15 kilometers per second and released an estimated 10^{17} joules of kinetic energy—an amount equivalent to energy released from the explosion of the most powerful nuclear devices or 20 million tons of TNT.

The transfer of this much kinetic energy from the meteorite to the target rocks produces the crater. Most



MOST IMPACT CRATERS on the earth are less than 200 million years old and are found in geologically stable regions called cratons (*shaded areas*). It is estimated that 1,000 more craters may remain to be discovered on land and the ocean floor.

of the energy propagates as a hemispherical shock wave that travels through the rock [*see illustration at right*]. The shock wave compresses the target rocks and propels them downward and outward from the point of impact. The shock wave can accelerate the target rocks to speeds of a few kilometers per second.

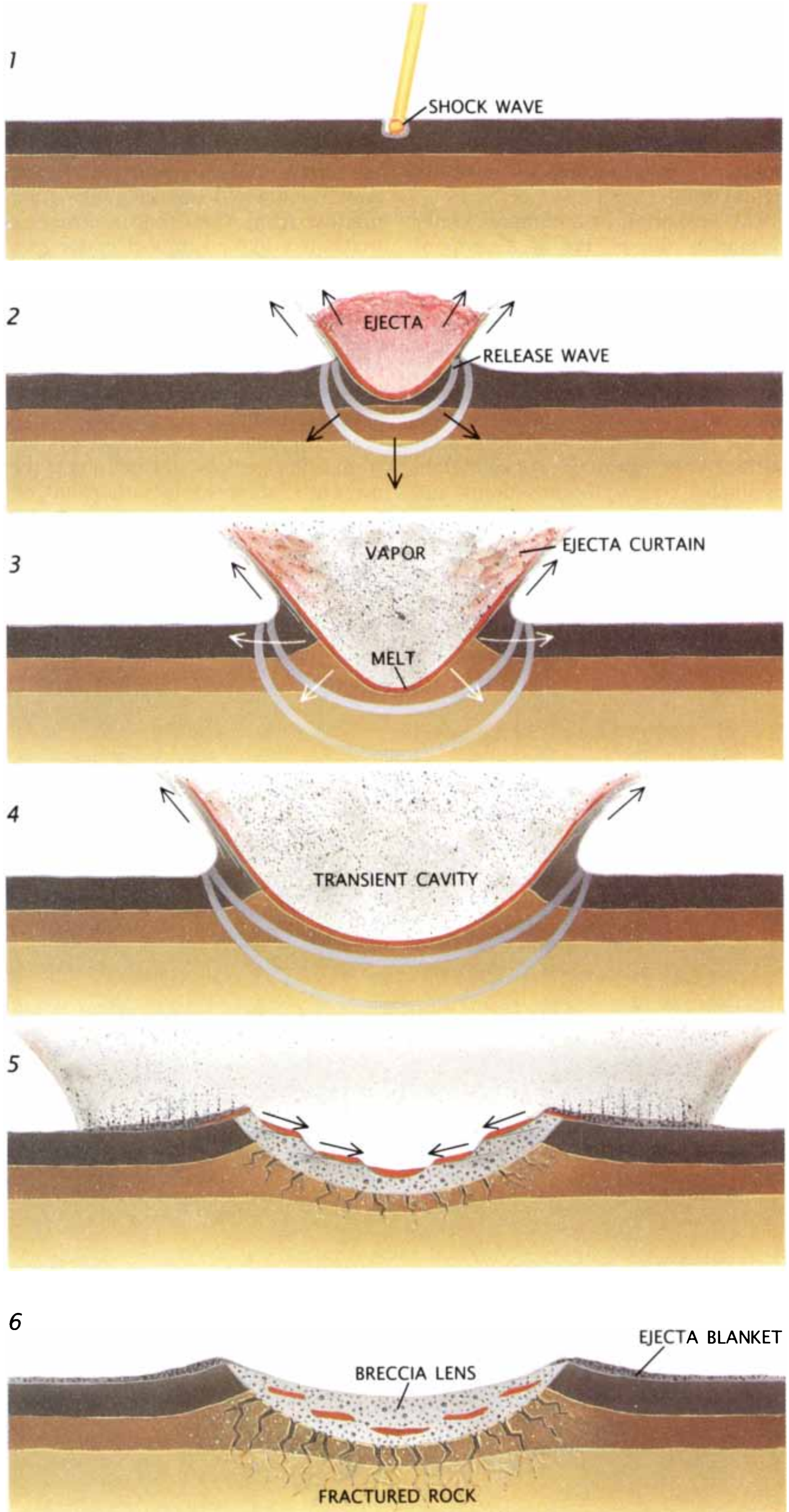
The shock wave precedes the release, or rarefaction, wave, which decompresses the target rocks. The release wave, like the shock wave, pushes the target material downward directly below the point of impact. Beyond this region the release wave creates more complicated dynamics. The release wave overtakes the target materials that are moving downward and outward from the shock wave. As the release wave interacts with the moving material, it deflects some of this material upward and outward.

This process ejects some target material from the center of the developing crater. The ejection of material from the center combined with the downward movement in the center creates a transient cavity lined with fractured target rock. Almost as fast as the cavity is formed, the fractured rock collapses inward. The collapsed walls form the breccia lens that partially fills the final, simple crater.

The diameter attained by a simple crater varies with the nature of the target rock. Geologists know the maximum size is related to rock strength, but they have not identified all the factors that determine the maximum diameter of simple craters. Simple craters can be as large as two kilometers in diameter, if the target rocks are sedimentary, and up to four kilometers across, if the target rocks are crystalline.

Terrestrial impact craters larger than four kilometers across exhibit the features characteristic of complex craters. Such craters can be 100 times wider than they are deep. Complex craters have uplifted central structures surrounded by an annular trough and a fractured rim. Between the central structure and the rim are various materials transformed by the impact: melted and brecciated rocks and shocked materials.

Initially a complex crater forms in much the same way as a simple crater does. As the transient cavity begins to grow, however, some of the rocks in the center rebound upward. The rebound effect lifts the floor of the transient cavity to form a central feature. The uplift in the center of a



METEORITE less than a few hundred meters across forms a simple crater when it collides with the earth (*1*). The impact produces a shock wave and a release wave in the bedrock (*2*). These waves compress, melt, vaporize and excavate the rocks. Thus, the waves create a cavity, which soon reaches its maximum depth (*3*) and diameter (*4*). The cavity walls collapse inward (*5*), leaving a lens of brecciated rock (*6*).

complex crater amounts to about one tenth of the crater's final diameter. For example, the uplift associated with the 100-kilometer-wide Manicouagan Crater in Quebec is estimated to be 10 kilometers, a distance more than a quarter of the thickness of the continental crust.

The formation of a complex crater resembles the effects of a drop of water as it falls into a pool. If one photographs the pool just after the drop hits the surface, one sees the water well upward near the point of impact, while ripples form in the surrounding region, accompanied by a spray of water droplets. But whereas the pool becomes a smooth, level surface once again, the rocks melted by impact freeze at some point, and the dynamic structures are preserved.

It is sometimes quite difficult to identify an impact crater, simple or complex, by its structure alone. Most recognized impact craters have eroded to form vague circular patterns in the terrain. Only the youngest craters have maintained anything like their pristine form. Even then a structure that looks like an impact crater may have been produced by volcanic activity or other geologic processes. Furthermore, because brecciated rocks are found at various sites that are not impact craters, they are not necessarily proof of impact.

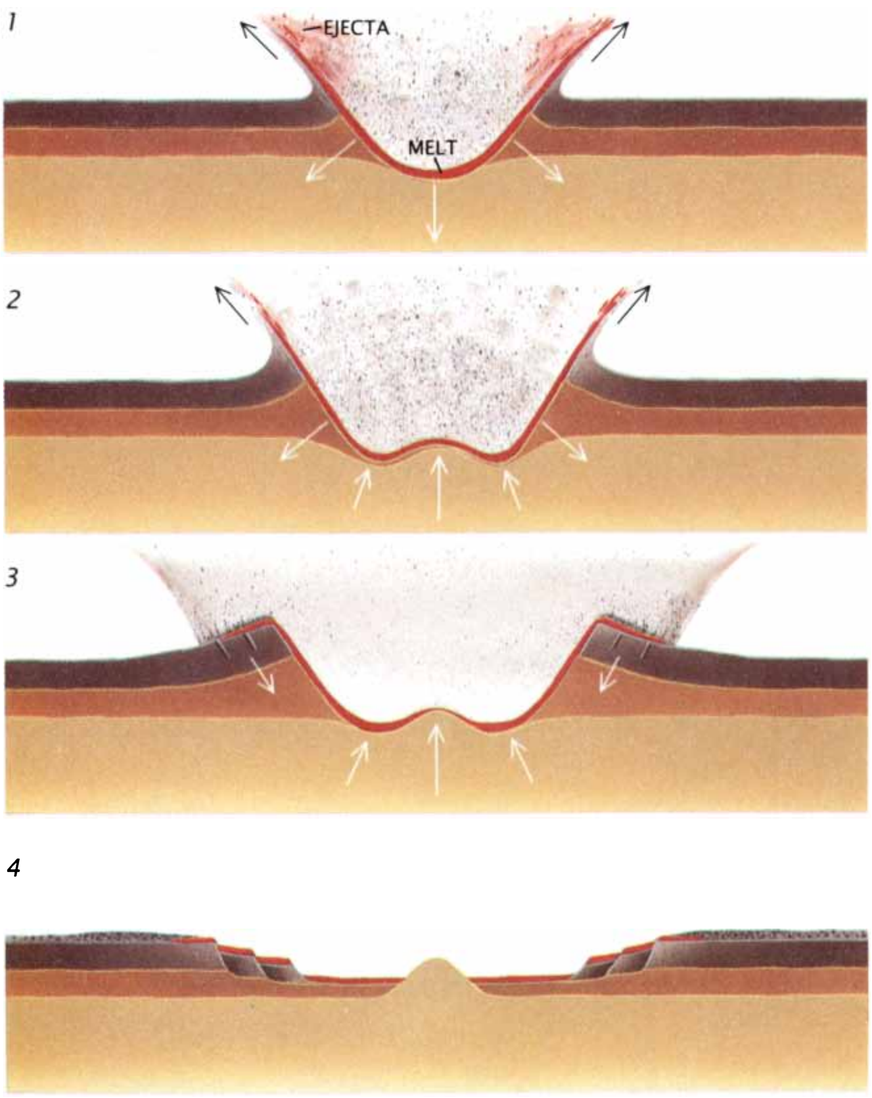
The critical piece of geologic evidence distinguishing impact craters from other geologic formations is the presence of shock-metamorphic effects. In simple craters, unfortunately, much of this evidence is buried within the breccia lens. As a result, many circular features of up to a few kilometers in diameter are suspected but not proved to be of impact origin.

It is considerably easier to confirm an impact origin for complex craters than for simple craters because the uplift in the crater's center exposes the shocked rocks. On the other hand, because complex craters are shallow and have uplifted central features, they are more likely to be mistaken for other types of formations. Historically, the unusual appearance of complex craters caused some workers to doubt that such features were created by the same processes that formed simple craters. Even today one reads scientific papers questioning the impact origin of a specific complex crater. This is particularly true for the largest and oldest structures, such as Sudbury in Canada and Vredefort in South Africa, which have estimated original diameters in the 150- to 200-kilometer range and have been heavily modified during almost two billion years.

Impact cratering is the only geologic process known to produce shock-metamorphic effects. Rocks do change under increased pressure and temperature in the normal geologic setting, such as in the course of orogeny (mountain building). Yet the pressures exerted during orogeny are generally less than one gigapascal and the temperatures less than 1,000 degrees C. Shock-metamorphic effects occur at about the same temperatures but at much greater pressures. The effects appear at about 10 gigapascals and can be produced at hundreds of gigapascals. Shock-metamorphic effects have never been discovered in an explosive volcanic environment. The pressures are simply too low [*see illustration on page 72*].

It must be admitted, however, that temperatures and pressures deep within the earth can equal those needed for shock metamorphism. Yet the conditions within the earth do not yield shock-metamorphic effects because they are applied over the course of millions of years. Shock metamorphism requires the virtual instantaneous application of high temperatures and pressures. For example, a meteorite that produces an impact crater of about three kilometers across will generate shock compression for only a few hundredths of a second.

Many types of shock-metamorphic effects have been observed in lu-



COMPLEX CRATER forms when a kilometer-size meteorite strikes the earth at speeds of about 30 kilometers per second. The impact vaporizes and melts rock, forming a cavity (*1*). Rock beneath the point of impact rebounds upward (*2*), lifting the cavity floor (*3*). The rim of the cavity collapses to create the final crater (*4*).

nar rocks and meteorites. They have also been reproduced in the laboratory and in high-energy nuclear explosions. These experiments help geologists to differentiate the types of shock-metamorphic effects and the pressures required to produce them.

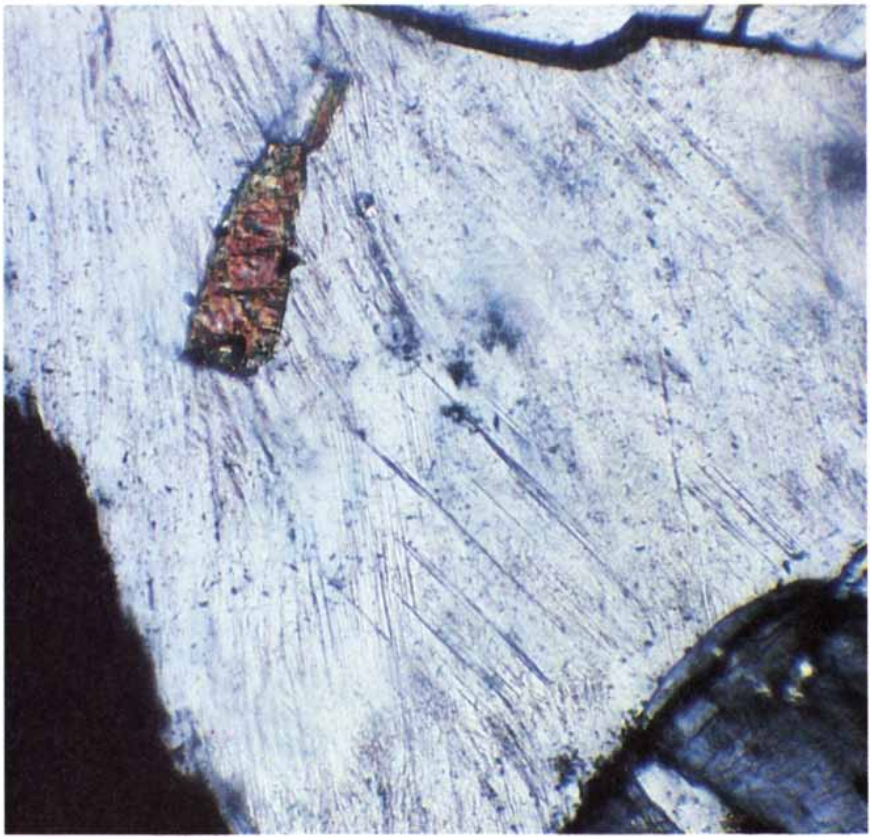
The most apparent shock effect is the formation of shatter cones: rocks fractured in a conical and striated pattern. Shatter cones form most readily in fine-grained rocks, such as limestone and quartzite, that have little internal structure. They develop best at pressures of between five and 10 gigapascals.

With the exception of shatter cones, all known diagnostic indicators of shock metamorphism are seen in individual minerals at microscopic scales. The shock effect that is perhaps best documented is seen in crystals of quartz and feldspar. At pressures of between 7.5 and 25 gigapascals, these minerals fracture in parallel planes that are spaced a few microns apart and that are oriented in directions dependent on the crystalline structure. The number of planes and their orientation are directly related to the shock pressure. One can therefore deduce the pressures that produced a particular sample by comparing the sample with a laboratory standard.

Shock can also produce various high-pressure materials, such as diamond from carbon or stishovite from quartz. Stishovite is an exceptionally good indicator of shock metamorphism. It forms at pressures of approximately 16 gigapascals. Although it can be generated within the earth at depths of from 500 to 600 kilometers, stishovite changes back into quartz before reaching the surface because it is highly unstable. Hence, stishovite found at the surface has been formed at the surface, and the only phenomenon known to generate the required pressures is meteorite impact.

At certain shock pressures, a mineral loses its crystalline structure altogether. Feldspar minerals break down at about 30 gigapascals and quartz at about 40 gigapascals. These pressures transform the minerals into so-called diaplectic glasses (from the Greek *diaplesso,* which means "to destroy by striking"). Diaplectic glasses, unlike conventional types, have the overall shape and composition of the original crystal. Yet diaplectic glasses have no apparent order or structure at the atomic level.

Most rocks shocked to 60 gigapascals or more melt completely. When these rocks cool and crystallize, they



SHOCK-METAMORPHIC EFFECTS result from the high shock pressures that meteorites produce on impact. Shatter cones (*top*) form at pressures of about five gigapascals; at approximately 10 gigapascals, quartz crystals (*bottom*) fracture along parallel planes spaced a few microns apart. These effects are called planar features.

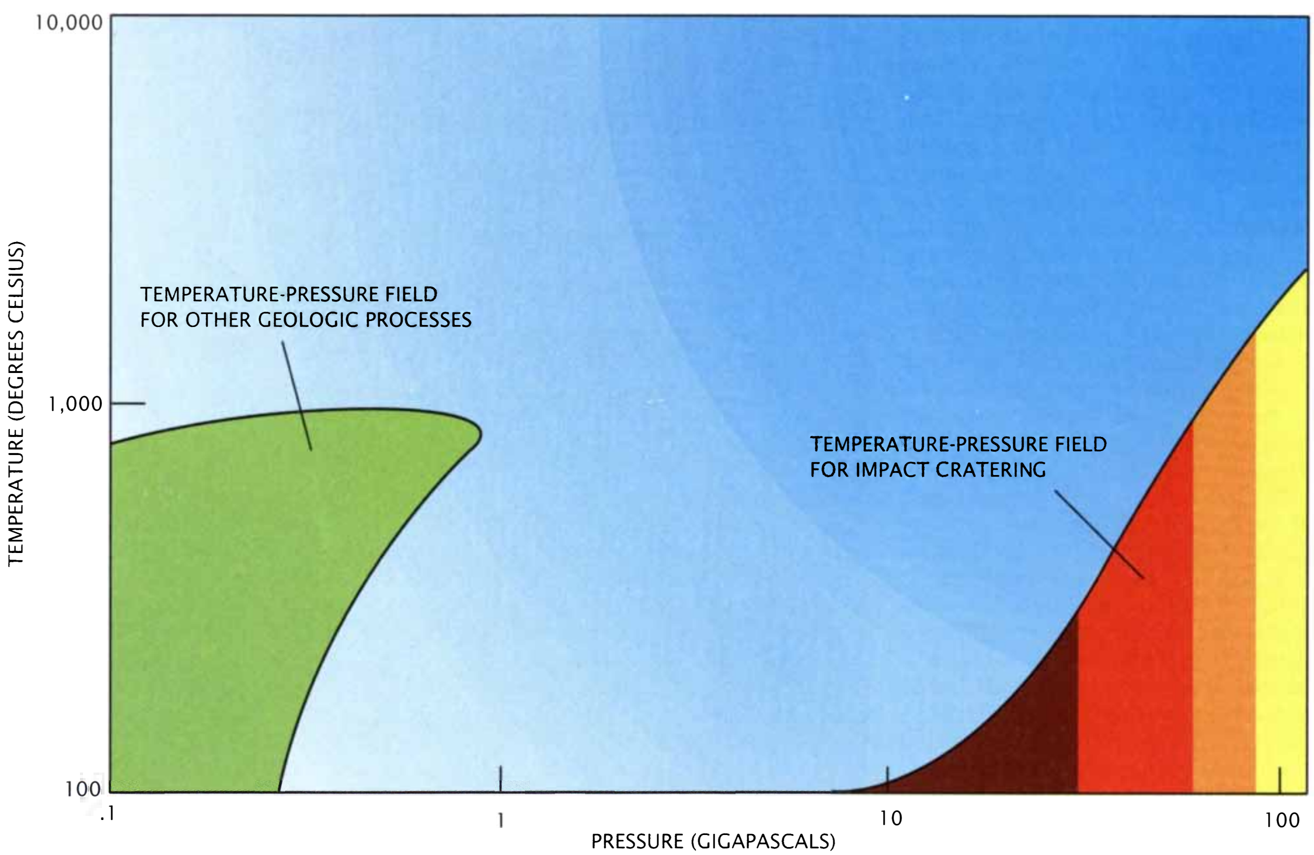
form impact melt rocks and glasses. Impact melt rocks can be found in the breccia lens of simple craters, in annular sheets at complex craters and among the fragments ejected from craters. Impact melt rocks can look like volcanic rocks, and, in fact, melt rocks at the Manicouagan Crater were first identified as volcanic rocks.

Impact melt rocks differ from volcanic rocks in several respects, however. First, impact rocks are not found in the context of common volcanic landforms. Second, they tend to contain large amounts of unmelted fragments from the local bedrock. Third, they have no internal cooling boundaries and appear to have erupted in one huge outpouring of lava. Finally, impact melt rocks can have a chemical composition unlike that of any volcanic rock. In general, an impact rock is a chemical mixture of the various target rocks that have been shocked above their melting point. In contrast, volcanic rocks have compositions determined by the melting of selected minerals under the conditions of pressure and temperature found beneath the earth's crust.

Some impact melt rocks have unusual amounts of certain trace elements that could only have come from a meteorite. For example, the concentration of nickel in an impact melt rock can be 20 times greater than that in the local bedrock. It is improbable that the nickel originated from the bedrock or from farther beneath the earth's surface. Nickel and other elements such as platinum, iridium and cobalt are more concentrated in the iron cores of planets than in their crusts, because they migrate to the core during planet formation. For this reason, these elements are called siderophiles, which means "iron lovers." Siderophilic elements are therefore not abundant in volcanic rocks and other rocks of the earth's crust. This is not the case, however, for old asteroids and comets that formed independently of the planets. Hence, a high concentration of siderophilic elements in a melt rock is a good indicator of meteorite impact.

Investigations of shock metamorphism and impact cratering have led to a number of interesting hypotheses about the earth's evolution. Perhaps the most interesting and controversial case of impact is one in which the shock-metamorphic effects have been detected but the crater has not been found. In 1980 a group of geoscientists led by Luis W. Alvarez of the University of California at Berkeley was studying a clay stratum at sites in Italy, Denmark and New Zealand. The clay stratum had been deposited at the end of the Cretaceous period and at the start of the Tertiary period—a time 65 million years ago when most of all the living species became extinct. The group discovered that the clay layer contained an anomalously high concentration of siderophilic elements. They concluded that the elements came from the impact of a very large meteorite, possibly 10 kilometers in diameter. The energy released from such an event is estimated at 10^{23} joules, equivalent to the explosion of 100 trillion tons of TNT or to about a million eruptions of Mount St. Helens. Alvarez and his co-workers proposed that this stupendous impact disrupted the earth's ecology and caused the mass extinction.

The hypothesis has many rivals. Notably, Charles B. Officer and Charles L. Drake, both of Dartmouth College,

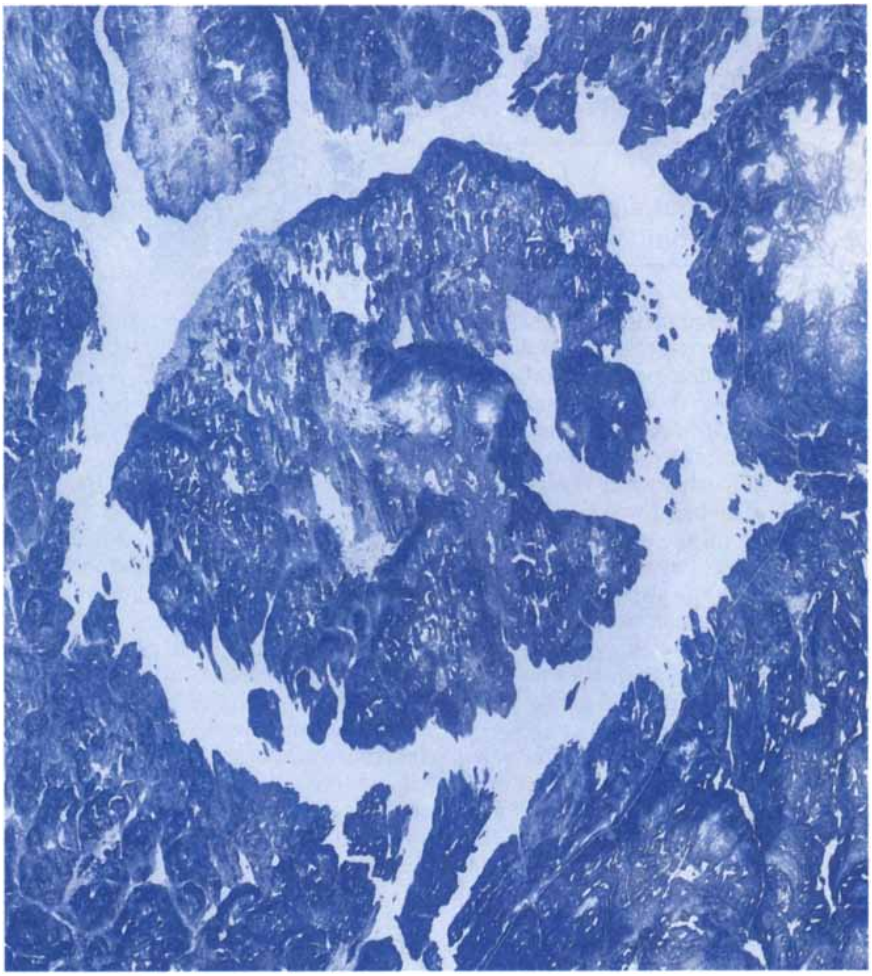


PRESSURES AND TEMPERATURES produced by impact cratering are much greater than those generated by other geologic processes, such as volcanic activity and earthquakes. As shock pressure increases from 10 to 100 gigapascals, mineral crystals first develop planar features (*brown*), then degenerate into glasses (*red*), then melt (*orange*) and finally vaporize (*yellow*).

have put forward other mechanisms, such as volcanism. Since Alvarez's original discovery, however, the siderophilic anomaly at the Cretaceous-Tertiary boundary has been confirmed worldwide. Considerable physical evidence for the impact hypothesis has also been provided by the discovery of shocked minerals, in particular quartz and feldspar with planar features, at Cretaceous-Tertiary boundary sites. On the other hand, an apparent setback to the impact hypothesis is the absence of an impact crater of the expected size (150 to 200 kilometers) and age (65 million years). The crater may yet be discovered on land or in the ocean. Perhaps, as some have proposed, one should not expect to find a large crater, because the Cretaceous-Tertiary event was the result of a shower of asteroids or comets that produced many small craters.

An impact event comparable to that suggested for the Cretaceous-Tertiary event is estimated to occur approximately every 100 million years. For this reason, some investigators have advanced that other mass extinctions were also caused by the impact of meteorites. At present these claims are not supported by physical or chemical evidence.

Considering what may have happened during the Cretaceous-Tertiary event, other investigators have suggested that a high rate of cratering hampered early biological evolution on the earth. Indeed, evidence from the oldest surfaces of the moon indicate that more than four billion years ago the cratering rate in the earth-moon system was perhaps 100 times higher than the cratering rate in more recent geologic time. Hence, it has been postulated that life did not have a chance to develop for an extended period until the cratering flux had subsided to a level approaching that of the present day.



ANNULAR LAKE in Quebec is now part of the 210-million-year-old complex crater named Manicouagan. The lake (here frozen and covered by snow) is 75 kilometers in diameter; according to some estimates, the crater was once 100 kilometers across.

Recent studies of impact cratering have also provided insights into the earliest periods of geology. For example, the rate of impact cratering declined at about the time, four billion years ago, when the oldest-known surface rocks were formed. Workers have speculated that this intense bombardment was sufficient to prevent the preservation of surface rocks older than four billion years.

Another contribution of cratering research to the study of geology has been a hypothesis about the origin of the moon. Computer models have shown that the moon may be the result of an impact between the protoearth and a body the size of Mars. This hypothesis was initially put forward 20 years ago when it was discovered that the chemistry of lunar samples was similar to that of the deep rocks of the earth's mantle. An additional observation was that the moon had an unusually small iron core.

Recent computer models have indicated that the impact of a Mars-size object on the earth would create an expanding cloud of vapor that would orbit the earth. The vapor would be composed of a mixture of elements from the projectile and the earth's mantle and would have little core material. Over a few hundred years the vapor would condense and accrete to form an orbiting satellite, such as the moon. The fact that the origin of the moon can be modeled successfully as the result of a giant impact does not mean the moon actually formed that way. The models do satisfy many of the chemical and physical properties of the earth-moon system, however, and planetary geologists regard the models seriously. Whatever the case, the models suggest the fundamental role that impact may have played in the evolution of the earth and the solar system.

FURTHER READING

SHOCK METAMORPHISM OF NATURAL MATERIALS. Edited by B. M. French and N. M. Short. Mono Book Corporation, 1968.

IMPACT AND EXPLOSION CRATERING: PLANETARY AND TERRESTRIAL IMPLICATIONS. Edited by D. J. Roddy, R. O. Pepin and R. B. Merrill. Pergamon Press, 1978.

METEORITE CRATERS. Kathleen Mark. University of Arizona Press, 1987.

TERRESTRIAL IMPACT STRUCTURES. Richard A. F. Grieve in *Annual Review of Earth and Planetary Sciences*, Vol. 15, pages 245–270; 1987.

IMPACT CRATERING: A GEOLOGIC PROCESS. H. J. Melosh. Oxford University Press, 1988.

SCIENCE AND BUSINESS

Milliken & Co.

Managing the quality of a textile revolution

They call it "Magnolia." Rectangular and squat, the fabric finishing plant stretches over 17 acres of South Carolina peach country. Inside, machines that turn greige, or undyed, cloth into colored fabrics clatter and belch steam as thousands of yards of polyester wind through the operation. The plant was built in 1963 to capitalize on the old golden rule of textile making: cut costs by processing as much fabric as possible in every batch. Now Magnolia has an additional mandate. It must also produce small amounts of very high quality material as part of a "just in time" manufacturing program.

Magnolia, and its parent company, Milliken & Co., are at the cutting edge of a revolution in the textile business. Led by chairman Roger Milliken, this privately held manufacturer is spearheading a strategy that blends technology with a drive for quality so as to shrink the time required to turn raw materials into finished garments. In this way, Milliken hopes to help stop the dissolution of a mature industry: soft goods.

It is a pitched battle. In 1989 soft-goods imports won more than 60 percent of the U.S. market. But Milliken has already proved itself on other ground. Last year the textile company beat the likes of Hewlett-Packard and Texas Instruments to become one of the two firms awarded the U.S. Commerce Department's award for quality, the nation's top prize for manufacturing prowess. (Xerox was also recognized, for the excellence of its business machines.)

Just-in-time clothes, turmoil at NTT, dapper dinosaurs, soul of a new budget

The battle plan launched by Milliken and the soft-goods industry is called quick response. It is a way to knit all levels of the business—from fiber producers, to textile manufacturers, to apparel makers and finally to retail stores, in great part by exploiting electronic networks. If the plan works, those who buy clothes should get better products and U.S. soft-goods producers should preserve their market, the program's proponents say.

DESIGNING CARPET SQUARE or other fabrics with a computer system is often a collaborative effort between customers and the textile manufacturer, Milliken & Co. Customers select such characteristics as pattern and color; overnight, Milliken makes a sample of the product.

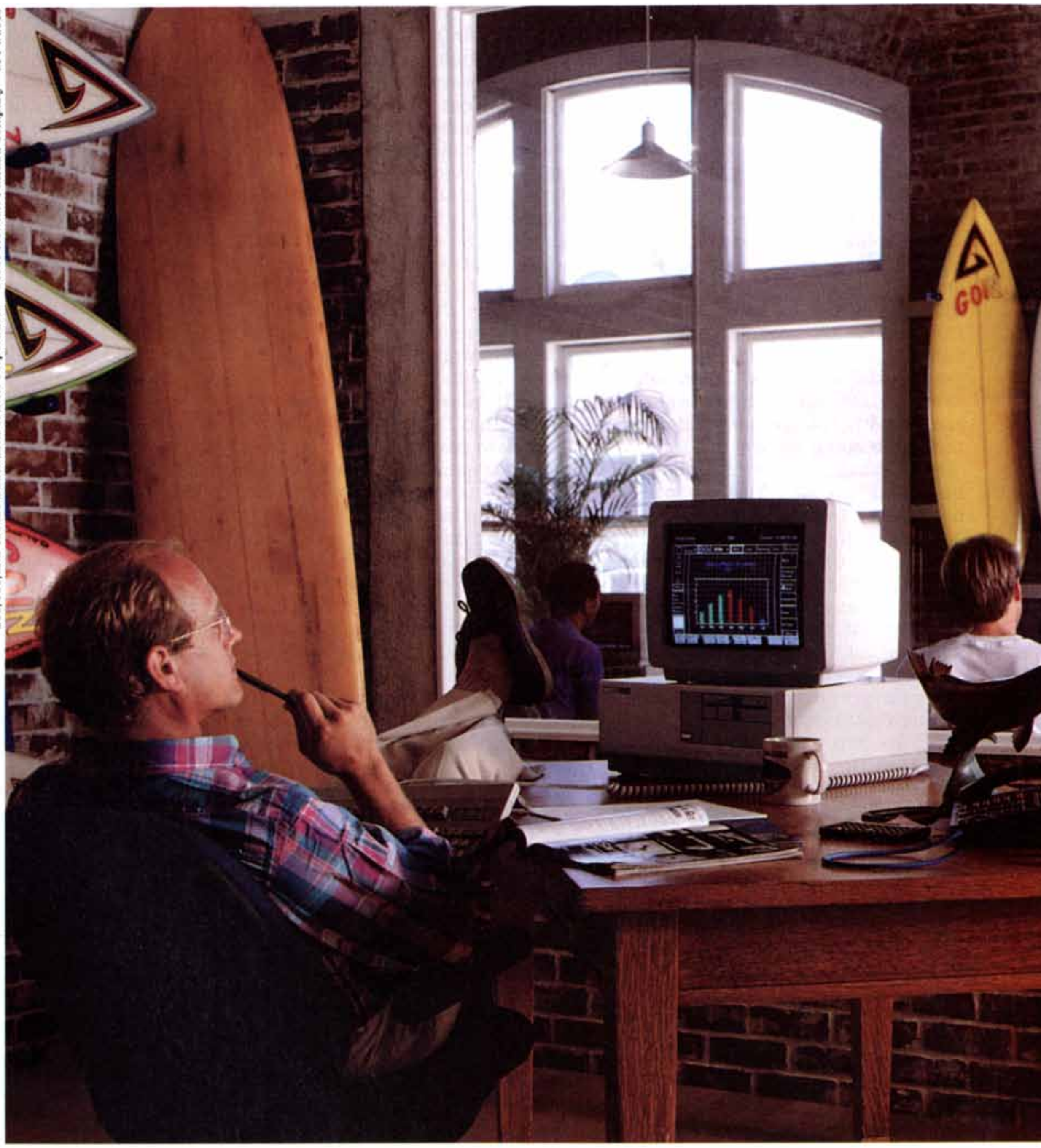
Quick response was born when U.S. fabric and garment producers realized they could not best their foreign competitors strictly on price, says Bob Frazier, a director at Kurt Salmon Associates, a consulting firm. (Wages for many foreign fabric workers are still extremely low.) U.S. producers turned to their obvious, home-court advantage: the short distance between domestic producers and retail stores.

Typically, fabric and apparel producers have needed more than a year to fill retailers' orders for clothes. Retailers have had to forecast what their customers might be likely to buy and then order the garments accordingly. The result, explains Andrew J. Vecchione, who directs Milliken's quick-response program, has been unsold clothes at the end of the season that retailers must discount.

Quick response, Vecchione explains, opens a channel between the garment producers and the shelves in a store. In such a program, textile and garment makers agree to ship small batches of garments to retailers throughout the season. Retailers, who monitor which garments are selling, need only reorder those in demand. Fabric and apparel manufacturers need only make goods that are selling. Consequently, customers get what they want, retailers avoid the nightmare of unsold stock and so save money, and the U.S. soft-goods industry is preserved.

Building such a system is not easy. It relies heavily on an electronic network by which producers and retailers can swap data. Textile makers, accustomed to producing large batches of material, must also retool plants to be able to quickly meet the demand for smaller quantities. Magnolia, for instance, was refitted with new equipment so that it could run 500-yard batches of fabric in addition to its traditional 15,000-yard runs.

Before sending the cloth on to the apparel maker, the textile producer tags the fabric with standardized bar codes that describe the type of material, the color and so on. The apparel maker, who scans these codes as the material arrives, then immediately puts the cloth on the proper production line. Finished garments, in turn, are labeled with specific bar codes—



Robert was sold on Hewlett-Packard PCs while he was still in college.

He was studying engineering and contends an HP calculator was the secret to his success. Since then, HP LaserJet printers have been a big help to his growing company. So when he found he could get Hewlett-Packard reliability in a network of personal computers, Robert decided to stay with a sure thing.

With Intel 286, 386, and 486™-based models, HP's line of Vectra personal computers suits a range of business needs. Expansion is easily accommodated with up to 8 accessory slots, 1.3 gigabytes of hard disk storage, and 64 Mbytes of RAM.

Beyond this, HP's adherence to industry standards ensures compatibility. Now, and into the future.

But the most important feature is Hewlett-Packard quality. This, along with HP's network of authorized dealers, gives you all the assurance you need. For the dealer nearest you, call **1-800-752-0900, Ext. 142F**

There is a better way.

HEWLETT PACKARD

say, one identifying a woman's size-8 red shirt—so that the retailer can track precisely what goods are being sold. (These bar codes are scanned, in turn, by sales clerks when goods are sold.) The retail store then transmits data on sales back to the apparel maker, who electronically orders additional fabric from the textile mill.

A far more subtle underpinning of quick response is the need for high-quality goods at every step in the process. "Quality manufacturing and quick response are synonymous," says Gary Swank, a senior vice president at Haggar Apparel in Dallas. "Our buffer stocks [of fabric] have gone down," he adds, "so when we get the goods from the textile mill they have to be top."

Milliken's quality drive did not begin because the company's products were under fire. It began in Roger Milliken's head. When he read quality guru Philip Crosby's *Quality is Free* over the Christmas 1980 holiday, he began a personal crusade to reinvigorate the company. At the time, Milliken was "managed like the best managed companies in America," says Tom Malone, company president. "We had strong, hierarchical management. We compared the top to the bottom and focused on the people at the bottom who weren't doing so well."

That hierarchy was soon shaken, from top to bottom. Within a few years some 700 mid-level managers were reassigned into more production-level slots. "Quality is free" became Milliken's 1981 campaign.

By 1985 Milliken and Malone began thinking intently about the company's work force: What would get every one of them personally involved with Milliken's goal of achieving higher quality? "Everyone has got to have fans in the stands," Malone insists. "Can you imagine waking up in the morning—a production associate or a secretary—and saying to yourself, 'My goal is for Milliken to make more money than anybody's ever made so we can get written up in the newspapers?'" he demands. Instead, Malone emphasizes, managers must excite employees about work by giving them continual feedback and applause.

Milliken and Malone instituted an extensive program of corporate recognitions and awards. Among them are quarterly "sharing rallies" in which associates (no longer "employees") present their contributions to their peers—and to Malone and Milliken.

Another part of Milliken's push for quality, albeit one that company leaders are reluctant to discuss, has involved advanced technology. Milliken has "very definitely" been at the forefront of the industry in applying new technology, says Charles G. Tewksbury, president of the Institute of Textile Technology. Magnolia's operations, for instance, were converted to computer automation by the mid-1970's. Most of the plant's equipment has been upgraded two or three times since the 1960's. More recently Milliken has adopted some artificial-intelligence tools to help technicians solve problems on the factory floor. And computer programs that recognize verbal commands have been installed to speed quality checks.

The efforts seem to have paid off. Five years ago 75 percent of Milliken's deliveries to customers were on time; now 98 percent are on time. The number of defects in goods has been cut by more than half since 1981.

Have the programs also boosted revenues for the $2-billion company? "It's probably been a tough few years for Milliken," one observer muses. Because the company is privately held, it guards its costs and profits closely. But it makes no excuses for its investments. "We've just decided that this is the way we're going to run our company," Vecchione says. "There is no backup plan." —*Elizabeth Corcoran*

Of Mice and Men

How form affects function in monoclonal-antibody drugs

Monoclonal antibodies, touted as biological guided missiles since their discovery in 1975, are nearing the mark as drugs. Much like the body's own antibodies, which attack foreign invaders, monoclonal antibodies made in large quantities in the laboratory can bind to designated substances. Once bound, those substances are rendered impotent and may be destroyed by the immune system. The ability to zero in on a target has already made monoclonal antibodies extremely useful in diagnostics. Now they are being used to treat acute illnesses caused by circulating toxins and cells.

Over the next 18 months the U.S. Food and Drug Administration is likely to approve three monoclonal antibodies for therapy: one to treat bone-marrow rejection and two competing products to treat septic shock, a lethal bacterial infection. The companies making these drugs use different manufacturing processes, which confer distinct advantages as well as potential problems.

The only monoclonal drug to make it to market so far is Ortho Pharmaceutical's Orthoclone OKT3, which treats kidney rejection; it was approved in 1986. OKT3 works by blocking circulating *T* cells from attacking a graft organ because it is foreign. OKT3 does not typically bother immature *T* cells. These grow up likely to perceive the new kidney as "self" and to leave it alone.

Foiling *T* cells also underlies the next monoclonal product likely to win FDA approval. The drug, developed by Xoma in Berkeley, Calif., treats graft-versus-host disease. In this illness, transplanted bone marrow rejects its new host by sending *T* cells to attack the patient's tissue. Xoma's product, which is called XomaZyme-CD5 Plus, marks these rampaging *T* cells for destruction. Xoma has recently started testing the drug against rheumatoid arthritis, insulin-dependent diabetes and inflammatory bowel disease.

In line behind these antirejection drugs are two varieties of monoclonal antibody for treating septic shock, a bacterial illness that kills some 85,000 hospital patients a year in the U.S. Conventional antibiotics kill the bacteria, but it is the endotoxins, or poisons, the organisms release that can make the infection fatal. Monoclonal antibodies bind to the endotoxin and so signal the body to dispose of it.

Two companies—Xoma and Centocor of Malvern, Pa.—are vying for a market that could be worth hundreds of millions of dollars. The drugs, likely to cost between $1,500 and $2,000 per treatment, are both made in spleen cells. The difference in the companies' approaches lies in the source of these cells and the way they are handled. It is a matter of mice and men.

Centocor selects antibody-producing spleen cells from humans who have died of septic shock. Xoma injects mouse spleen cells with bacterial endotoxin so that the cells will produce antibodies to it. These cells are then "immortalized" by being fused with a continuously growing cell—a cancer cell in Xoma's case, a virus in Centocor's. This fusion creates a new hybrid cell called a hybridoma, which secretes monoclonal antibodies.

The process of making a hybridoma occurs just once; the companies' next step is to produce the drugs in large quantities. Xoma injects its monoclonal antibodies into live mice, where they proliferate and collect in fluid in the mouse's abdominal cavity. They are then extracted and purified. Centocor grows its hybridomas in tanks to churn out the antibodies.

Because mouse proteins differ from human proteins, patients receiving a drug made in mice may make their own antibodies to it. "If that happens, the next time you give any mouse monoclonal, it can't do its job. It gets cleared by the body almost immediately," explains Curtis L. Scribner, chief of the FDA unit that reviews clinical trials of monoclonal antibodies. Xoma says half of its patients have had a HAMA (human-antimouse-antibody) response. "The FDA will weigh the risks and benefits," Scribner says, "but I can't see HAMA delaying approval of anyone's drug."

Centocor thinks a human antibody will work better in a human. The company brushes off a common criticism of human monoclonal antibodies: that they are expensive as well as difficult to produce and store, especially in large quantities. The firm says its proprietary technology has solved these problems. But there is the risk that the final product could pass on to a patient oncogenes or viruses from the immortalization process.

A third approach to monoclonal antibodies combines the strengths of both mouse and human systems. "The pharmaceutical industry is clearly heading for 'humanized' antibodies," according to Ivor Royston, director of clinical immunology at the University of California at San Diego. One technique uses protein engineering to remodel mouse antibodies so that they look more human. Another fuses human and mouse components. Eventually the tinkering should build a better antibody. —*Deborah Erickson*

ALTERNATIVE CLEANING AGENTS for printed-circuit boards, intended to replace chlorofluorocarbons, are being developed at a Du Pont laboratory near Wilmington, Del.

Atmosphere of Uncertainty

Businesses dependent on CFC's cooperate, innovate and wait

This June in London more than 40 nations will reassess their plan to phase out chlorofluorocarbons, the ozone-depleting chemicals employed as refrigerants and used to make insulants and solvents. The meeting's probable outcome is a complete ban on CFC's in many countries by the year 2000.

Far less predictable is how the proposed amendments to the Clean Air Act will affect the alternatives to CFC's that manufacturers are now developing. Some of these compounds deplete ozone, although far less avidly than do their predecessors. The uncertainty is irking CFC makers and worrying users—as well as forcing former competitors and customers to cooperate.

At particular risk are hydrochlorofluorocarbons (HCFC's), which deplete ozone only 2 to 10 percent as much as CFC's do. An amended Clean Air Act could ban them by 2015. Like CFC's, HCFC's keep appliances such as refrigerators cool for long periods because of the chemical stability that chlorine provides. But that stability also enables the compound to catalyze ozone-destroying reactions in the stratosphere. Manufacturers have begun making small quantities of HCFC's for testing, but they warn that if Congress enacts a ban to cap the chemicals' commercial life span, they will not produce them in large quantities.

Even the Environmental Protection Agency thinks Congress should not indulge in such unilateral policy-making. "The U.S. should renegotiate international protocol, so all countries will be signatories to the law," says Stephen R. Seidel of the EPA.

"How can government legislate the demise of HCFC's now, before alternatives to these alternatives are in sight?" asks F. Anthony Vogelsberg, Jr., Du Pont's environmental manager of Freon products. Profit is not manufacturers' sole concern, adds Paul Dugard of ICI. "We're also thinking about all the users who will have to redesign their equipment to accommodate these alternatives."

While the fate of HCFC's hangs in the air, redesign centers on ozone-friendly alternatives called hydrofluorocarbons (HFC's). Du Pont plans to supply commercial quantities of an HFC called 134a by late 1990. Also this year, ICI and Allied-Signal will break ground on factories to make 134a. It should be available by 1991 or 1992. Because HFC's are less efficient refrigerants than chlorine-based chemicals, they cannot simply be "dropped into" existing systems. General Motors has worked since 1986 testing new compressors for its automobile air conditioners. It expects to be able to phase in 134a for the 1994 model year.

Replacing the CFC-based solvents that degrease parts and clean the electronic sensors that automobile systems use is more difficult, says Gerald F. Stofflet of GM. With the Clean Air Act amendments proposing tougher standards, "we've got to be darn sure that whatever we end up using can do the job." In the eyes of the home-appliance industry, federal laws are creating a regulatory "double whammy" by demanding improved energy efficiency at the same time that known chemical technology is being phased out.

Other strictures complicate the move to CFC alternatives. Toxicological testing is far more stringent than when CFC's were introduced in the 1930's. The tests, including evaluation of neurotoxicity and effects on reproduction, have firms behaving in an unusual way: they are cooperating.

Through the Program for Alternative Fluorocarbon Toxicity Testing (PAFT) formed in January, 1988, 14 international companies have pooled funds to test the five leading ozone-friendly candidates. "Almost anyone can figure out what the possible alternatives are, so there are no secrets," explains George M. Rusch, PAFT's toxicology chairman. Preliminary results look good, he says, but the final data will not be compiled until 1993 or 1994.

"We can't base our business on assumptions," complains Leo Soorus of Whirlpool. But companies are assuming that the EPA will eventually regulate "venting," or the release of CFC's into the atmosphere. Consequently, some firms, including Whirlpool and GM, will have CFC-recovery programs up and running this year.

CFC recycling is not just a way of cleaning one's own house. "As CFC's become scarce and expensive, but the chemicals are still needed to run $135 billion of existing equipment, recycling is turning into a business opportunity," says R. Colin Dayton of National Refrigerants in Plymouth Meeting, Pa. His company started recycling CFC's last summer, when the EPA tacitly encouraged industry by ruling that used refrigerant was no longer a hazardous waste. "Now I pay companies for their CFC's," he says. "But I tell them, 'In five years you'll be paying me.' They laugh a little."

The search for benign chemicals is not just a CFC issue, suggests Braden R. Allenby, an environmental attorney for AT&T. The challenge is how to design products in a world where the environment is the limiting factor. He reasons, "We need to develop holistic, sustainable technologies." —*D.E.*

Breakup?

The view from inside a beseiged NTT

Early this spring Japan's government may well announce a plan to break up the country's largest corporation, Nippon Telegraph and Telephone (NTT). The proposal raises the ironic prospect that government action might cripple a company's technical prowess. U.S. observers have a rare opportunity to watch another version of their own history: the breakup of AT&T. Recently Scientific American *talked with top NTT executives to learn their views.*

"There is a contrary wind against large enterprises," observes Kohji Ohboshi, NTT's senior vice president for corporate planning. NTT, which employs more than 270,000 people and serves some 50 million customers, cannot be blamed if it takes the climate personally. "NTT is too big—it has to go the way of the Japan National Railways," says the chairman of a leading Japanese bank. (JNR was dismantled by the government in 1987.)

NTT leaders, though, are reluctant to split up the company. At the heart of their concerns is the fate of NTT's 11 research laboratories. "If the breakup takes place, it's quite elementary that [R&D] capability will plummet," Ohboshi says.

Perhaps. The Japanese government has its eye on another ball: the presumed beneficial effects of competition. "There's the potential for dramatic drops in long-distance costs," says Yang-Soon Lee of Booz Allen & Hamilton in New York City. "And there's a distinct possibility that NTT could become more nimble" after a breakup, he adds. The government started exploring this venue in 1985 when it transformed NTT from a government-run monopoly into a private company owned by more than 1.5 million shareholders and exposed both NTT and Japan's international telecommunications company, Kokusai Denshin Denwa, to competition.

So far privatization has wrought some changes. There are now more than 850 firms in Japan's telecommunications industry, most of which offer value-added networks for distributing data, as well as electronic and voice mail. Yet NTT still holds more than about 97 percent of the communications market. Prices have declined somewhat: early this year NTT lowered its charge for a three-minute long-distance call from 330 to 280 yen ($2.27 down to $1.93), while still asking only 10 yen ($.07) for local calls—a bargain price in high-cost Japan.

NTT executives argue that new systems technology will push long-distance prices lower. "Take ISDN [the integrated-services digital network] fiber optics," says Keiji Tachikawa, an associate vice president for long-range planning at NTT. "There will be exponential declines in costs and in tariffs [of domestic phone calls] over the next decade or two," he predicts.

To accelerate this trend, NTT has been boosting its research budget. Last year NTT's research spending rose to $1.7 billion, or 4.3 percent of revenues. (In contrast, AT&T spent $2.65 billion, or 5.2 percent, on research in 1989.) Since 1985 NTT has also established about 150 subsidiaries and joint ventures to market technology developed at the labs.

Still, the Ministry of Posts and Telecommunications remains critical of NTT's business performance and eager to reconfigure the firm. One plan would divide NTT into 11 regional companies that provide both long-distance and local service; another would follow the AT&T model by splitting up the company into 11 local services firms and one long-distance company. In this case the labs might remain with the long-distance group.

But NTT executives do not want to take that path. They maintain that such a strategy would doom NTT's efforts to install a national ISDN by 1999 and to improve the company's research efforts. "The integration of the telecommunications network [in the U.S.] became a major problem following the splitting up of AT&T," said Haruo Yamaguchi, NTT's president in a speech. Adds Ohboshi: "Users want one-stop accountability. That's not possible if there's a breakup and one customer has to negotiate with a dozen companies!"

According to Tachikawa, NTT will not be able to afford a large research and development budget if the company is split up, because the revenues of a long-distance company will be only about one fifth that of NTT's current revenues.

"NTT [would] like to be a leading power in R&D," Tachikawa says. "In some areas we have caught up with Western countries, so now we have to create new technology, to contribute to progress in the world." NTT has achieved some good results, Tachikawa says, pointing to work in high-density memory chips, methods for making optical fibers and optical-electronics research. Nevertheless, "we don't have enough new things," he adds. "[We are] not so proud of the results so far." The company is keen to expand its basic research efforts in such areas as neural computing.

Perhaps even more critical is the effect a breakup might have on applied development, according to NTT leaders. "Break up NTT, divorce the labs from the regions by attaching them to a long-distance company, deprive the researchers of daily contact with users and what happens to incentives [for innovation]?" asks Tachikawa. "It's very important [for the company's 3,000 researchers] to have contact with the users, to provide a service," Ohboshi adds. "If you cut off the regions from R&D, how would that actually work in practice?" he asks.

Yet it is conceivable that Japan's telecommunications giant will be dismantled, even NTT executives concede. "Only," Ohboshi asks, "give us a few years' grace—time to cut the NTT payroll by another 40,000, time to see what privatization has meant—yes, and time to find out what our labs can do." —*Henry Scott Stokes, Tokyo*

Cashcowosaurus

Robotic dinosaurs put meat on museums' fiscal bones

Kids love dinosaurs. So do museum directors, especially when the creatures move, growl and attract large, paying crowds. Financial pressures, as well as curiosity, have created a market for a growing herd of robotic dinosaurs. Even paleontologists like them, observes Frank H. Talbot, a museum director for the Smithsonian Institution in Washington, D.C.

Robotic dinosaurs were introduced to U.S. museums in 1982 by a San Juan Capistrano, Calif., company called Dinamation. The firm's founder first spot-

ROBOTIC TRICERATOPS, *one of a herd of mechanical prehistoric beasts, is helping to boost museum revenues and to spark public interest in paleontology. Photograph courtesy of Kokoro.*

ted Japanese-made dinosaurs at an amusement-park convention and convinced the manufacturer, Kokoro, to let him market the beasts in the U.S. Dinamation began making its own dinosaurs in 1986; Kokoro opened an office in Los Angeles in late 1987.

Both companies offer their newfangled prehistoric creatures for sale or rent. Renting broods of robotic dinosaurs is popular with museums "at least as much for the revenue they bring as for educational purposes," says Robert M. West, a paleontologist and director of the Cranbrook Institute of Science in suburban Detroit, Mich. "It's up to the institution to make them more than a sideshow," he adds. Cranbrook, for instance, has simulated an excavation site and hosts both auditorium lectures and computer games featuring dinosaurs.

The Pink Palace Museum in Memphis, Tenn., views traveling robotic-dinosaur shows as loss leaders that tempt people into more hands-on educational programs. "People stand out in the cold to see the creatures," says Brad Evans, a spokesperson. A robotic *Triceratops* on display near the entrance turns its head, lumbers forward a step and bellows—all for a quarter. In five years it has almost earned its purchase price of $45,000.

Although they live with the reality of show biz, both Dinamation and Kokoro retain leading paleontologists to keep them up to date on discoveries. They incorporate these findings into their new models and even retrofit those dinosaurs already in the field. "Oh, sure, we changed the plates on *Stegosaurus*," recalls Dinamation's Anthony Bland, when paleontologists concluded that the reptiles' dorsal ridges were staggered, not side by side. Three years ago the mechanical dinosaur tails stopped dragging, after researchers found that tiny bones, apparently evolved from a vertebra, must have supported the appendages. And of course, *Brontosaurus* is no more. The long-necked vegetarian, which now has a broader head, is really *Apatosaurus*.

Although better skeletal specimens give rise to new, improved robots, the manufacturers are not hindered by what paleontologists do not know. "Fossil skin is brown," points out Jun Shimizu of Kokoro, "but nobody knows what the colors really were." The purple patches on his *Pachycephalosaurus* ("butting head") are not entirely frivolous, however. Researchers now think this family of dinosaurs was related to birds and may have been brightly colored.

Except for two species, no one knows what dinosaurs sounded like either. *Parasaurolophus* might have made a noise like that of a bassoon, because its big horn crest would resonate. *Apatosaurus*, which lacked vocal cords, might have sounded like a similarly bereft reptile, the snake. But such hisses might scare children, Shimizu thinks. His *Apatosaurus* robots make a low, throaty grunt. "After all, this is a show too, and people have to enjoy what they're experiencing," he says.

Tens of millions of people see robotic dinosaurs each year around the world, the companies estimate. Dinamation's creatures have even been escaping from science halls, to appear in shopping malls and convention centers. But ubiquity could undercut their appeal, some members of the museum community suggest. Dinamation, however, likens the busy itinerary of its 300 dinosaurs—which travel in groups of five to 18—to Disney rereleasing *Bambi*. "There's always a five-year-old who hasn't seen it," says Wendy Keiper, of Dinamation.

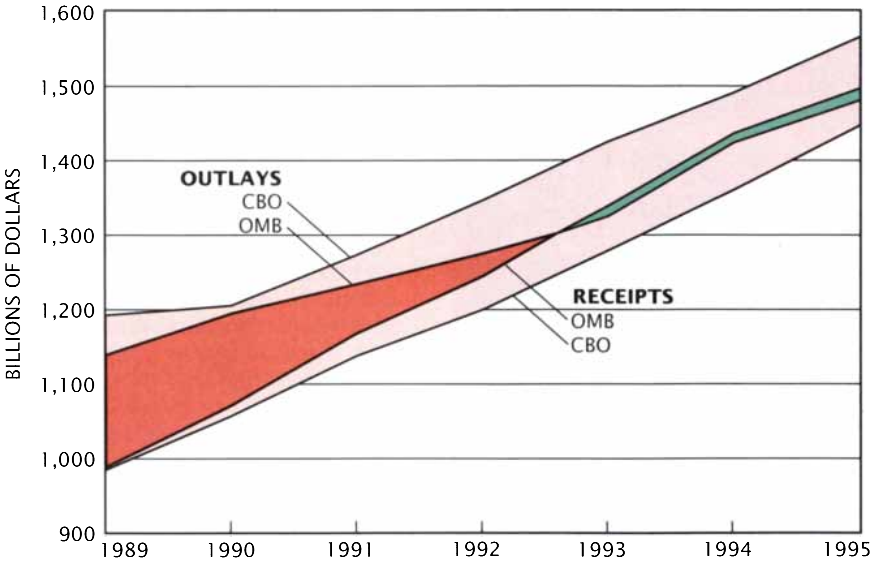
It is not just the plentitude of robotic dinosaurs that might add to ennui, West suggests. "It's the dino-everything. Dinosaur pencil sharpeners, mugs, T-shirts, stickers, bedroom slippers. How long can this go on?" For a while yet, Evans predicts. "Dinosaurs lived for 140 million years, humans have for perhaps 100,000. Instead of considering dinosaurs too dumb and slow to survive, we're asking, Why did they last so long?" —*D.E.*

THE ANALYTICAL ECONOMIST

Budgeting smoke and mirrors

Every January, journalists from mastheads around the country swarm to Washington, D.C., for the great midwinter rite: the unveiling of the federal budget. Perhaps some agency has statistics on how many forests of pulp trees are converted into the documents that government minions distribute at the jamboree of press briefings. Veteran reporters try to outdo one another by asking ever more pointed questions about ever more obscure budget line items. The flood of data (bearing up the occasional twig of information) fills the nation's publications.

What does one really learn from an administration's draft of the budget? The numbers may change as Congress wrangles over the budget until September. (The fiscal year begins on October 1.) Might a proposed budget hint



***PROPOSED FEDERAL BUDGET** will lead to different amounts of revenues and expenditures depending on the growth of the U.S. economy. The Office of Management and Budget, which predicts relatively robust growth, expects that the government will run a deficit until 1993* **(dark red)** *and a surplus in later years* **(green).** *The Congressional Budget Office, which has lower growth expectations, forecasts a continued budget deficit through the mid-1990's* **(light red).**

at the future direction of the economy? Is it important at all?

"Strongly yes and strongly no," says George B. Shaw, director of planning at the Planning Research Corporation, a government contractor in McLean, Va. Congress actually leaves about 95 percent of an administration's budget unscathed, points out Stanley E. Collender, director of federal-budget policy at Price Waterhouse in Washington, D.C. Roughly 60 percent of the spending is dictated by law and covers such expenses as Medicare and interest on the debt. Much of the rest goes for "nuts-and-bolts obligations from prior years," Collender adds.

What the Congress and the administration fight over, then, is the 5 percent or so—less than $60 billion of the $1.2 trillion budget for fiscal 1991—that represents changes in program funding. That may seem a puny sum, but the dollars are important because they support such discretionary programs as scientific research and development. A tiny cut in these funds may amount to a wholesale dismantling of some activity.

Such discretionary spending is particularly contentious because no one knows just how much money will be available. Federal expenditures are constrained by receipts that come mostly from personal and corporate taxes. Those revenues rise or fall with the health of the economy. Projecting receipts is strictly a matter of best-guess estimates, Collender says.

The administration is betting, for instance, that the economy will grow at a brisk 2.6 percent this year and 3.3 percent in 1991 and that the Federal Reserve will keep interest rates relatively low on such instruments as three-month Treasury bills (6.7 percent this year and 5.4 percent in 1991, down from 8.1 percent in 1989). Inflation and unemployment are predicted to remain unchanged.

"It's not as wildly out of line as many of the Reagan economic assumptions," one congressional staff member concedes. "Still, the administration is extremely optimistic about the future." The Congressional Budget Office paints a less vibrant picture: 1.7 percent growth this year, rising to 2.4 percent in 1991, and higher interest rates.

The bottom-line differences, based strictly on economic growth, amount to more than $20 billion. In addition, "technical differences" in how the administration and the CBO tally up the dollars create a further $11-billion discrepancy.

That $31-billion difference of opinion causes some anxiety in Congress. "Say we adopt the president's budget whole in the spring," the staff member says. If the economy grows more slowly than predicted during the summer, thus depressing the government's revenues, Congress will have to scramble to find another $12 billion in funding cuts by autumn, he says. Yet even though Congress complains about the administration's optimistic outlook, it may happily adopt those economic projections as a way of postponing cuts, notes John H. Makin, who directs fiscal-policy studies at the American Enterprise Institute.

Congress has been waving a red flag of its own, too; the so-called peace dividend. The birth of a new order in Eastern Europe raises the possibility of significant cuts in the U.S. defense budget—but not in fiscal 1991. Secretary Richard V. Cheney reportedly started the speculation about such a dividend when he suggested some $180 billion would be cut from defense over the next few years.

"Cheney was his own worst enemy," Collender says. First the secretary estimated what the defense budget would be if it grew faster than inflation, as proposed by the Reagan administration. Then he calculated what he was likely to get. The difference amounted to $180 billion—but not a dividend.

In the short term, even scaling back defense costs money. The Navy reckons it will pay far more than $100 million to decommission two aging nuclear cruisers. Closing bases is a slow process at best. Canceling large weapon-production contracts incurs stiff penalties.

Eventually, though, holding defense spending constant—at roughly $290 billion—over the next five years could generate savings of $150 billion compared with last year's projections, Makin says. The savings would most likely appear in two to three years' time, he adds, and simply help the government meet its goal of reducing the deficit to zero by 1993.

If the budget process is, in fact, an exercise in smoke and mirrors, why do people still watch it? With all its flaws, Shaw says, the budget still highlights an administration's priorities. "Once a year we get a sense of where the government says it's going," he observes. "Then we see if it puts its money where its mouth is." He and other contractors point to the administration's relatively strong boost for the National Aeronautics and Space Administration. On the other hand, the environmental president has called for pushing the spending of the Environmental Protection Agency up to $5.8 billion in 1991, followed by a slow decline to $5.2 billion by 1995.

"There's no strong vision of the future," Collender contends, particularly in contrast to the ideologies that marked the budgets under Ronald Reagan or even Jimmy Carter; "I call it a budget without a soul."

—Elizabeth Corcoran and Paul Wallich

In Bangkok, good things come in enormous packages.

There was a time in Thailand when pachyderms meant political power, and a white elephant was worth its weight in gold. While times have changed—elephants today are used to symbolize good luck instead of wealth—there are still a number of big surprises in store for the businessman in old Siam.

Mai pen rai.

Pronounced "my pen rye," this is translated as "it doesn't matter," and it conveys the Thai attitude toward business and life as well. It means it's best to avoid open conflicts and negotiate patiently —shouting will get you nowhere.

Shy Thais.

Thais seldom invite foreigners home—entertaining is done in restaurants. Use a private dining room at the Oriental Hotel (tel. 236-0400) for special occasions.

The how of the wai.

Instead of a handshake, try the Thai form of greeting, the *wai* (pronounced "why"). Bring your palms together, fingers up, and bow—it's most respectful.

Northwest notes.

Northwest flies more people to Asia from the U.S. than any other airline. And, in addition to our enhanced First and Executive Class service—featuring our award winning cuisine—we offer something that no other U.S. airline can: the knowledge and insight that come from more than 40 years of helping people do business in Asia. For international reservations, just call your travel agent or Northwest at 1-800-447-4747. To find out more about doing business in Asia, call extension 176, at 1-800-553-2215.

NORTHWEST AIRLINES

Advanced Light-Water Reactors

New designs that include "passive" safety features can make nuclear power more attractive, but only if attention is paid to economics, effective management and rigorous training methods

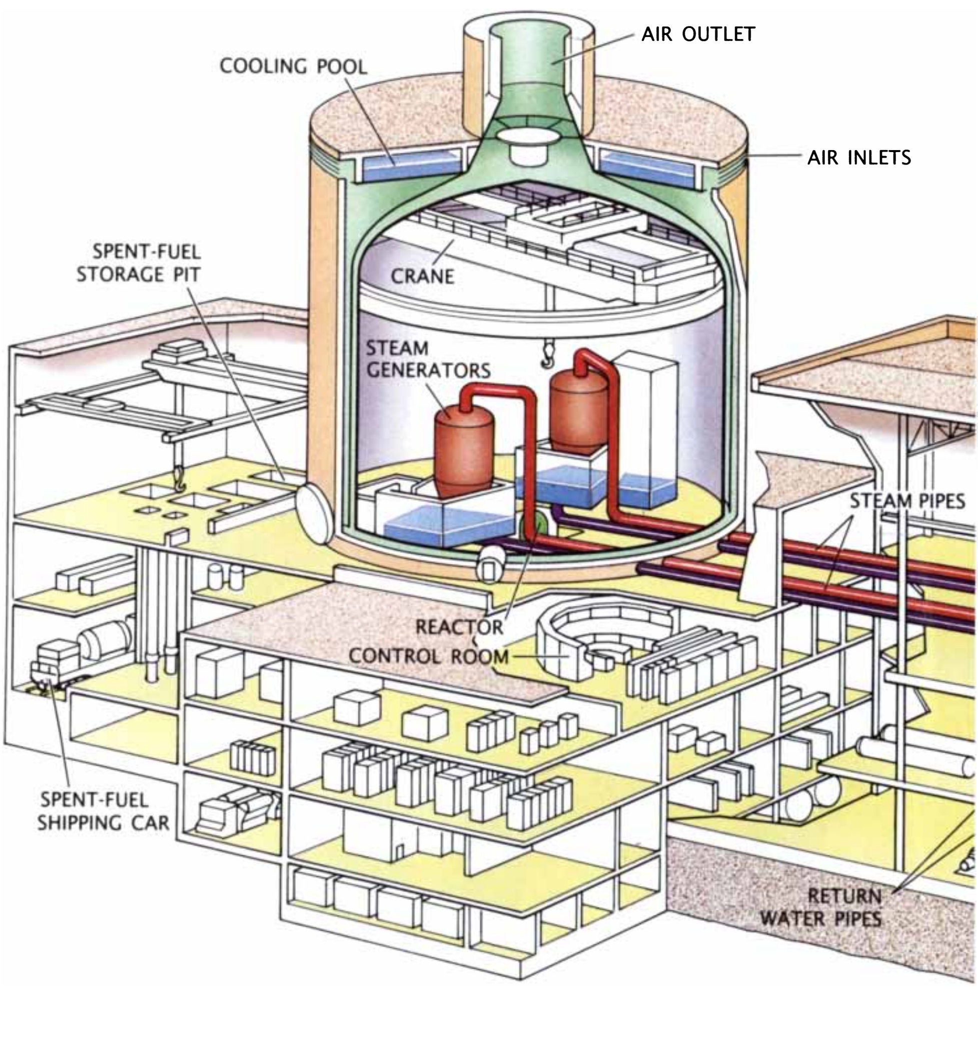
by Michael W. Golay and Neil E. Todreas

One fifth of the world's 560 nuclear power plants operate in the U.S., supplying one fifth of that country's electricity. Many industrialized countries are even more reliant on nuclear power—France and Belgium, for example, derive nearly 80 percent of their electricity from this source. Although the vast majority of nuclear power plants have operated safely and economically, a few notable failures have given the technology a bad reputation and stymied its development for more than a decade. Concerns about nuclear power are balanced, however, by fears that the burning of fossil fuels—the main alternative to nuclear energy—might lead to global warming. Will nuclear power recover its former position in the list of attractive energy options if the public becomes more concerned by its absence than by its presence?

Worries about the safety of nuclear power were seriously exacerbated by the reactor accidents at Three Mile Island in 1979 and at Chernobyl in 1986. At Chernobyl the radioactive pollution was widespread and more than 30 people died; at Three Mile Island the damage was confined to the plant site and nobody was harmed. The difference in the number of casualties can be attributed to the fact that the core of the Chernobyl plant, unlike that at Three Mile Island, lacked an effective containment building—a design that had been strongly criticized before the accident by the United Kingdom Atomic Energy Authority. Even though the Soviet design differed markedly from those in the West, its dramatic failure shocked public opinion worldwide and led to calls for more stringent safety standards.

The experience of industrialized countries has also revealed vulnerabilities in current nuclear power plants that were not fully understood when the plants were designed. Among the most important are human error, materials degradation, accidents resulting from chains of subtle failures and the economic sensitivity of nuclear projects to disruption and poor workmanship. The identification of these problems led to demands that nuclear power must be much safer and more economical to be acceptable.

Some critics attribute the industry's shortcomings to the light-water reactor (LWR), a design employed in all U.S. and most foreign nuclear power plants. The critics favor the development of radical designs that do not employ water as a coolant. Other ana-



NEW LIGHT-WATER REACTOR (LWR) forms the basis of the power plant shown in this cutaway illustration. The reactor, called the AP-600 because it employs an advanced pressurized-water design and generates 600 megawatts of power, is sealed in a concrete cylinder called the containment structure. A closed loop of pressurized water

lysts advocate the continued refinement of the LWR into so-called advanced LWR designs. The program to develop advanced LWR's has been more vigorous and diverse than those of other reactor technologies, in part because it is based on much greater operating experience.

The engineering problems of nuclear power are complicated; the underlying principles are simple. Reactors produce heat by splitting the nuclei of such fissile elements as uranium and plutonium, packaged as fuel shaped into rods or pellets. Fission begins when a fissile isotope absorbs a neutron and decays into lighter elements, releasing energy in the form of recoiling fission fragments, gamma rays and a shower of energetic neutrons. The neutrons stimulate other atoms to fission, releasing yet more neutrons in a self-sustaining process called a chain reaction. Reactors tame the chain reaction by controlling the neutron population with neutron-absorbing control rods and by reducing the neutrons' kinetic energy with moderating materials. The fuel, the control rods and the moderator together constitute the reactor's core. Flowing coolant extracts the heat liberated in fission to make steam, which in turn spins the turbines that drive the electric generators. (Such cooling must continue even if the chain reaction is shut down, or else the decay of radioactive fission products will overheat the core, causing it to release its contents.)

The experience of nearly four decades has winnowed out designs for four basic types of reactor: the heavy-water reactor (HWR), the gas-cooled reactor (GCR), the liquid-metal-cooled reactor (LMR) and the LWR.

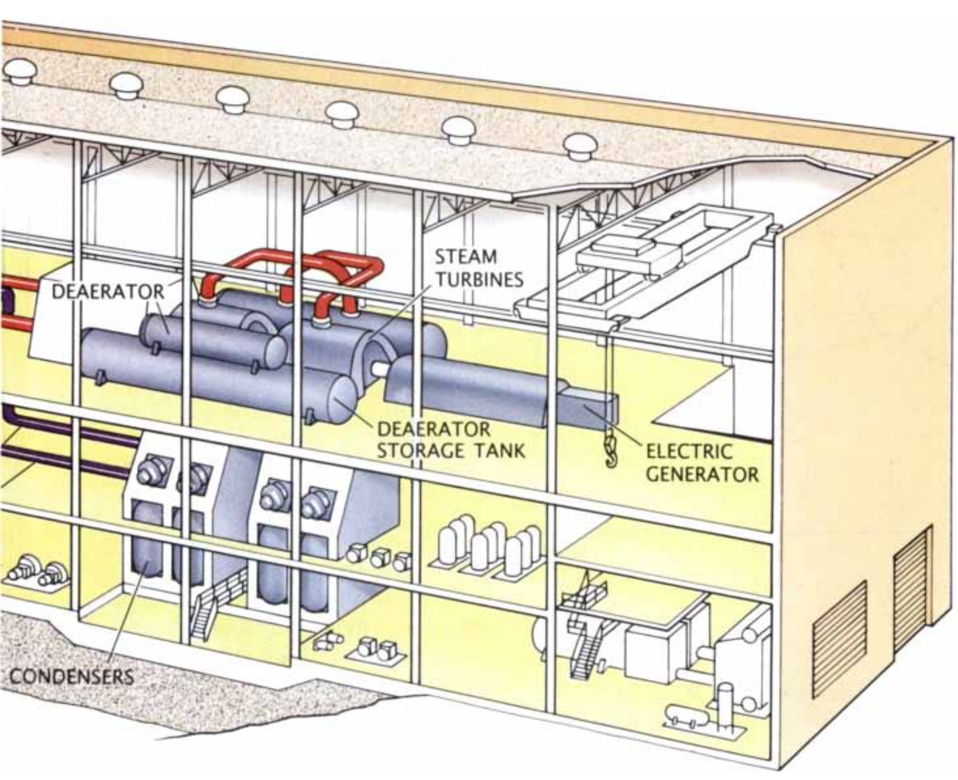
In heavy-water reactors the core is cooled and moderated by "heavy" water, so called because some of the hydrogen atoms have been replaced by deuterium, a rare, heavy isotope of hydrogen. The deuterium's cost is offset by the reactor's ability to use natural uranium (in which the rare, highly fissile isotope uranium 235 has not been concentrated).

Gas-cooled reactors have been built since 1956, at the beginning of the civilian nuclear power era [see "Gas-cooled Nuclear Power Reactors," by Harold M. Agnew; SCIENTIFIC AMERICAN, June, 1981]. Their major advantage lies in their theoretical ability to operate at temperatures above 700 degrees Celsius, considerably higher than the 330 and 550 degrees C generally achieved in LWR's and LMR's, respectively. Because engines that convert heat into other forms of energy work more efficiently at higher than at lower temperatures, the GCR typically converts 40 percent of its heat energy into electricity, compared with 33 percent for LWR's. Although the nuclear power industries of Great Britain and France were founded on versions of this technology, the high capital costs and low reliability associated with the GCR have since led both countries to abandon it in favor of LWR's. High-temperature GCR technology has been explored to a limited extent in the U.S. and West Germany, where it has achieved a rather uneven technical record.

Liquid-metal-cooled reactors, designed in "breeder" configurations, produce more fuel than they consume. They do this by converting nonfissile uranium 238 (more than 99 percent of natural uranium) into fissile plutonium 239, a reactor fuel. Because breeders consume uranium so efficiently, they can give substantial electric-energy independence even to countries that lack indigenous uranium resources. Another advantage of LMR's is the high average energy of their neutrons, which can induce fission and promote rapid decay of many heavy isotopes. Recent work indicates LMR's may be valuable as burners of nuclear wastes.

There are two drawbacks to LMR's. First, the liquid-metal coolant—generally sodium—is reactive with air and water and so requires expensive measures to prevent chemical fires and explosions. Second, the plutonium that the LMR's produce can be made into atomic bombs. Some are afraid breeder reactors would make plutonium so widely available that an interna-

MICHAEL W. GOLAY and NEIL E. TODREAS are professors of nuclear engineering at the Massachusetts Institute of Technology. Golay directs the M.I.T. Center for Advanced Nuclear Power Studies there. He has a doctorate in nuclear engineering from Cornell University and taught at the Rensselaer Polytechnic Institute before joining the M.I.T. faculty in 1971. Todreas headed the department of nuclear engineering at M.I.T. through June, 1989, and chaired the industry review committee that oversaw the Department of Energy's analysis of the accident at Three Mile Island. He has a doctorate in nuclear engineering from M.I.T. and worked in the Atomic Energy Commission's reactor program before joining the M.I.T. faculty in 1970.



conveys heat from the reactor core to steam generators. Pipes (*red*) issuing from the tops of the generators carry steam to turbines on the right; water condensed from spent steam returns in parallel pipes (*purple*). The plant is monitored in the main control room. Details were provided by the Westinghouse Electric Corporation.

tional traffic in the metal would result, impairing efforts to limit the proliferation of nuclear arms. The existence of ample indigenous uranium and political opposition to nuclear proliferation have led the U.S. to limit its LMR development to closed-facility concepts that consume all the plutonium produced. Other countries continuing to press ahead with the technology include Japan, France, West Germany, the U.K. and the U.S.S.R.

Light-water reactors, so called because ordinary water serves as both the coolant and the moderator, come in two versions: the boiling-water reactor (BWR) and the pressurized-water reactor (PWR) [*see illustrations on pages 85 and 86*]. The BWR has a single thermal cycle wherein the core directly heats the water to generate steam. The PWR has two cycles: one employs a circulating loop of pressurized (hence, nonboiling) water to extract heat from the core; the other uses that heat to generate steam.

Although the BWR's direct-cycle design requires fewer large, high-pressure components than the PWR, greater efforts must be taken to shield the BWR's steam pipes (which tend to circulate radioactive material). The extra expense tends to balance the BWR's savings in heavy hardware. Both designs are in service throughout the industrialized world, in versions manufactured by companies in the U.S., France, West Germany, Japan, Sweden and the U.S.S.R.

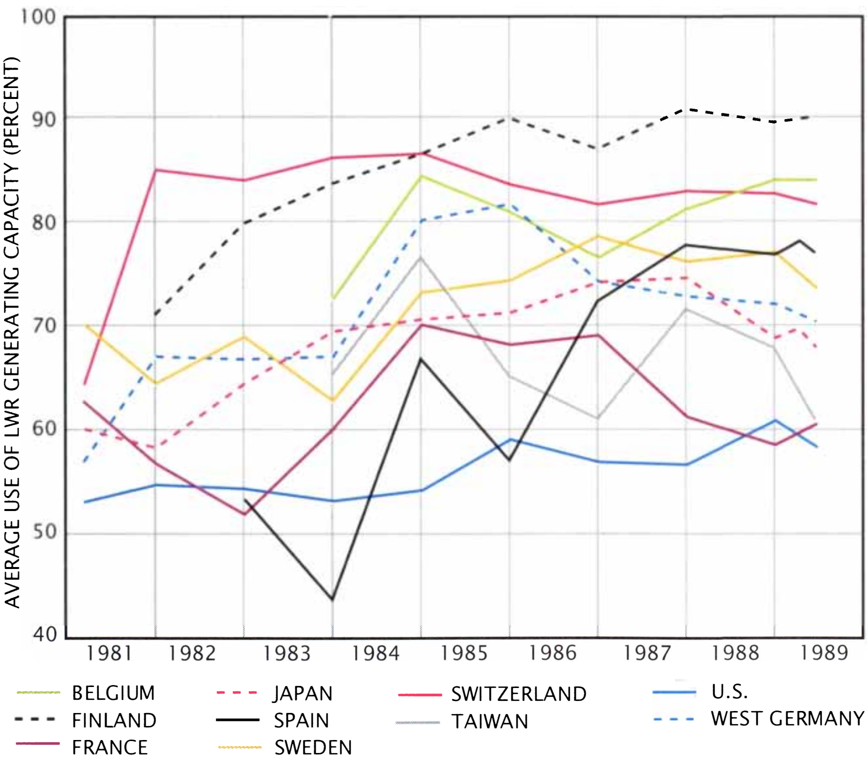
The first commercial LWR, a plant generating 68 megawatts of electricity (MWe), became operational in Shippingport, Pa., in 1957. It is agreed to have been one of the most successful nuclear power plants ever built, having operated safely and reliably for more than 30 years. (It is now being decommissioned because its age and small size have made it uneconomic.) The plant was designed and built by the Westinghouse Electric Corporation in only four years versus the from 10 to 14 years required in the U.S. for reactors now near completion. Other industrialized countries now take about seven years to do the same job.

LWR's were originally designed for the U.S. submarine program, and they have since become standard in nuclear power stations. Three contractors—Combustion Engineering, Inc., the Babcock and Wilcox Corporation and Westinghouse—each developed a version of the PWR for large-scale commercial use. The General Electric Company developed the BWR. In the 1950's and 1960's several companies built LWR plants in risk-sharing partnerships with the federal government and the utilities.

Then, in an effort to establish the market for nuclear power, various contractors offered customers turnkey contracts according to which fully operational plants were delivered at an attractive price, determined beforehand. The growing demand for nuclear plants obviated the need for such contracts by the late 1960's, when the utilities themselves began building plants (with capacities of up to 1,250 MWe). Experience has shown that some of these utilities were unprepared for the resulting financial and technical burdens. The proliferation of designs was such that today virtually every operating plant is different. This made it hard for the utilities to learn from the experience of others and hindered the establishment of uniform safety regulations. As a result of higher costs, growing political opposition and slowing growth in the demand for electricity, no plant ordered since 1974 has been completed.

Still, LWR's are alike in their most important part: the nuclear steam-supply system, which includes the reactor and its cooling system. (The residual systems vary much more.) Further, the similarity of designs for nuclear steam supply is common not only in the U.S. but also in Japan and the European Community, where most of the technology for LWR's has been purchased directly from the U.S. or indirectly, through foreign licensees that have since become independent manufacturers. Nuclear power, therefore, can be added to the list of technologies pioneered by the U.S. and mastered by other nations with which the U.S. now struggles to compete.

It is notable that LWR plants in the U.S. have been able to supply electricity to market for a lower proportion of their operating time than has been the case in many other countries [*see illustration at left*]. A 1986 study, which involved one of us (Golay), found that the disparities result from differences in management and professionalism at individual plants, rather than in political or industrial structure. Some U.S. plants performed as well as any in the world, but oth-



PERFORMANCE OF LWR PLANTS varies substantially among industrialized countries, largely because of differences in management style. The above graph shows the percentage of LWR generating capacity used by the electric utilities of 10 nations. The relatively poor showing of the U.S. is striking because it cannot be explained by differences in hardware, safety regulatory systems or nuclear industry structure. It stems, instead, from the way in which the plants are run: although many perform as well as any in the world, the rest pull the average down. The U.S. has recently improved in other areas, however, such as the number of unplanned shutdowns.

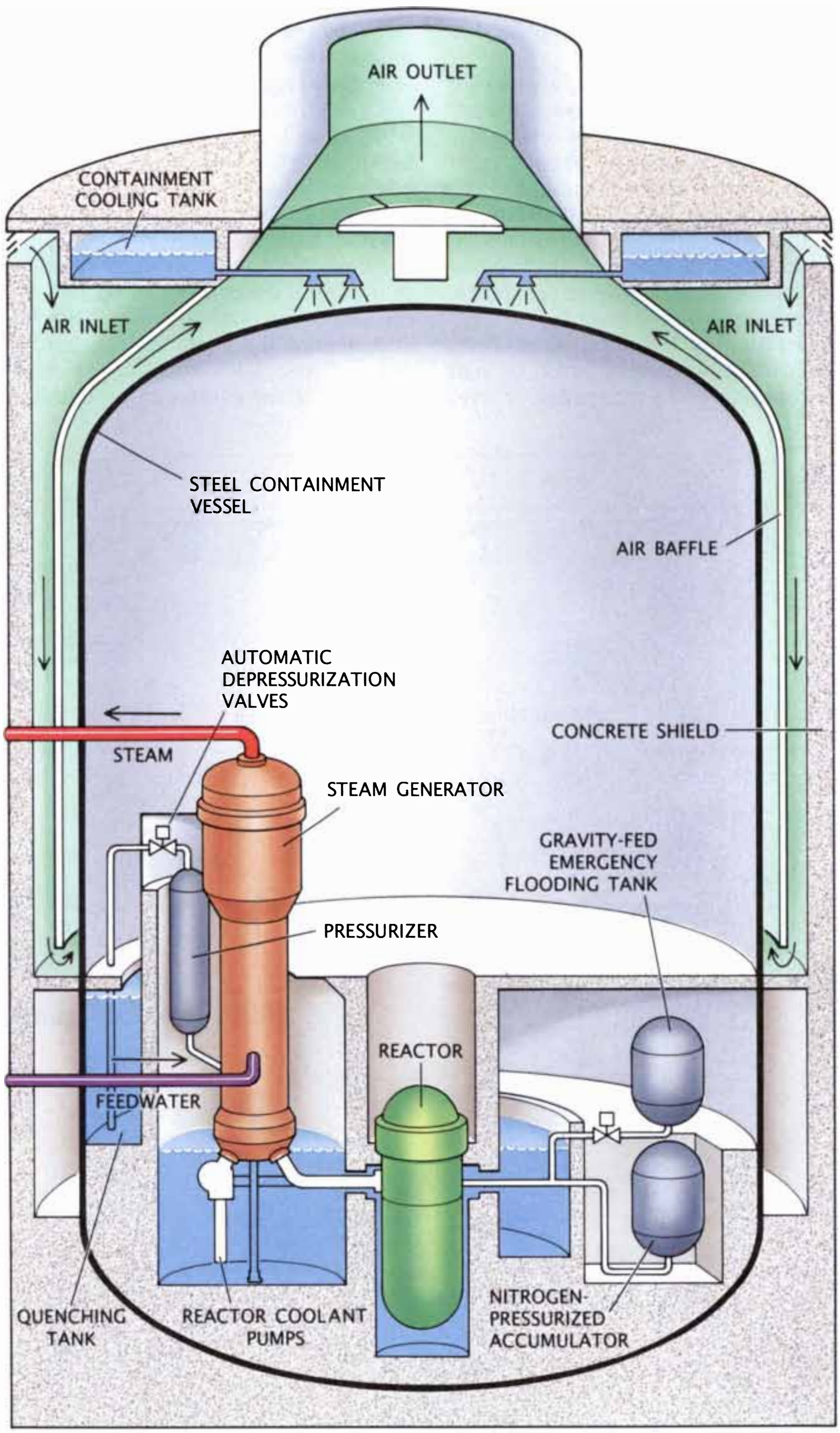
ers performed poorly enough to drag down the average. It seems that vigorous anticipation of problems and attention to detail are key to good plant performance. Good management, therefore, is as important to the success of nuclear power as design is.

Each reactor technology provides examples of success and failure [*see illustration on page 87*]. Most of the failures resulted from either human errors or subtle phenomena that were hard to anticipate; only a few failures can be traced to fundamental conceptual flaws. Therefore, it is wrong to try to predict a technology's prospects with reference to its early difficulties; instead such difficulties should be regarded as an opportunity for improvement. The late Admiral Hyman Rickover, champion of the nuclear submarine, once noted that the most alluring reactor concepts were those that had not yet been built. Only afterward do their hidden faults become evident.

We think each of the various reactor types has a place in the technological armamentarium. Their virtues are complementary: LWR's are understood best of all. GCR's can provide high-temperature heat and offer potential for improved simplicity, safety and fuel economy. LMR's can breed fuel, perhaps burn nuclear wastes and offer safety improvements. The promise of the GCR's and LMR's will remain uncertain, however, until they gain operational experience. Thus, it seems prudent to advance all three technologies simultaneously: advanced LWR's could be deployed immediately, GCR's or LMR's in the next century.

How can a nuclear power concept meet the differing safety and economic requirements of the utilities and the general public? Each side is concerned with the same issues but emphasizes them differently. Each has an effective veto over the other, and neither regards nuclear power as essential in the short term.

What do the utilities and the public want, and how might their differences be bridged? The utilities want any proposal for a nuclear power to be economically attractive not only today but throughout the plant's operating life. They fear that changing regulatory standards (driven either by genuine safety considerations or by shifts in public opinion) may force them to retrofit their plants at enormous expense or shut them down altogether. They also worry that state public-utility commissioners may deny them recovery of their investment (in the form



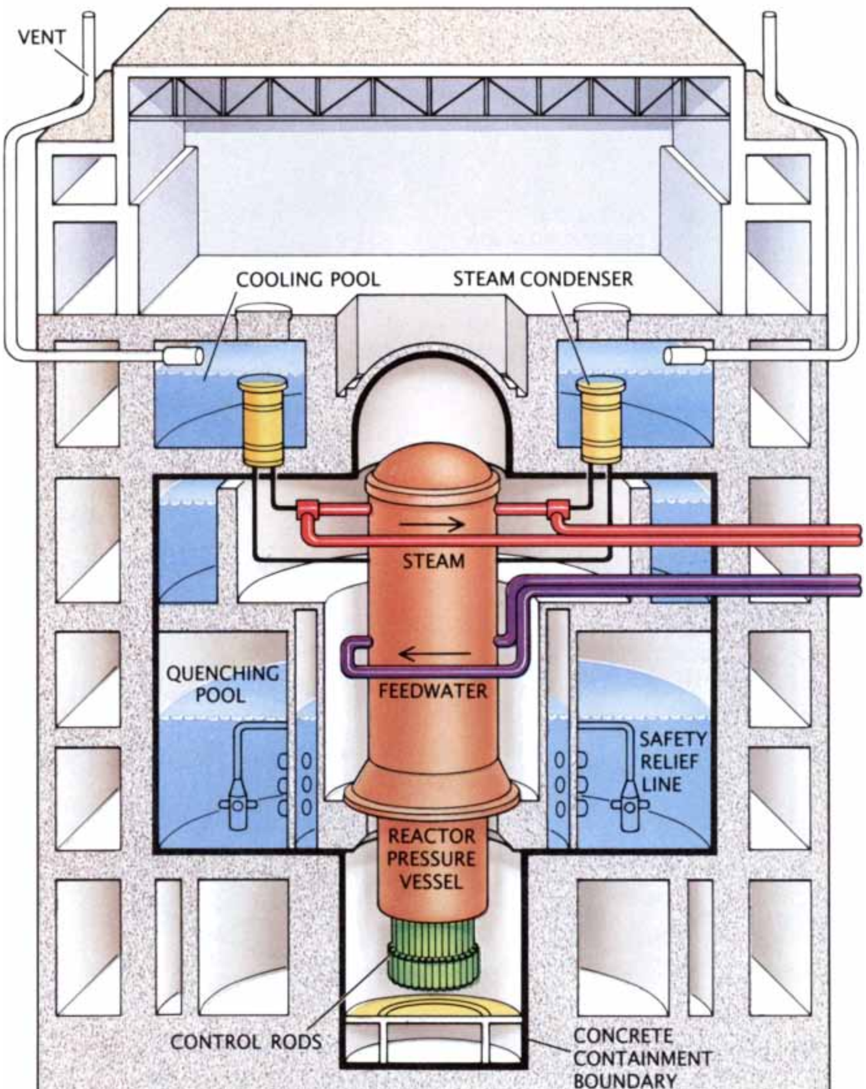
PROPOSED PRESSURIZED-WATER REACTOR design calls for semipassive safety devices, which operate with few moving parts and little human intervention. The design places the reactor at the bottom of the containment building so that gravity-fed cooling water continuously covers its core. Depressurization valves on the steam generator open above a certain pressure, venting steam into a quenching tank that condenses it into water. The 1.7-inch steel walls of the containment vessel enclose a much larger space relative to the reactor's power than is found in conventional plants; the vessel can therefore withstand a prolonged buildup of pressure. Foot-thick concrete walls protect the reactor from natural calamities. Should an accident occur, a double wall formed by a steel air baffle functions as a chimney through which air can move by natural convection. These currents, together with the evaporation of water fed by gravity to overhead sprinklers, cool the vessel and reduce the pressure inside it, preventing a containment rupture through which radioactive materials could be released. The sketch is based on a Westinghouse design.

of higher rates) should the plant cost more than expected.

The public, in contrast, wants electricity at an acceptably low price and with certain guarantees of safety. Yet people apparently cannot agree on how much safety is enough. Some ask that a plant be unlikely to have an accident during its lifetime; others demand that it be theoretically impossible for the plant to have an accident at all (an impossible demand).

Because this debate concerns values that cannot be reduced purely by logic to common units (of money or even human lives), it cannot be settled by quantitative arguments. But one can attempt to analyze safety as rationally as one analyzes costs by applying what is called probabilistic risk assessment (PRA). This technique, developed in the aerospace industry, estimates the likelihood and severity of most accidents. The information is incorporated into the preliminary draft of a plant's design, PRA is applied to the new design and the process is reiterated until the goals for acceptable risk are met. Such goals are framed in terms of the estimated probability of each possible accident, multiplied by the gravity of its consequences. The chain of possible events resulting from a failure of any of the components—coolant seals, pressure valves, automatic controlling devices and so on—can be identified from knowledge of the plant's design. The results of PRA are limited, however, by uncertainties in the data and in the understanding of human behavior and certain physical phenomena. Still, the exercise itself will sometimes suggest simple ways to reduce risk. One such way is to incorporate "passive" safety features, which rely on gravity or natural convection rather than "active" features, such as pumps. Semipassive systems incorporate only a few active devices—valves, for example.

It is important to note that risk reduction often benefits a plant's owner even more than the public because it is far more likely that an accident will destroy the value of a plant than injure the people who live nearby. The accident at Three Mile Island, for instance, eliminated the equity in the plant but did not harm the surrounding population.



SIMPLIFIED BOILING-WATER REACTOR (SBWR) design envelopes the core in three concentric structures: a reactor pressure vessel, in which heat from the reactor directly boils water into steam; a concrete chamber (*outlined in black*) and water pool, which contains and quenches steam vented from the reactor in an emergency; and a concrete building, which acts as a secondary containment vessel and shield. Excessive pressure in the reactor automatically opens valves that release steam into a quenching pool, reducing the pressure. Water from the quenching pool can, if necessary, flow downward to cool the core. The buildup of containment pressure is limited by cooling provided by evaporation from a pool on top of the containment building. The illustration is based on a design by the General Electric Company.

For a nuclear plant to provide electric power that is both economical and safe, it must satisfy many subordinate requirements: inexpensive construction, maintenance and fueling; long life and stable operation; and reliable shutdown and cooling. Safety and efficiency goals do not conflict during the design process until each factor is pushed beyond acceptable limits; thereafter, further gains in one area usually come at the expense of another. At that point, designers maximize one factor as much as they can without compromising the others to an unacceptable degree.

Historically, economic goals have received primary emphasis—accompanied by a serious concern for safety. This strategy is now pursued most vigorously in Japan, where the Ministry for International Trade and Industry (MITI) has supported two new concepts: the Advanced PWR (APWR), a Westinghouse design, and the Advanced BWR (ABWR), a General Electric design. Both designs are expected to achieve significant improvements in the efficiency of fuel consumption and sizable reductions in capital and operational costs. The Tokyo Electric Power Company is expected to receive a construction permit for the world's first ABWR in 1991.

In the APWR the cores are longer-lived and larger in volume than those of conventional PWR's, which serves to minimize the lost availability of power during shutdowns for refueling. Use of improved steam generators also promises long-life operation. The electric signals by which the plant's components are monitored and actuated have also been greatly simplified. The many copper wires typical of older reactors, for example, have been replaced by a compact, multiplexed fiber-optic system. Automatic mechanical and electronic devices have supplanted human operators in the performance of many emergency duties. Reliability has been enhanced by designing according to more conservative thermal and stress criteria.

The ABWR shares many control innovations of the APWR. Unlike conventional BWR's, the ABWR recirculates water through sealed pipes within the reactor vessel instead of through external pipes, which are less secure. Reactor power is controlled by varying the volume of bubbles in the water within the core, thereby controlling the degree to which the water moderates the chain reaction. The control systems are driven electromechanically rather than hydraulically, thereby reducing maintenance. Safety systems are more redundant (there are more backup systems) and make fewer demands on the plant's operators.

In the U.S. the effort to promote the adoption of still more advanced LWR designs is more diverse but less vigorous than in Japan. The cooperative effort involves the Electric Power Research Institute (EPRI), which is the research agency of the U.S. electric-utility industry, the U.S. Department of Energy, Combustion Engineering, Inc., Westinghouse and General Electric. The group is focusing its efforts on designs that incorporate both evolutionary change and semipassive safety features; the efforts have also led to formulation of corresponding plant-performance requirements. Top priority has been given to obtaining complete approval by the Nuclear Regulatory Commission of any design before construction begins in order to limit the financial risk involved.

France has developed an evolutionarily improved PWR. Called the N4, it has a capacity of 1,450 MWe, greater than that of any other LWR. It uses automation to reduce the possibility of human error and to improve core-cooling reliability; it also incorporates a filter into the containment building to trap radioactive releases from severe core accidents.

Several programs focus on both economic and safety performance. In Great Britain, the Central Electric Generating Board (CEGB) has begun a program to replace that country's system of advanced gas-cooled reactors with PWR's. Lately, privatization and the breakup of the CEGB have put most of the program on hold. The exception, a plant called the Sizewell B unit, is scheduled for completion in 1994; it is the first in the world to be designed with the extensive application of probabilistic risk assessment. Among the improvements wrought by this statistical technique were the physical separation of safety-related systems from other systems, the multiplication of redundancy in safety systems, the reduction of human involvement in the early stages of accident control and the construction of a double-walled containment structure.

One way to reduce risk is to build smaller reactors, which contain less radioactive material and in principle can be cooled more easily in case of a malfunction. There may even be an economic justification for such a strategy. The 54 nuclear utility companies in the U.S. have system capacities ranging from 2,000 to 32,000 MWe, and individual units have generating capacities of about 1,000 MWe (enough to supply the needs of about 500,000 households). Critics have said such units are too large: many utilities prefer to increase their capacities in smaller increments.

The open question is whether the capital costs of such small reactors can be made competitive with those of larger units, which can enjoy certain economies of scale if they are managed well. There is some reason to think so. Small plants could be built at the factory and shipped to their destinations by rail or barge (most power plants are near bodies of water in order to have ample water for cooling). Such standardized manufacturing is generally faster, cheaper and more precise than on-site construction. Still, the problem of attracting the investment capital and of sustaining such factories remains unsolved.

EPRI has encouraged two efforts to develop smaller advanced LWR's with capacities of 600 MWe or less, one under the direction of General Electric, which is developing a BWR, the other under Westinghouse, which is developing a PWR. Each design calls for plants that will stabilize for up to three days without the intervention of operators, that rely on semipassive systems and that have greatly fortified containment methods to retain radioactive materials even in the event of an accident. Since both reactors have fewer parts than conventional plants, they would be easier to maintain and less likely to break down.

The simplified BWR employs semipassive systems, such as a reservoir of water placed above the core, so that

LIGHT-WATER REACTORS

SUCCESSES: Plants in France, Switzerland, West Germany, Japan, Sweden, Finland, Korea and most plants in U.S.

ECONOMIC FAILURES: Midland, Mich.; Zimmer, Ohio; Rancho Seco, Calif.; Marble Hill, Ind.; Shoreham, L.I.

SAFETY FAILURE: Three Mile Island 2, Pa.

LIQUID-METAL-COOLED REACTORS

SUCCESSES: Phenix, France; Dounreay, U.K.; Fast Flux Test Facility, Wash.; Experimental Breeder Reactor II, Idaho.

ECONOMIC FAILURE: SNR 300, West Germany.

SAFETY FAILURES: Fermi I, Mich.; Experimental Breeder Reactor I, Idaho.

GAS-COOLED REACTORS

SUCCESSES: Peach Bottom, Pa.; AVR, West Germany.

ECONOMIC FAILURES: Fort St. Vrain, Colo.; the gas reactor programs in France and U.K.; Thorium High-Temperature Reactor-300, West Germany.

SAFETY FAILURE: Windscale, U.K.

SUCCESSES AND FAILURES of the major reactor technologies are listed in this table. A plant is an economic failure when it costs so much or provides power so infrequently that it is not worth operating; it is a safety failure when an accident ends its operation. Three Mile Island plant 2, for example, is a safety failure because an accident in 1979 put it out of commission. The Shoreham, L.I., plant is an economic failure because state authorities will not allow its owners to recoup their investment.

EMPHASIS ON ECONOMICS

All existing LWR's
Large-plant (1,200 –1,400 megawatts electric) LWR programs in Japan, France and U.S.
Small (200–600 MWe) barge-mounted PWR's (U.S., U.K., France and West Germany)

ATTEMPTS TO BALANCE SAFETY AND ECONOMICS

Small (600 MWe) PWR (AP-600, U.S.)
Small (600 MWe) BWR (SBWR, U.S.)
Sizewell B (U.K.)

EMPHASIS ON SAFETY

PIUS (PWR, Sweden)
PIUS (BWR, U.S.)
ISER (PWR, Japan)
SIR (PWR, U.K. and U.S.)

SAFETY AND EFFICIENCY lack common units of measurement and therefore are balanced subjectively, as in the three basic strategies exemplified by the designs listed above. Concepts that mainly emphasize economic efficiency dominated early LWR development because no other strategy could create a market for the new technology. Designs primarily emphasizing safety have not yet been implemented commercially, in part because they are considered to be economically unattractive.

the water can be fed to the reactor by gravity rather than by pumps, which might fail. Other gravity-driven features include a high-pressure core-cooling system and a system that cools the containment building by evaporation of water stored in an elevated pool. The design eliminates the possibility of a failure in the recirculation pumps by instead basing core recirculation on natural convection.

In the AP-600 [*see illustration on page 84*], passive safety measures include a chimney that cools the containment building through the flow of outside air (moved by natural convection rather than fans). During an accident such cooling, facilitated by gravity-fed sprinklers that wet the top of the containment building, lowers the pressure inside the vessel and thereby reduces the likelihood of a containment rupture. A more subtle, elevated passive cooling system employs hydrostatic pressure—the weight of a column of water—to force emergency cooling water into the core. Some possible loss-of-coolant accidents are also eliminated by using sealed pumps.

The earliest and most radical passive safety plan is embodied in PIUS (Process Inherent Ultimately Safe), a PWR concept designed in Sweden but never built. The design, which takes into account a considerably larger range of potential accidents than do the other reactor concepts discussed here, specifies that the plant be able to shut itself down and automatically cool itself even if it is sabotaged or simply abandoned [see "Rethinking Nuclear Power," by Richard K. Lester; SCIENTIFIC AMERICAN, March, 1986]. PIUS is unusual since it generates only 500 MWe and operates at a lower temperature and coolant pressure than do current PWR's. The reactor and primary cooling loop are immersed in and communicate with a pressurized pool (contained by a prestressed concrete vessel) of water laced with boric acid, which not only absorbs heat in the event of a mishap but also infiltrates the primary loop if pressure falls as the result of a failure in a component. Natural convection then circulates pool water through the core, where the boric acid can absorb neutrons, shutting the reactor down while cooling it. Only a failure in the prestressed concrete vessel—which is considered very unlikely—could prevent the passive system from working.

The major drawback to PIUS may be its difficult maintenance and its apparent tendency to shut down very easily. Small perturbations of the primary cooling system might lead to unnecessary shutdowns. Once such a shutdown occurs, it could take at least a day to remove boron from the primary system in order to restart the plant. It is not yet clear whether such concerns are justified.

A number of important factors will affect the development of advanced reactor technologies in the U.S. First, the U.S. has an abundant fossil-fuel supply, particularly coal, and therefore has less to lose from an unwise technology policy than do other, less well-endowed countries. Nuclear power is economically important to the U.S., but it is not strategically vital. Nevertheless, the development of advanced reactors is important because it provides technology for future energy needs and supports a potential export industry that can reduce the trade deficit and exert influence on the nuclear-policy decisions of other countries.

Most national nuclear power programs are in decline, a process that typically begins when demand for new plants dwindles. Existing industrial capacity then becomes superfluous, investment plummets and the intellectual and economic foundations for the development of advanced technologies weaken. In the U.S. this process has been exacerbated by inadequate funding and lack of leadership.

Periodically the federal authorities are pressed to focus the country's resources on the one or two "best" reactor technologies. But in the absence of a sustained program to provide the technical basis for such a choice, this approach is wrong for two reasons: there is no consensus on the criteria by which the technological options are to be winnowed, and the notion of selection is itself a distraction from the central problems posed by details, subtle vulnerabilities of the system and human performance.

Each of the nuclear reactor technologies has its merits. One cannot anticipate which reactor type will turn out to be most valuable under future conditions. Rather than reducing the number of alternatives, current reactor development efforts should strive to maintain many technological options from which to choose. Its extensive technological base makes the light-water reactor a leading option that should be exploited thoroughly.

FURTHER READING

THE SAFETY OF FISSION REACTORS. Harold W. Lewis in *Scientific American*, Vol. 242, No. 3, pages 33–45; March, 1980.

A PROCESS INHERENT ULTIMATE SAFETY BOILING WATER REACTOR. C. W. Forsberg in *Nuclear Technology*, Vol. 72, No. 2, pages 121–134; February, 1986.

INTERNATIONAL COMPARISON OF LWR PERFORMANCE. K. Hansen, E. Beckjord, M. Golay, E. Gyftopoulos, R. Lester and D. Winje. MIT Energy Laboratory Report MIT-EL 87-004, February, 1987.

AP-600 OFFERS A SIMPLER WAY TO GREATER SAFETY, OPERABILITY AND MAINTAINABILITY. Ronald Vijuk and Howard Bruschi in *Nuclear Engineering International*, Vol. 33, No. 412, pages 22–28; November, 1988.

SIMPLICITY: THE KEY TO IMPROVED SAFETY, PERFORMANCE AND ECONOMICS. R. J. McCandless and J. R. Redding in *Nuclear Engineering International*, Vol. 34, No. 424, pages 20–24; November, 1989.

What the Brain Tells the Eye

A circadian clock adjusts the sensitivity of the horseshoe crab's eyes. Study of the mechanism opens a window on the brain's control of sensory organs in other animals as well

by Robert B. Barlow, Jr.

The horseshoe crab (*Limulus*) was once thought to possess a primitive, simple eye that had been largely bypassed by evolution. In fact, evolution has served the crab well. Anatomical and physiological studies are showing that the 350-million-year-old animal has developed a complex, sophisticated visual system that incorporates elegant mechanisms for adapting its sensitivity to daily cycles of light and darkness.

Humans see only dimly at night, but the world of horseshoe crabs may be nearly as bright at night as during the day. Inquiry into the mechanisms by which *Limulus* performs this feat has added to knowledge of a most intriguing phenomenon. The brain and its sensory organs are not merely passive recipients of information from the outside world. Instead the brain actively controls those organs to optimize the information it receives.

Over the past decade my colleagues and I have explored in considerable detail how horseshoe crabs adapt their visual systems. Our most important finding is the discovery of a 24-hour biological clock in the crab's brain that transmits nerve signals to its eyes at night. These signals work to increase the eyes' sensitivity to light by a factor of up to one million. Oddly enough, this extraordinary nighttime increase in sensitivity went undetected until the late 1970's, even though the horseshoe crab's visual system is among the most thoroughly studied in the animal kingdom.

ROBERT B. BARLOW, Jr., has studied the visual system of horseshoe crabs for more than 25 years. He received a bachelor's degree in physics from Bowdoin College and a doctorate in biophysics from the Rockefeller University and is now professor of neuroscience in the Institute for Sensory Research at Syracuse University. He is also research professor of ophthalmology at the SUNY Health Science Center in Syracuse. From May through August, Barlow conducts his work at the Marine Biological Laboratory in Woods Hole; he reports that writing field notes underwater is not difficult on nights when the moon is full.

Indeed, *Limulus* has been a favorite of neurophysiologists for more than 50 years. H. Keffer Hartline inaugurated work on the *Limulus* eye in 1926 when he went to the Marine Biological Laboratory in Woods Hole in search of a "simple" eye that would lend itself to studying cellular mechanisms of retinal function. The eye comprises about 1,000 photoreceptor clusters, called ommatidia, each roughly 100 times the size of the cones and rods of the human eye. These receptors are the largest of any known animal.

Hartline and his colleagues missed discovering the influence the *Limulus* brain exerts on the eye because they excised the organ from the crab before recording the responses of its sensory fibers. Two graduate students at Syracuse University, Stanley J. Bolanowski and Michael L. Brachman, discovered the eye's circadian changes when I suggested as part of their coursework that they try to record responses from the eye without removing it from the animal. My colleagues and I had suspected that the eye's properties might change from day to night, but the two students were astonished to find that the eye's sensitivity varied on a daily cycle even when the animal was kept in complete darkness.

Such a circadian rhythm should not have come as a surprise, even in a creature as ancient as *Limulus*. Internal clocks in humans and other animals control bodily functions ranging from waking and sleeping to temperature and blood sugar levels. In mammals the clock is located in the suprachiasmatic nucleus, a small cluster of nerve cells located above the crossing point of the optic nerves. In birds it is located in the pineal gland, in the sea slug it is in the eye and in the horseshoe crab the clock lies in the anterior portion of the brain. The horseshoe crab's clock forms direct neural links with all the major eyes of the animal—lateral, median and ventral.

Most of the research on circadian rhythms in the horseshoe crab has focused on the brain's output to the lateral eye, whose daytime function is already very well understood. W. Henner Fahrenbach of the Oregon Regional Primate Research Center first found the dozen or so efferent fibers in the optic-nerve trunk that carry signals from the brain to the lateral eye. The fibers enter the retina and branch out profusely, forming as many as 100 synaptic connections with the cells of each ommatidium, about 100,000 connections in all.

Severing the optic nerve prevents the brain's signals from reaching the eye and blocks rhythmic changes in sensitivity. Yet if those signals are reproduced artificially, the eye responds just as it would to the brain's command. At night Bolanowski, Brachman and I record the neural activity of the clock from the cut end of the optic nerve connected to the brain; the next day we attach an electrode to the end of the nerve leading to the eye and play back the neural activity. The eye's sensitivity increases immediately; signals recorded from the brain at night can convert the eye to its highly sensitive nighttime state at any time.

A convenient measure of these daily changes in the sensitivity of the eyes is the electroretinogram, or ERG. To record an ERG, one places a conductive electrode on the corneal surface; a flash of light excites all the photoreceptors in the eye and generates a summed electric potential that is then recorded. The ERG is easy to measure and does *Limulus* no harm; data can be obtained from the same animal over long periods.

We have found that retinal responses (as measured by the amplitude of the ERG) are greater at night than during the day both in horseshoe crabs exposed to natural light and in those kept constantly in the dark. The amplitude begins to increase about the time of dusk, reaches a maximum during the early evening hours, begins to decline after midnight and drops back to its daytime level by dawn. This cycle of changes in retinal sensitivity can continue uninterrupted for at least a year; that is the longest period we have recorded potentials from a single animal left in the dark.

The internal rhythm has a period of about 24 hours; when the crab is exposed to natural light, signals from photoreceptor cells in its tail help to keep the circadian clock synchronized with the actual cycle of light and darkness. (Two small eyes on top of the crab, the median ocelli, send additional signals to the brain that increase the strength of the circadian response depending on the amount of ultraviolet light in the sky at night.)

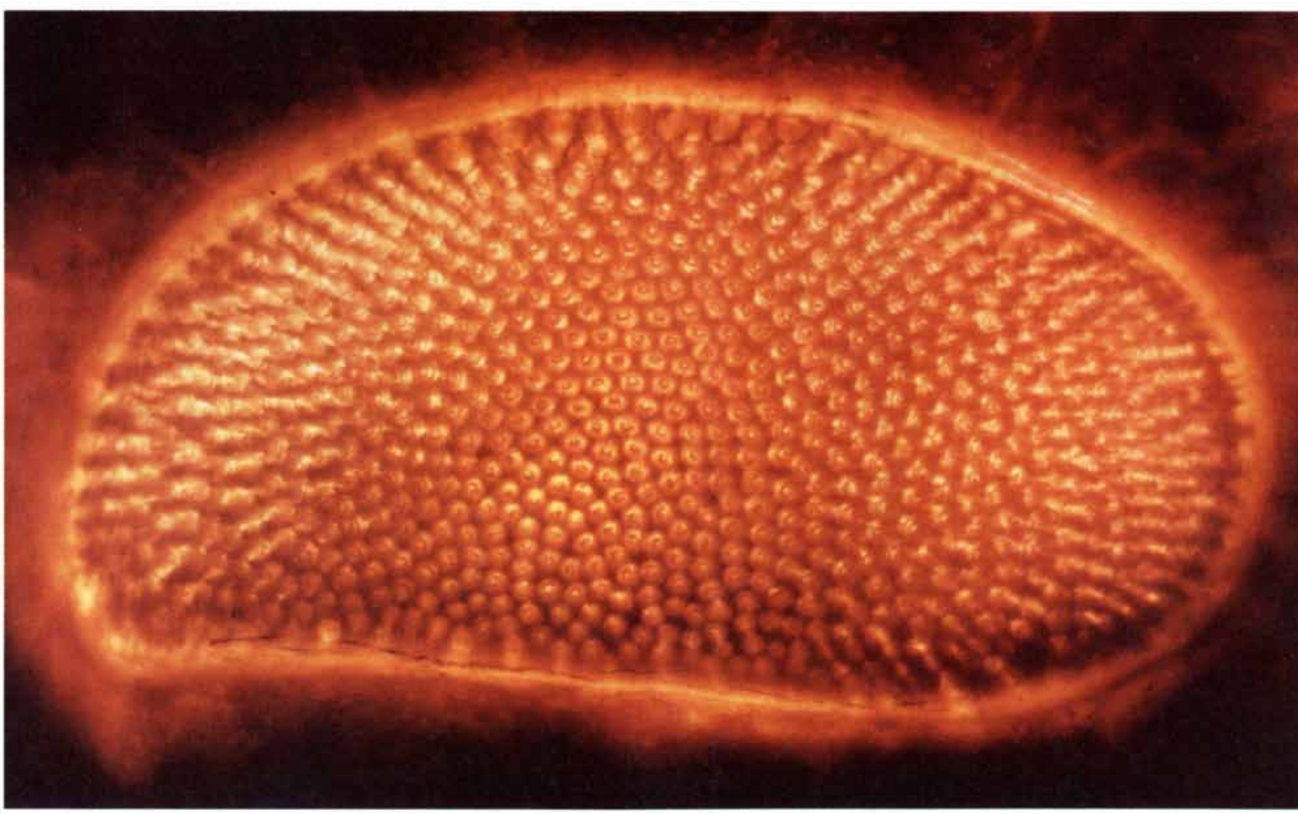
Once it had been determined that an internal clock controls the sensitivity of the *Limulus* eye, work began to elucidate the mechanisms underlying that control. At Woods Hole, Ehud Kaplan of the Rockefeller University, George H. Renninger of the University of Guelph, Takehito Saito of Tsukuba University and I inserted small glass electrodes through a hole in the cornea into the underlying photoreceptor cells. The cells' pattern of intrinsic noise (electric signals generated at random) and their response to light exhibited a distinct circadian rhythm. The most remarkable aspect of these changes was that they contravened the usual rules governing the behavior of both biological and artificial photodetectors. Increased sensitivity is usually accompanied by increased noise, but the signal generated by the crab's eye increases while its noise level decreases.

In both vertebrates and invertebrates, the first recordable signal of a visual stimulus is a change in the voltage across the membrane enclosing a photoreceptor cell. An intense flash causes the transmembrane voltage to decrease by as much as 50 millivolts and then to return to its previous level. Such a potential change is the sum of many so-called quantum bumps, each of which can be triggered by a single photon striking a light-sensitive rhodopsin molecule. Bumps may also occur spontaneously. When a horseshoe crab is kept in the dark, its cells frequently generate spontaneous quantum bumps during daylight hours; the eye in its daytime state is noisy. At night the signals sent by the circadian clock in the brain dramatically lower the rate of spontaneous quantum bumps and so lower the eye's background noise.

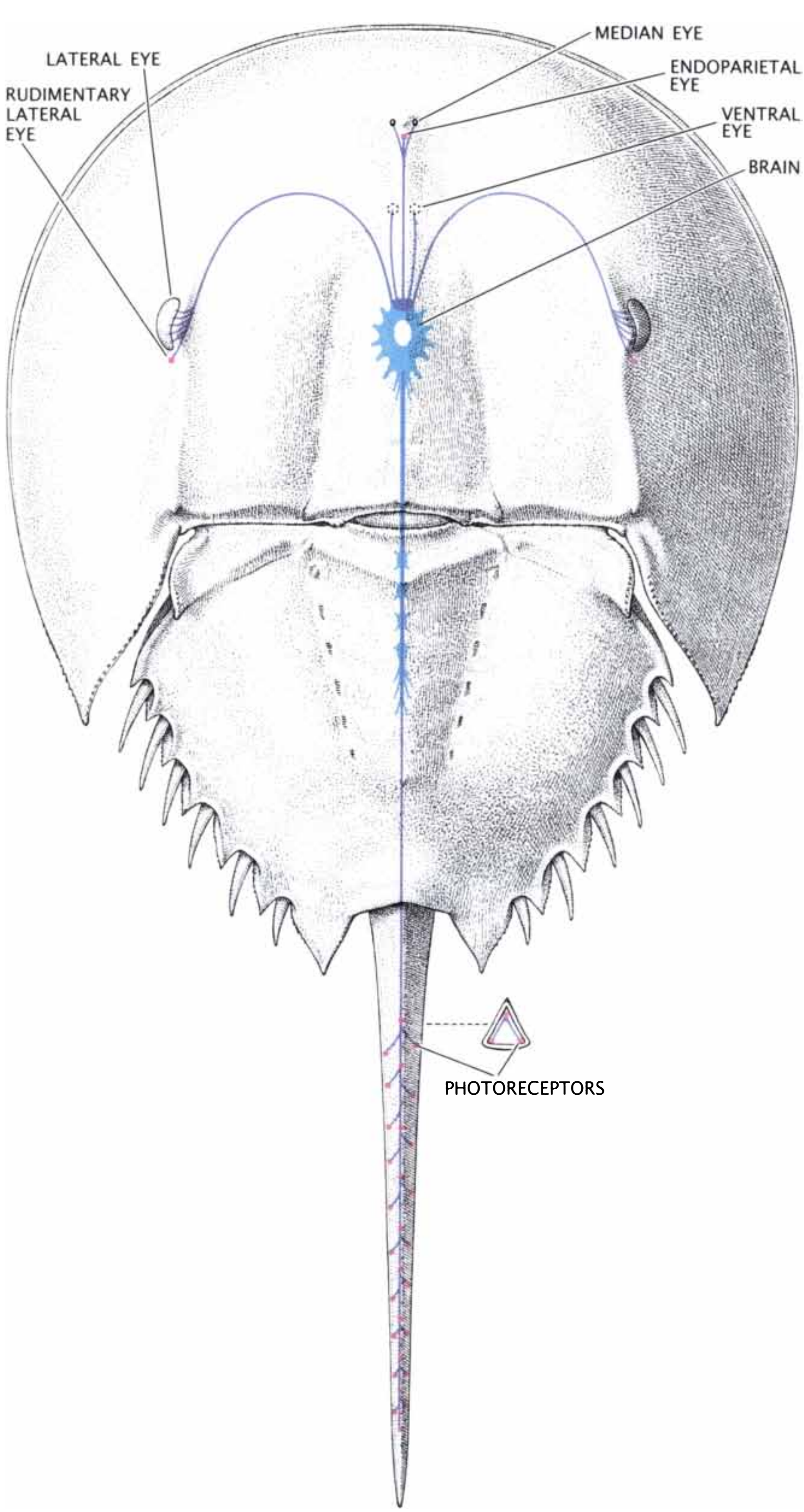
At the same time the eye's internal noise decreases, the response of each photoreceptor increases: structures inside the cell shift to improve the chance that a photon entering the eye will set off a quantum bump, and bumps generated at night last longer than during the day. The clock's output to the retina at night prolongs the quantum bump triggered by one photon so that the bump is likely to overlap with another bump and produce a membrane voltage change that will generate a nerve impulse.

This longer duration of quantum bumps, however, reduces the eye's response to rapid changes in illumination. As Ranjan Batra of the University of Connecticut and I have shown, the crab trades temporal resolution for increased sensitivity.

The other changes in the crab's eye increase sensitivity even further, but



COMPOUND EYE of the horseshoe crab (*Limulus*) is perhaps the best studied of all neural systems. It comprises about 1,000 ommatidia, each consisting of a lens and photoreceptor cells. The author and his colleagues have found that the ommatidia undergo drastic structural and functional changes at dusk and again at dawn in order to adapt to changing levels of light.



EYES OF *LIMULUS* are shown in this sketch of the animal's visual system. A clock in the anterior part of the brain sends out signals that control the sensitivity of the lateral and median eyes. Photoreceptors throughout the length of the tail synchronize the clock to daily cycles of light and darkness. Additional signals to the brain from the small median eyes enhance the degree of adaptation to darkness in accordance with the amount of ultraviolet light those eyes receive from the sky at night.

they do so at the expense of spatial resolution. Steven C. Chamberlain of Syracuse and I found that each ommatidium contains an aperture that works much like the iris in the eyes of vertebrates. During the day the aperture constricts to reduce the number of photons that reach the underlying photoreceptors. At night it dilates to increase the photon catch.

The clock's signals also cause the photosensitive region of the ommatidium (the rhabdom) to fold in on itself and move closer to the aperture. This response maximizes the chance that a photon entering the eye will strike a rhodopsin molecule. Finally, the clock weakens the lateral inhibitory interactions in the retina that enhance its response to edges in the visual field. The eye becomes more sensitive to light but less sensitive to contrast. Furthermore, at night the fields of view of adjacent ommatidia overlap partially so that the crab's vision is slightly blurred.

Such trade-offs are widespread in the animal kingdom. Humans, for example, switch from cone receptors to rod receptors in response to dim light. Cones cluster densely in the fovea at the center of the retina; they provide high acuity but require at least moderate light levels. Cones also have high temporal resolution, but they cannot sum their inputs: each has its own direct connection to the optic nerve. Rods, in contrast, are numerous on the periphery of the retina. Their acuity and temporal resolution are poor, but they can readily detect single photons of light, and several rods can generate a summed output that then excites a single neuron.

Most animals whose eyes contain both rods and cones can rely on ambient-light changes to control sensitivity. In contrast, the horseshoe crab, whose retina contains only a single kind of photoreceptor, adapts to day and night by means of an internal clock. The clock changes the receptors' properties in anticipation of changes in light intensity.

How do the signals from the circadian clock in the brain that travel along efferent nerve fibers to the eye cause such drastic cellular changes? Barbara-Ann Battelle and her colleagues at the University of Florida have studied this question along with me, Renninger and Leonard J. Kass of the University of Maine. We have found that the neurotransmitter octopamine is one key chemical. It couples to membrane receptors in the photoreceptor cell. These sites on the membrane activate the enzyme adenylate cyclase, which in turn mediates the production of cyclic AMP inside the photoreceptor. The cyclic AMP plays its familiar role as a second messenger; in this instance it passes on the effect of the octopamine by triggering cytoskeletal changes that modify the shape of the photoreceptors as well as changes in ion channels that enhance the cell's response to photons. It is likely that another neurotransmitter, as yet unidentified, acts in concert with octopamine to produce the nightly changes in the crab's eye. Injecting octopamine or analogues of cyclic AMP into the eye during the day pushes the retina toward its nighttime state, but the alteration is incomplete.

Octopamine and other factors increase the sensitivity of the eye. What decreases it? Herman K. Lehman of the University of Arizona, Chamberlain and I have developed a model based on a push-pull mechanism: signals from the brain beginning around dusk release neurotransmitters that push the retina to its nighttime state. These signals wane toward dawn, allowing a circulating hormone, also as yet unidentified, to pull the retina back to the daytime state in anticipation of dawn. (Bruce G. Calman of the University of Florida, Lehman and I found that eyes kept in a tissue-culture medium and deprived of circadian signals do not return to the daytime state unless certain blood extracts are added to the medium.)

It is essential that the eye return to the daytime state before dawn; its function can be seriously impaired if it is exposed to even moderate levels of light in the nighttime state. During the day, it appears, the circulating hormone reduces the eye's sensitivity and thus protects it from light.

Although the circulating hormone is directly responsible for the transition to the daytime state, the horseshoe crab's clock controls the metabolic process that renews the eye's light-sensitive structure. Chamberlain and I found that the clock's nightly signals prime the retina for a massive disassembly and rebuilding of rhodopsin-containing membrane that is triggered by the dawn's light. The clock's signals are also a necessary precursor to additional protective structural adaptations that the ommatidia undergo during the day in response to bright illumination.

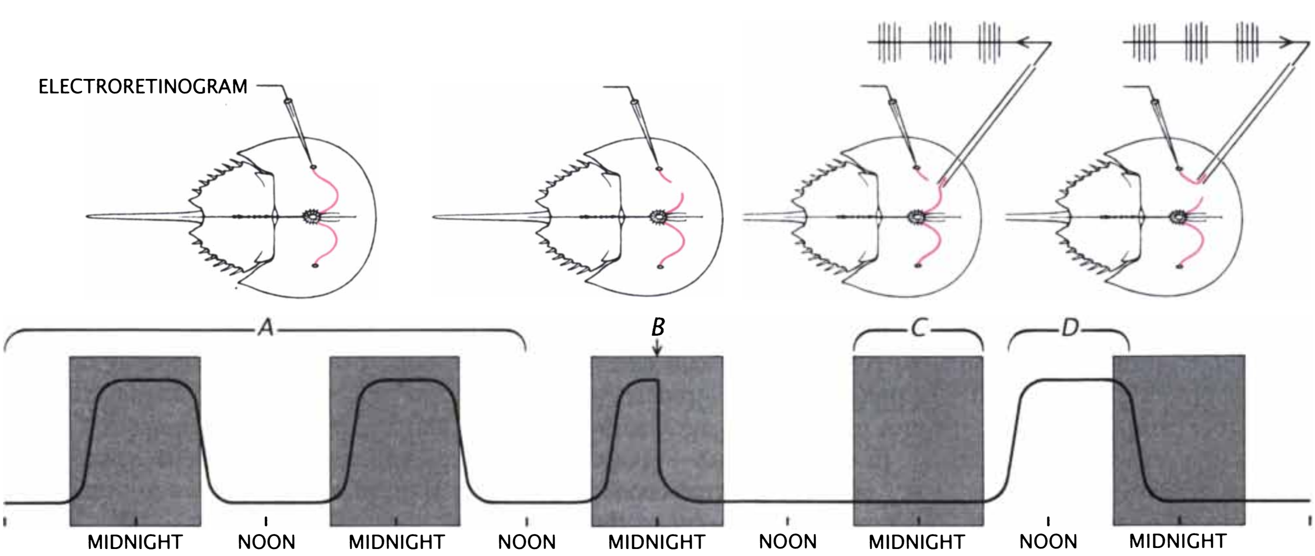
Discovery of a well-orchestrated daily cycle in the horseshoe crab's eye intensified a long-standing question: What does *Limulus* use its eyes for? Hartline often joked that he was "studying vision in a blind animal." Although the visual system of the horseshoe crab has been studied intensely for more than 60 years, no one had, until recently, found a role for vision in the animal's behavior. It does not apparently use its eyes to find food. When diving with horseshoe crabs, I found they would move away from shadows that passed over one of their eyes—such as those made by shifting a clipboard underwater in moonlight. Although the response appears designed to avoid predators, I do not know of any animal that would prey on *Limulus.*

One complex activity that horseshoe crabs engage in is mating. Every spring the animals migrate from deep water to protected beaches along the eastern coast of North America. Males and females pair off near the time of high tide, and then the females, with males clinging to their backs, build shallow underwater nests near the water's edge. Shoreward migrations are greatest during the nighttime high tides of the full and new moons.

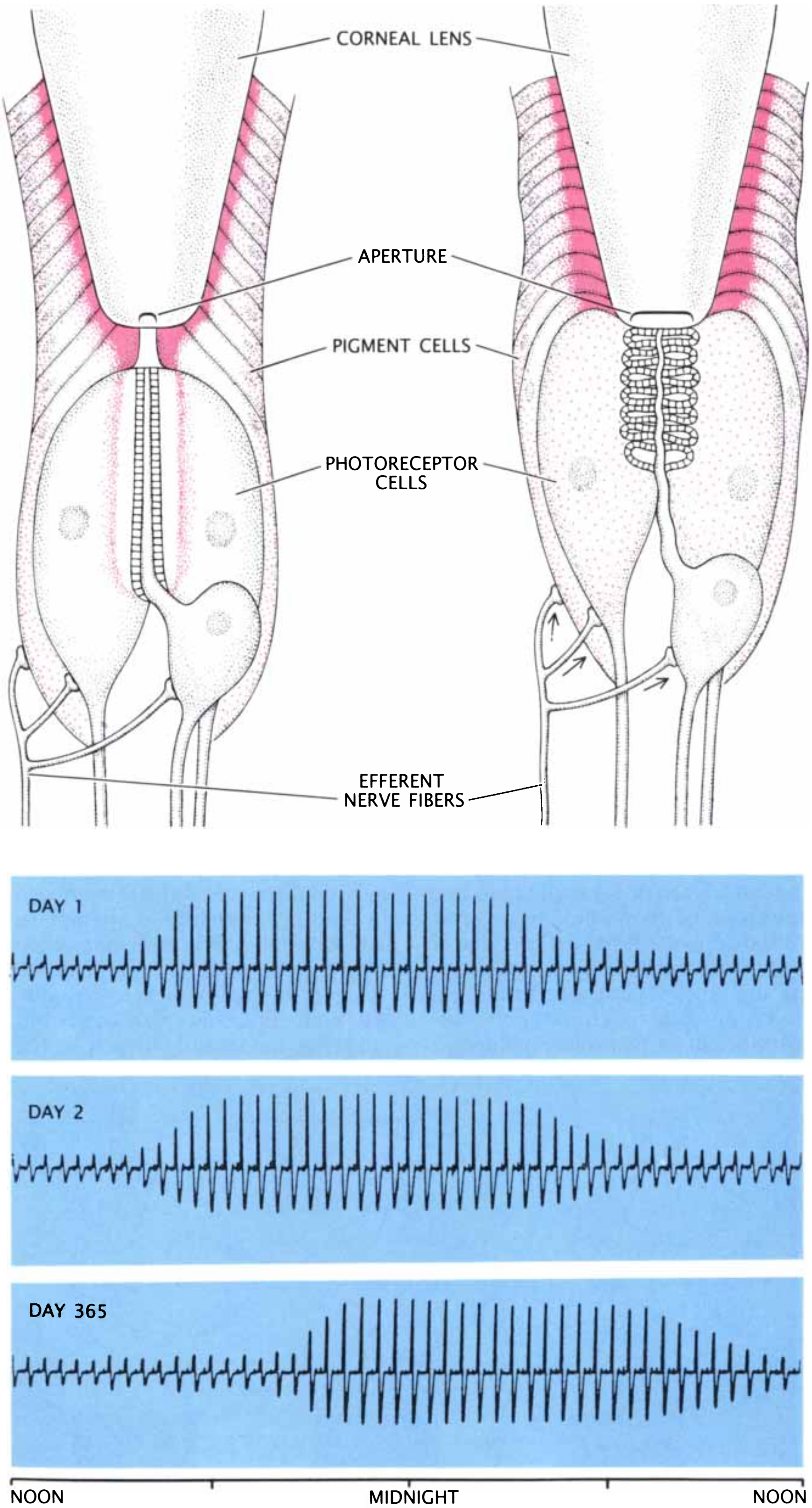
Leonard C. Ireland, then at Woods Hole, Kass and I designed an experiment to test the hypothesis that vision plays some kind of role in horseshoe crab mating. We placed cement castings of female horseshoe crab shells, as well as other shapes, on the sea bottom along the water's edge on a beach near Woods Hole and observed the behavior of the males who swam by. Black castings of female shells were the most attractive; they evoked the entire mating behavior of males: approach, mounting and release of sperm. Castings of other shapes and various shades of gray attracted males but did not elicit mating behavior.

Clearly, chemical cues were not involved: our cement castings did not give them off. Furthermore, male crabs that had been blinded by placing covers over their eyes did not approach any castings, much less attempt to mate with them. It is clear horseshoe crabs use their eyes to find mates.

Indeed, the crabs appear to see quite well. Maureen K. Powers of Vanderbilt University, Kass and I videotaped the



OPTIC NERVE conducts signals from the horseshoe crab's brain to its eye in this experiment performed by the author and his colleagues. The brain of an intact crab (*a*) sends signals to the eye at night that increase its sensitivity to light. Cutting the nerve (*b*) interrupts the flow of signals, and the eye's response falls to its low daytime level. If nighttime signals from the brain are recorded from the optic nerve (*c*) and then played back to the eye (*d*), the eye's sensitivity increases immediately.



LIGHT-SENSITIVE STRUCTURES in the horseshoe crab's eye show striking changes from day (*top left*) to night (*top right*). The aperture that restricts the entry of light into the photoreceptors widens, pigment granules that absorb stray light disperse and the rhabdoms (light-sensitive membranes) fold to present a greater surface for illumination. In addition to these structural changes, the eye undergoes physiological alterations that increase the neural response to each photon of light captured. Electroretinogram recordings (*bottom*) show that the cycle of increased and decreased sensitivity persists even when the animal is kept in complete darkness.

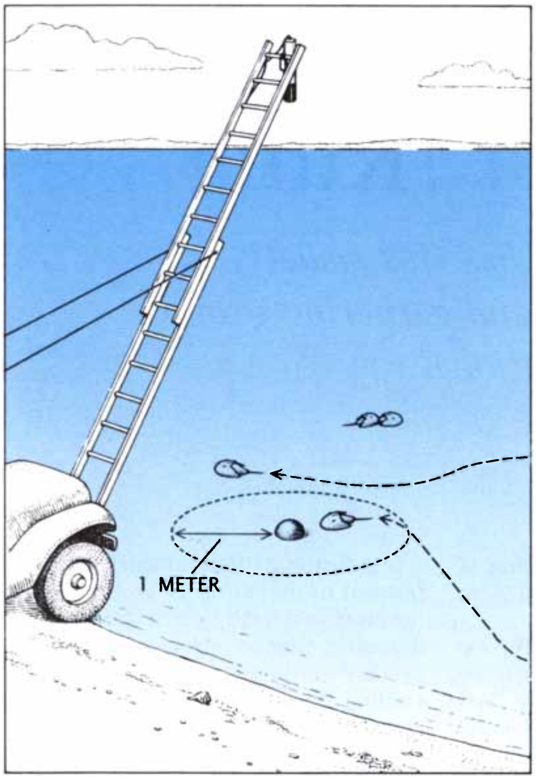
behavior of males around submerged castings of females during the day and at night. (At night we fitted the camera with an image intensifier that amplified ambient light by a factor of 20,000.) The male's sighting distance (the proximity required for a crab to turn toward the casting and approach it) was about one meter during the day and .9 meter at night. The millionfold increase in retinal sensitivity compensated almost completely for the decrease in light intensity after sunset.

At a distance of about one meter a male horseshoe crab's image of the female is constructed from the responses of very few ommatidia, perhaps as few as four. The male may be operating near the optical limits of its compound eye as it searches for females, but nonetheless it succeeds. And although horseshoe crabs sacrifice acuity at night to catch as many photons as possible, apparently they can still see one another.

A male horseshoe crab, then, can turn toward an object whose image in the brain is formed by signals from just a few ommatidia. This raises the question of just what information the eye, modulated as it is by circadian signals, sends to the brain. The horseshoe crab's eye is one of the few networks of neural cells whose connections and behavior have been modeled exactly. The equations that describe its response to static images have been known for about 30 years, those describing the response to moving images for nearly 20. No realistic model exists for any neural network of similar size or complexity.

In theory, a comparison between the response of cells in the *Limulus* eye and in the model to the same image could significantly advance the goal of understanding what information the eye must transmit for the brain to see. In practice, until recently such a computation from the equations for individual ommatidia taxed even the largest and fastest computers. Only with the advent of massively parallel computers has a network containing 1,000 elements and tens of thousands of connections become tractable.

Syracuse graduate students Ramkrishna V. Prakash and Eduardo C. Solessio and I have programmed 1,000 processors in a Connection Machine to model the horseshoe crab retina; we feed simulated images from an underwater camera to the machine and compute the eye's response. Kaplan and I have also implanted electrodes in horseshoe crabs to record the response of individual photoreceptors



VIDEO CAMERA and image intensifier record the response of male horseshoe crabs (*left*) to cement castings designed to resemble the female *Limulus* to a greater or lesser degree. A cast of a female (*top right*) evokes the complete mating response (approach, mounting and release of sperm) when a male horseshoe crab passes within sighting distance. Male crabs approach other objects (*bottom right*) briefly, but they do not remain in their vicinity or attempt to mate with them.

as the animal crawls across the bottom and sees the same image presented to the video camera.

So far the results are encouraging and intriguing. One rather surprising result of the simulations is an understanding of the crucial role of motion in the crab's vision. The eye responds much more strongly to objects that move across its visual field than to stationary ones. Furthermore, the simulation indicates that the eye's combination of spatial and temporal responses may be optimized for detecting objects the size of another *Limulus* moving across the field of view at rates a horseshoe crab would be likely to travel. (To the eye of a moving crab, stationary objects appear to be in motion, so the male detects castings of females as well as it would detect the real thing.)

The complex interaction between the brain and the eye of *Limulus* is only one example of the intricate relations between the brain and the sensory organs of almost all animals. The pioneering neuroanatomist Santiago Ramón y Cajal first uncovered two-way communication between the brain and the eye of a bird in 1889; he found connections between neurons in the upper brain stem and neurons in the retina. In 1971 Frederick A. Miles of the National Institutes of Health showed these connections carry signals that change the way the retina codes spatial information and so should alter the way a bird sees its world.

Similar efferent connections have been found in many other animals, from the nerves that heighten overall sensory response in some fish to those that transmit signals from the brain to the ear in humans and other primates. Efferent neural connections from other parts of the brain outnumber the afferent connections from the optic nerve in the human lateral geniculate nucleus, where the initial stages of visual processing are performed. It appears that the brain, as much as the eye, determines how people see.

People and birds are complicated, however; no one knows exactly how they see, much less how their brains modulate that vision. The work done on simpler neural systems such as *Limulus* may help elucidate such questions in more complicated species. Ultimately a series of ever more complex studies, founded on work on the horseshoe crab, may explain how the incomplete and unstable picture that sensory organs provide, modulated both by the brain and the environment, gives rise to such direct and incontrovertible impressions as the image of a sunset, the smell of a rose or the sound of a Bach fugue.

FURTHER READING

CIRCADIAN RHYTHMS IN *LIMULUS* PHOTORECEPTORS: I. INTRACELLULAR STUDIES. Robert B. Barlow, Jr., Ehud Kaplan, George H. Renninger and Takehiko Saito in *Journal of General Physiology,* Vol. 89, No. 3, pages 353-378; March, 1987.

VISION AND MATING BEHAVIOR IN *LIMULUS.* Robert B. Barlow, Jr., Maureen K. Powers and Leonard Kass in *Sensory Biology of Aquatic Animals.* Edited by Jelle Atema, Richard R. Fay, Arthur N. Popper and William N. Tavolga. Springer-Verlag, 1988.

CENTRIFUGAL PATHWAYS TO THE RETINA: INFLUENCE OF THE OPTIC TECTUM. Hiroyuki Uchiyama in *Visual Neuroscience,* Vol. 3, pages 183-206; 1989.

CIRCADIAN RHYTHMS IN INVERTEBRATE VISION. Robert B. Barlow, Jr., Steven C. Chamberlain and Herman K. Lehman in *Facets of Vision.* Edited by D. G. Stavenga and R. C. Hardie. Springer-Verlag, 1989.

WHAT THE BRAIN TELLS THE EYE (videotape). Robert B. Barlow, Jr. Woods Hole, Mass., Technical Video Ltd., 1989.

The Transformation of the Kalahari !Kung

Why after centuries of stability has this society, an apparent relic of ancient hunting and gathering groups, abandoned many of its traditional ways?

by John E. Yellen

We study history to understand the present. Yet sometimes the present can help to clarify the past. So it is with a San-speaking people known as the !Kung—a group of what were once called African Bushmen. (The exclamation point is pronounced as a click.) Dramatic changes now occurring in the !Kung culture are illuminating a major problem in anthropology: Why did most hunting and gathering societies disappear rapidly after coming in contact with societies that kept domesticated animals and plants?

This swift disappearance is puzzling. After all, hunting for animals and gathering wild plants was a robust enough strategy to ensure the survival of anatomically modern human beings from their emergence more than 50,000 years ago until some time after the first animals and plants were domesticated, roughly 10,000 years ago. Conventional wisdom suggests that many traditional societies, recognizing the nutritional advantages of herding and agriculture, simply abandoned their old practices once they learned about newer subsistence strategies. Yet a number of observations indicate that dissatisfaction with foraging is apparently the wrong explanation in many instances.

Archaeologists have shown, for example, that foraging can actually be more beneficial than herding and farming. Detailed analyses of skeletal remains reveal that in parts of North America a shift to agriculture was in fact detrimental to nutrition, health and longevity for certain groups. Similarly, in modern times it has become clear that when droughts strike southern Africa, groups that rely heavily on hunting and gathering tend to be affected less severely than groups that depend primarily on water-hungry herds and crops.

Moreover, foraging probably is not as taxing and unfruitful as it is stereotypically portrayed. Richard B. Lee, when he was a doctoral student at the University of California at Berkeley in the 1960's, found that the !Kung, who at the time were among the few groups in the world still obtaining most of their food by foraging, did not live on the brink of starvation, even though they inhabited the harsh Kalahari Desert. (The !Kung occupy the northwest corner of Botswana and adjacent areas of Namibia and Angola.) Indeed, they spent only several hours each day seeking food.

What, then, accounts for the decline of foraging societies? No one can say definitely, but glimmers of an answer that may have broad application are emerging from studies focusing on the recent changes in the !Kung way of life. Today young boys no longer learn to hunt, and some of the behavioral codes that gave the society cohesion are eroding. One major catalyst of change appears to have been a sudden easy access to goods. Perhaps a similar phenomenon contributed to the demise of past foraging societies.

It is fortunate that a rather detailed portrait of the !Kung's traditional culture was compiled before the onset of dramatic change. Many investigators deserve credit for what is known, including the independent anthropologist Lorna Marshall, who began studying the group in 1951, and Irven DeVore, Lee and other participants in what was called the Harvard Kalahari Project. One aim of the project, which officially ran from the late 1960's into the 1970's (and in which I participated as a doctoral student), was to understand how traditional hunting and gathering societies functioned.

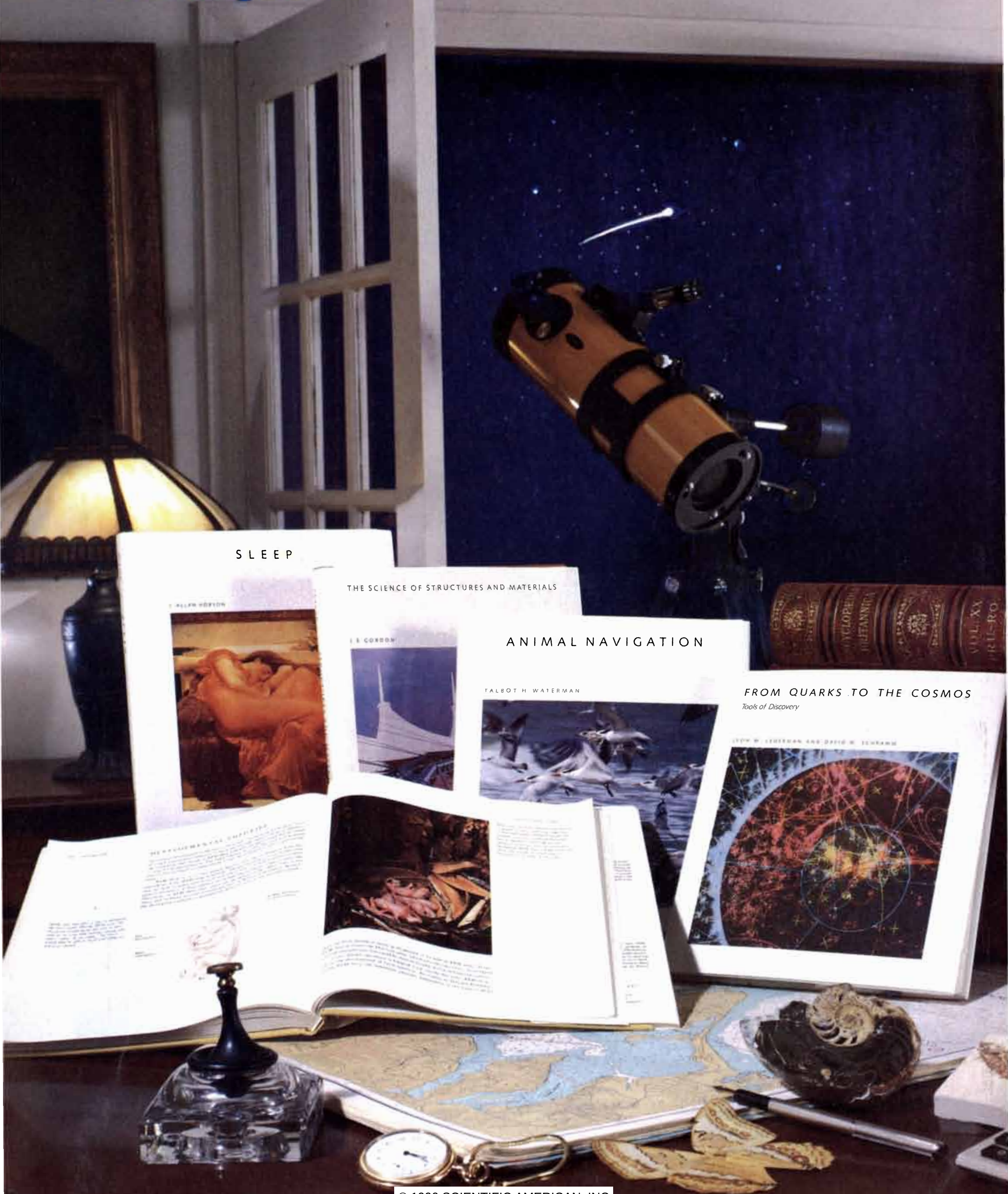
Any description of the !Kung begins most appropriately with a brief history of the peoples in southern Africa. Before the start of the first millennium A.D., Africa south of the Zambezi River was still populated exclusively by foragers who almost certainly were of short stature, had light-brown skin and spoke what are called Khoisan languages (all of which, like those in the San group, include clicks). In the still more distant past, the various groups had apparently shared a common language and culture and then, as they spread out, adapted to the specific conditions of the regions where they settled. Some had adjusted to the seasonal cold of the Drakensberg Mountains, others to the coastal areas (with their wealth of fish), and still others to the drier conditions of the deserts and other inland areas.

The various groups were what archaeologists call late Stone Age peoples; their knife blades and scraping tools were made of stone and specialized for particular tasks. As yet there were none of the hallmarks of so-called Iron Age peoples: domesticated goats, sheep and cattle; cultivated

JOHN E. YELLEN, who received a doctorate in anthropology from Harvard University in 1974, is program director for archaeology at the National Science Foundation. Before joining the foundation in 1977, he was a research associate at the National Museum of Natural History of the Smithsonian Institution. Yellen conducted ethnographic and archaeological studies of the !Kung for some 15 years, often in collaboration with his wife, Alison S. Brooks of George Washington University. Today he and his colleagues are excavating sites in Zaire for clues to the emergence of complex hunting and gathering cultures.

!KUNG MEN traditionally secured food by hunting (*top*). Today they also keep herds of domesticated animals (*bottom*), and boys are no longer taught to hunt. These changes in behavior are just some of the many that have become strikingly apparent in the past 20 years. The !Kung, who were once highly mobile and could carry all their possessions on their backs, now have trunks filled with goods and are much more sedentary, less interdependent and more enmeshed with outsiders.

Begin a voyage of discovery that takes you from quarks to the cosmos, and beyond.



Join the Scientific American Library now by examining *From Quarks to the Cosmos* for 15 days risk-free.

You are invited to embark upon a journey to the newest frontiers of science. From the mysteries of the brain . . . to the most spectacular phenomena in the universe.

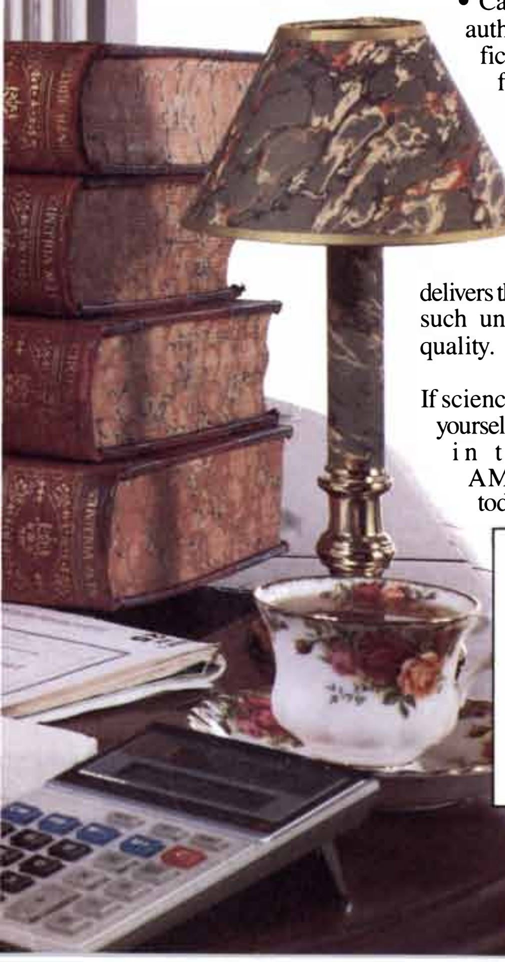
Join us and you will explore the most exciting discoveries on the scientific horizon today as a member of the SCIENTIFIC AMERICAN LIBRARY.

From Molecules, *a polarized light view of light heating oil; courtesy Manfred Kage/Peter Arnold, Inc.*

Written by leading scientists in accessible, compelling prose.

SCIENTIFIC AMERICAN LIBRARY books are created in a unique collaboration between our editors and some of the world's most renowned scientists (including many Nobel laureates). The result is a series of books that are readable *and* authoritative. Each volume . . .

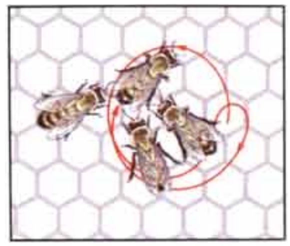
- Covers a topic of major interest with immediate relevance to ongoing research.
- Carries the authority of an author who has made significant contributions to the field.
- Is written in clear, compelling prose that's uniquely accessible even to the non-expert.



No other book club delivers the literature of science with such uncompromising editorial quality.

If science is your passion, reward yourself. Begin your membership in the SCIENTIFIC AMERICAN LIBRARY today.

Beautiful, richly illustrated volumes that are made to endure.



From The Honey Bee, *foragers dance to indicate a food source; drawing by John Hatzakis, adapted from K. von Frisch,* Tanzsprache und Orientierung der Bienen.

From the moment you receive your first book, you'll see that SCIENTIFIC AMERICAN LIBRARY volumes are produced to exacting standards. Each work:

- Features abundant photographs, paintings and drawings in full, vibrant color.
- Includes many helpful diagrams, charts, and maps to graphically illustrate the text.
- Is impeccably printed on high quality paper and clothbound for uncommon durability.

The books in the SCIENTIFIC AMERICAN LIBRARY are fascinating reading *and* they make a handsome addition to your library.

Join now by sending for your premier volume without risk or obligation.

Join the SCIENTIFIC AMERICAN LIBRARY now by previewing *FROM QUARKS TO THE COSMOS,* written by 1988 Nobel Prize-winning physicist Leon Lederman and cosmologist David Schramm. This intriguing book explores the revolutionary discoveries made through the powerful merger of particle physics and cosmology.

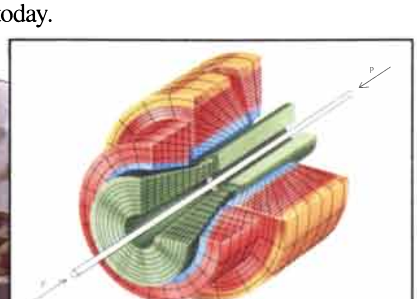
From Eye, Brain, and Vision, *the neurobiology group at Harvard Medical School, 1963, including Nobel Prize winners David Hubel, standing right, and Torsten Wiesel, seated right. Photograph by Joseph Gagliardi.*

This is just the first of many fascinating volumes you'll receive. In months to come you'll discover books covering a wide range of subjects, including *Molecules, Sexual Selection,* and *Eye, Brain, and Vision.*

As a member, you'll enjoy:

- A NEW BOOK EVERY TWO MONTHS or so, shipped direct to your doorstep.
- 15-DAY FREE EXAMINATION. Review each book without risk. You may keep it or return it for full credit.
- A LOW, MEMBERS-ONLY PRICE. Pay just $24.95 (25% off the retail price) for each book you keep.
- NO PREPAYMENT OR MINIMUM PURCHASE. Pay only for the books you choose to keep.
- OUR GUARANTEE that you'll be fully satisfied or *you may cancel at any time without penalty.*

TO JOIN, WRITE TO SCIENTIFIC AMERICAN LIBRARY, P.O. BOX 646, HOLMES, PA 19043, OR CALL TOLL-FREE 1-800-345-8112.



◄ *From* From Quarks to the Cosmos, *a schematic of a particle collider detector; drawing by George Kelvin.*

SCIENTIFIC AMERICAN LIBRARY

AN EXCITING NEW GENRE IN THE LITERATURE OF SCIENCE grains such as millet and sorghum; pottery; and smelted and forged iron and copper.

The first Iron Age influences appeared in southern Africa some time early in the first millennium A.D., when, according to the archaeological record, occasional goods and domesticated animals were introduced, presumably by trade with peoples in more northern territories. The items were soon followed by Iron Age settlers themselves. These newcomers from the north spoke mostly Bantu languages and, compared with the foragers, were taller and darker-skinned. Either directly or indirectly, all of the foraging groups were eventually exposed to the new settlers and technologies and, later, to waves of European intruders: the Dutch and the Portuguese beginning in the 15th and 16th centuries and then the English and Germans as well.

Artifacts as well as journals of European settlers indicate that some of the hunting and gathering groups were exterminated by the intruders. In most other instances, according to clues provided by genetic studies, linguistic analyses and other methods, groups broke up (forcibly or otherwise), often merging with their new neighbors through intermarriage.

In certain cases, foragers were able to maintain a distinct genetic and cultural identity. Some of them, changing many practices, became transformed into new cultures. (For example, the first Dutch settlers, arriving at the southern tip of Africa, met "Hottentots," Khoisan speakers who herded flocks of sheep, goats and cows.) In the Karroo Desert of South Africa and in the northern Kalahari, however, a few hunting and gathering societies—among them, the !Kung—not only stayed intact but also apparently held onto many of their old ways.

Indeed, even as late as 1968, when I first visited the Kalahari as part of the Harvard project, most !Kung men and women in the Dobe region of Botswana still dressed in animal skins and subsisted primarily by hunting and gathering. (Dobe is the site most intensively studied by the project; the people there are, by all indications, quite representative of the !Kung over a broader area.) It is true that iron had long since replaced stone in tools, and plastic and metal containers had supplanted their ceramic counterparts. Yet men still hunted with bows and poisoned arrows, and women set out daily with digging sticks to seek edible plants.

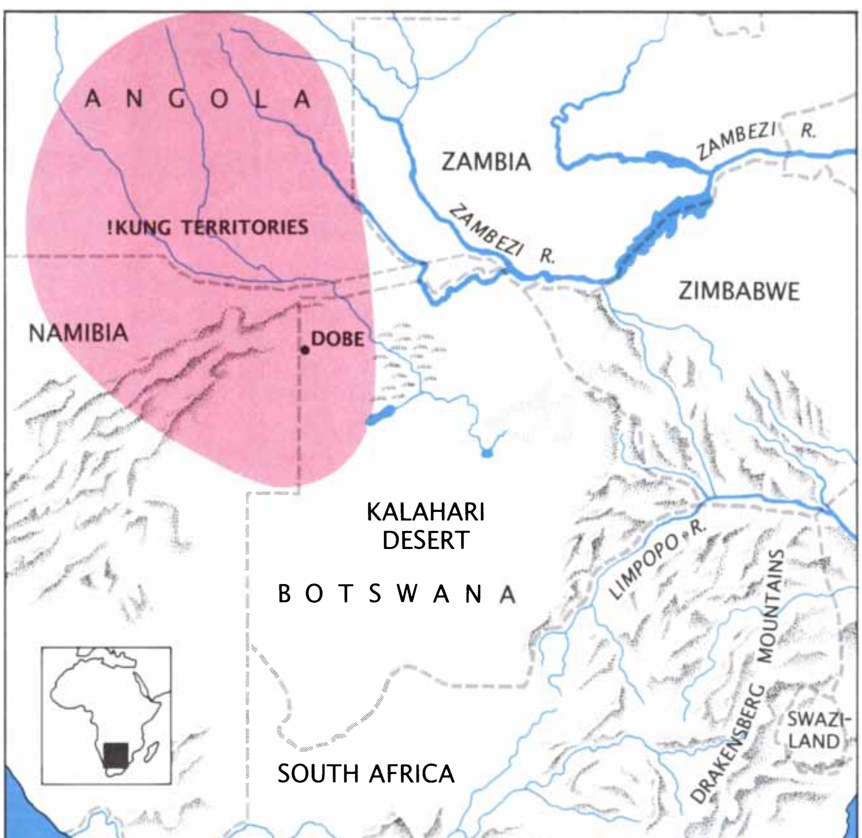
At least it seemed to us that the people we met were behaving much as their ancient ancestors had. Some scholars dissent from that view, contending that the forerunners of 20th-century foragers were probably altered radically by contact with Iron Age peoples. If so, they say, modern foragers, including the !Kung, may reflect but little of the past.

In my view, strong evidence suggests that the !Kung studied in the early years of the Harvard project were very much like their distant forebearers. For example, I have determined that the range of stone tools excavated from what is now !Kung territory remained remarkably constant into the late 19th century (when the grandparents of modern !Kung adults would have been born). This finding means that the region was probably populated continuously by one cultural group and that its foraging and manufacturing practices remained essentially unaffected by Iron Age influences.

What were the traditional ways of the !Kung? Observations made back in the 1950's and 1960's reveal that the group's strategy for obtaining food—and in fact its entire social organization—was exquisitely adapted for survival in the Kalahari. There, rainfall can vary dramatically from year to year and region to region, giving rise to profound shifts in the availability of food.

When it came to securing food, the !Kung followed what I call a generalist strategy. Rather than specializing in the pursuit of a limited number of species, as could be done in more predictable environments, they cast their foraging "net" broadly and so could usually find something to eat even if favored foods were in short supply. Remarkably, Lee found that males hunted more than 60 animal species, ranging in size from hare to buffalo. Females recognized more than 100 edible plant species, collecting perhaps a dozen varieties in a single day.

Certain accepted foraging guidelines minimized competition for the desert's limited resources. For example, groups of people were loosely



SOUTHERN AFRICA is home to many indigenous groups of San speakers (formerly known as Bushmen), including some who lived essentially as hunters and gatherers, or foragers, well into the 20th century. The !Kung, perhaps the best studied of the San, occupy the Kalahari Desert in parts of Botswana, Namibia and Angola. Much of what is known of the group has been gleaned from anthropological and archaeological studies conducted by a number of investigators in the Dobe region of Botswana.

organized into bands, and each band had the right to seek food in specified areas. During the dry season the members of a single band would congregate, setting up camps near a water hole (a year-round source of drinking water) understood to belong to that band. From the camps, individuals or small clusters of people would fan out each day to forage. During the rest of the year, when rainfall was more frequent and rain collected in shallow depressions in the ground known as pans, bands would disperse; small groups foraged in less trafficked areas, staying for as short as a day (and rarely as long as two months) before moving on.

The band system actually made it easy for people to migrate to more desirable places when the territory allotted to a given band was unproductive or becoming depleted. Band membership was rather fluid, and so a family could readily join a different band having more luck.

Consider the options open to a husband and wife, who would have had few possessions to hamper their travels. They could claim the right to join the bands available to both sets of parents, which means that at least four territories were open to them. Moreover, they could join any band in which their brothers or sisters had rights. If the couple also had married children, they might, alternatively, forage anywhere the children's spouses could; indeed, parents frequently arranged their children's marriages with an eye to the accompanying territorial privileges. Individuals could also claim band memberships on the basis of certain less direct kinship ties and on friendship.

The social values of the !Kung complemented this flexible band system, helping to ensure that food was equitably distributed. Most notably, an ethic of sharing formed the core of the self-described !Kung system of values. Families were expected to welcome relatives who showed up at their camps. Moreover, etiquette dictated that meat from large kills be shared outside the immediate family, which was obviously a sound survival strategy: a hunter who killed a large antelope or the like would be hard pressed, even with the help of his wife and children, to eat all its meat. By distributing his bounty, the hunter ensured that the recipients of his largess would be obliged to return the favor some time in the future.

Similarly, individuals also established formal relationships with nonrelatives in which two people gave

TYPICAL RAINY-SEASON CAMP of the past was often maintained for just a day or two before its inhabitants moved to a new location. When people were in camp, they congregated around cooking hearths, which were placed just outside lean-to-like huts. Dry-season camps were similar in form but were occupied for weeks or months. Today many families no longer leave "dry-season" camps during the rainy season.

HUNTER digging for tubers exemplifies the !Kung's traditional, generalist approach to obtaining food. The !Kung coped with the Kalahari's unpredictable food supply by being flexible. If one food was unavailable, they accepted a range of substitutes.

TIME PERIOD	LARGE MAMMALS GREATER KUDU/CATTLE	MEDIUM MAMMALS STEENBOK/GOAT	SMALL MAMMALS SPRINGHARE/ PORCUPINE	REPTILES AND AMPHIBIANS PUFF ADDER/BULLFROG	BIRDS GUINEA FOWL/ CHICKEN
1944–1962	1.40	2.40	1.80	0.80	1.20
1963–1968	2.86	2.86	2.14	1.71	1.86
1970–1971	2.33	3.33	2.33	1.67	1.33
1972–1975	2.00	2.75	2.00	1.25	2.25

AVERAGE NUMBER OF SPECIES

!KUNG DIET remained varied between 1944 and 1975, according to an analysis of animal bones excavated from dry-season camps at Dobe. The author identified and counted the number of bones at each camp to learn the relative numbers of the large, medium and small mammalian species and the reptilian, amphibian and bird species consumed at each camp during four periods. (Selected examples are shown.) The balance across categories changed little, indicating that variety was maintained, as was the !Kung's generalist food-securing strategy. The persistence of a diverse diet even after domesticated animals were acquired in the 1970's indicates that the group had not become dependent on their herds, which apparently were viewed as foraging resources like any others. Hence, the popular notion that dissatisfaction with foraging caused hunting and gathering societies of the past to abandon their old way of life does not seem to hold for the Dobe !Kung.

each other gifts such as knives or iron spears at irregular intervals. Reciprocity was delayed, so that one partner would always be in debt to the other. Pauline Weissner, when she was a graduate student at the University of Michigan at Ann Arbor, analyzed those reciprocity relationships and concluded that individuals purposely selected gift-giving partners from distant territories. Presumably it was hoped that a partner would have something to offer when goods were difficult to obtain locally. Hence, in the traditional !Kung view of the world, security was obtained by giving rather than hoarding, that is, by accumulating obligations that could be claimed in times of need.

Clearly, mobility was a critical prerequisite for maintaining reciprocity relationships over long distances and for making it possible to move elsewhere when foraging conditions were unfavorable. The !Kung system of justice had the same requirement of ready movement. Like many other traditional foraging groups, the !Kung society was acephalous, or headless: no one was in charge of adjudicating disputes. When disagreements became serious, individuals or groups of disputants simply put distance between themselves, claiming membership in widely separated bands. As long as everyone could carry their few possessions on their backs, and so could relocate with ease, the approach worked well.

The traditional !Kung, then, were well suited to the Kalahari. They were generalists who lived by the ethic of sharing, ensuring that those who were less successful at finding food could usually be fed nonetheless. Because families owned no more than they could carry, they were able to travel at will whenever resources became scarce or disputes too heated.

By 1975, however, the !Kung were undergoing a cultural transition—at least so it seemed by all appearances at Dobe. I left there in 1970 and returned in the middle of the decade. I found that, in the interim, many families had taken on the ways of the neighboring Bantu. A number had planted fields and acquired herds of goats along with an occasional cow. Fewer of the boys were learning to hunt; traditional bows and arrows were still produced but mostly for eventual sale on a worldwide curio market. The people wore mass-produced clothing instead of animal skins, and traditional grass huts were for the most part replaced by more substantial mud-walled structures, which were now inhabited for longer periods than in the past.

An influx of money and supplies had clearly played a part in many of these changes. Botswana became an independent nation in 1966, after having been the British protectorate Bechuanaland. The new government began to encourage the keeping of livestock and the development of agriculture, such as by giving donkeys to the !Kung for pulling simple plows. And it arranged for the routine purchase of traditional handicrafts (for example, bead necklaces), thereby injecting extraordinary sums of money into the community. Later, when the !Kung in Namibia (then a colony of South Africa) were brought into the South African Army, the !Kung in Botswana received more infusions of cash and goods, mainly via interactions with kin.

Yet the exact meaning of such surface changes remained unclear. To what extent did the livestock and fields, the new clothes and the sturdier huts reflect a weakness in the glue that held !Kung society together? Why had the men and women, who had long been successful as foragers and who were not coerced into changing, decided to take on the burdens of herds and crops and to otherwise allow their mobility to be compromised? Archaeological work I undertook at Dobe between 1975 and 1982 (first as a research associate at the Smithsonian Institution and then as an employee of the National Science Foundation), together with observations made by other workers during the same period, provides some hints.

To be frank, when I returned in 1975, a methodological question preoccupied me. I hoped to learn about what happened to the bones of hunted

Introducing...

CARE *of the* SURGICAL PATIENT

from SCIENTIFIC AMERICAN *Medicine*

Because the quality of your care depends on the quality of your information.

Treating pre and post operative patients poses a unique set of challenges. Yet in one way it's no different than any other practice issue.

Doing it well takes the right information.

That's why SCIENTIFIC AMERICAN *Medicine* is pleased to announce the publication of *CARE of the SURGICAL PATIENT.*

The definitive resource on pre and post-operative care.

CARE of the SURGICAL PATIENT gives you ready access to the most authoritative and current information on pre and post-operative standards available anywhere.

Written and designed by prominent surgeons under the supervision of the American College of Surgeons' Committee on Pre and Postoperative Care, *CARE of the SURGICAL PATIENT* provides two volumes – over 1,500 pages – of practical information on both critical and elective care.

And, *CARE of the SURGICAL PATIENT* is updated twice a year, with each surgeon-author reviewing his own specialty. Updates include new information on significant topics, such as current developments on AIDS.

In short, *CARE of the SURGICAL PATIENT* presents the standards for pre and postoperative treatment. You simply won't find a more important resource. Or one that organizes its information in such an intelligent way.

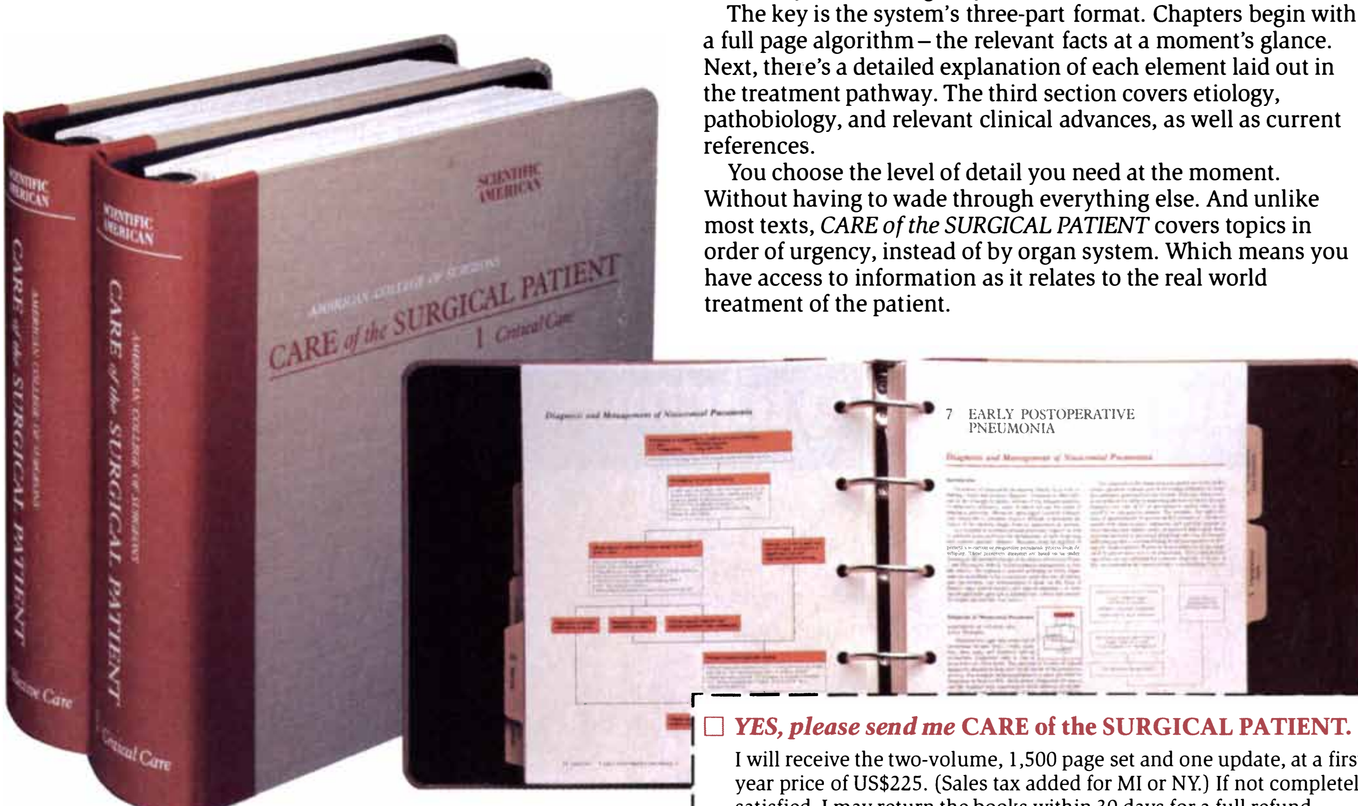
A unique system for rapid information retrieval.

CARE of the SURGICAL PATIENT is designed to get you the information you need, the way you need it.

Quickly. And intelligently.

The key is the system's three-part format. Chapters begin with a full page algorithm – the relevant facts at a moment's glance. Next, there's a detailed explanation of each element laid out in the treatment pathway. The third section covers etiology, pathobiology, and relevant clinical advances, as well as current references.

You choose the level of detail you need at the moment. Without having to wade through everything else. And unlike most texts, *CARE of the SURGICAL PATIENT* covers topics in order of urgency, instead of by organ system. Which means you have access to information as it relates to the real world treatment of the patient.



Try CARE of the SURGICAL PATIENT *Free for 30 days.*

You'll find it the most valuable resource on pre and post-operative care that's ever been published. And if you're not satisfied, just return it. No risk. No obligation.

CARE of the SURGICAL PATIENT, from SCIENTIFIC AMERICAN *Medicine*. No other resource helps you keep up better.

And the better you keep up, the better your care.

☐ ***YES, please send me*** **CARE of the SURGICAL PATIENT.**

I will receive the two-volume, 1,500 page set and one update, at a first-year price of US$225. (Sales tax added for MI or NY.) If not completely satisfied, I may return the books within 30 days for a full refund.

☐ Check enclosed ☐ MasterCard ☐ VISA ☐ Bill me

Acct.# ______________________ Exp. Date ________

Name ______________________

Address ______________________

____

City ____________ State ________ Zip ________

Signature ______________________ Specialty ________

Or call toll free 1-800 345 8112.

SCIENTIFIC AMERICAN MEDICINE 415 Madison Avenue, New York, NY 10017 7530 animals after the carcasses were discarded and became buried naturally in the ground; such information was important for developing archaeological techniques to determine how people in the past killed, butchered and cooked animals. I thought that by locating the remains of old cooking hearths at Dobe, around which families ate, I might gather a good collection of bones—the remains of meals—on which to test a few ideas. Later I realized the data I had collected in the course of this endeavor might also say something about the transformation of the !Kung.

As part of my studies I identified and mapped the locations of huts and their associated hearths dating back to 1944. I then dug up bones that had been dropped in and around the hearths and identified the species to which they belonged. In visits made after 1975, I no longer collected bones, but I continued to map contemporary camps; in the end, I accumulated almost 40 years of settlement data.

The camps were usually occupied by the same extended family and close relatives, such as in-laws, although the specific mix of individuals changed somewhat from year to year. At the older sites, where all visible traces of occupation had disappeared, the huts and the hearths (which were normally placed outside a hut's entryway) were identified with the help of family members who actually remembered the placements.

My data supported the conclusion that by the mid-1970's long-standing !Kung values, such as the emphasis on intimacy and interdependence, were no longer guiding behavior as effectively as they once did. The data also indicated that, despite appearances to the contrary, the !Kung had retained their foraging "mentality." These generalists had taken up herding as if their goats and few cows were no different from any other readily accessible foraging resource. This surprising discovery meant that factors other than a failure of the food-securing system were at the root of the !Kung transformation.

My sense that traditional values were losing their influence over behavior came mainly from my analyses of the maps I had drawn (combined with other observations). Traditional !Kung camps, as depicted in the first 25 or so maps, were typically arranged in a circle, and most entrances faced inward. The huts were also set close together, so that from the entrance of one of them

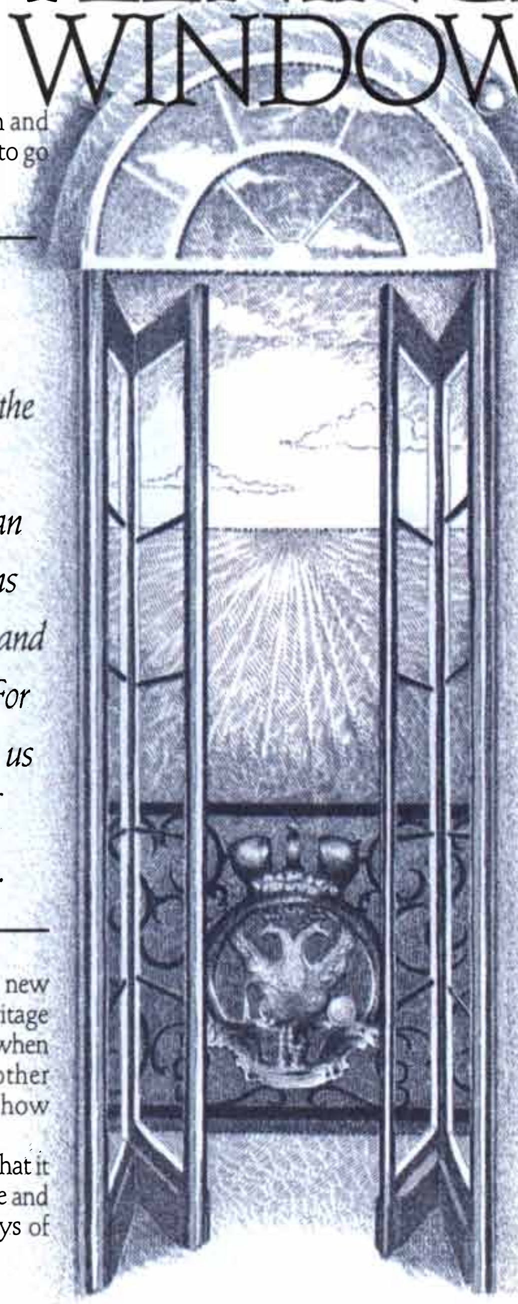
GLASNOST REOPENS A "WINDOW TO THE WEST."

CAN WE SEE OUR OWN HORIZONS FROM LENINGRAD'S WINDOW?

NEWBORN citizens, at a ceremony in the Baby Palace of Leningrad, are awarded a medal that bears the words "Born in Leningrad."

To begin life in Leningrad at this moment in history holds the potential of witnessing changes as momentous as any in this great city's eventful past.

With glasnost, today's Leningrad has the sound of rock concerts, allows the revival of religion and creates co-ops that permit people to go into business for themselves.



Leningrad, a city of golden domes and glittering canals, opens its windows further to the Western world as a result of glasnost. *Glasnost is Russian for "openness"—which means acceptance of ideas, cultures and the free flow of information. For Americans, glasnost reminds us of the openness that is one of our country's basic principles.*

For Americans, openness is not a new idea. It has been a part of our heritage since our nation's beginnings. And when we witness its appearance in other countries, we are reminded of how much it has made us what we are.

The people of Chevron believe that it is important to our nation to educate and inform, share information and ways of life, and bring the openness of ideas to a watching world. That's why we underwrite the National Geographic Television Specials. Chevron is proud to salute the National Geographic Society on its 25th anniversary of producing award-winning television.

To visit one of the most beautiful cities in the Soviet Union, explore its turbulent past, see the splendor of its art and get a glimpse of its future through its resilient people, watch and listen to "Voices of Leningrad," an all-new National Geographic Special on Public Television.

It's produced by the National Geographic Society and WQED/Pittsburgh and underwritten by the people of Chevron.

Tune in Wednesday, April 4th. Check your local listings for show times on PBS.

Do people believe that openness can make a world of difference?

People Do.

Chevron

SCIENTIFIC AMERICAN

CORRESPONDENCE

Offprints of more than 1,000 selected articles from earlier issues of this magazine, listed in an annual catalogue, are available at $1.25 each. Correspondence, orders and requests for the catalogue should be addressed to W. H. Freeman and Company, 4419 West 1980 South, Salt Lake City, Utah 84104. Offprints adopted for classroom use may be ordered direct or through a college bookstore. Sets of 10 or more Offprints are collated by the publisher and are delivered as sets to bookstores.

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 flat fee of $1.50 per copy of each article or any part thereof. 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., 21 Congress Street, Salem, Mass. 01970. Specify CCC Reference Number ISSN 0036-8733/89. $1.50 + 0.00.

Editorial correspondence should be addressed to The Editors, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. 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, N.Y. 10017.

Address subscription correspondence to Subscription Manager, SCIENTIFIC AMERICAN, P.O. Box 3187, Harlan, IA. 51593. Telephone inquiries: 1-800-333-1199, U.S. only; other 515-247-7631/32. The date of the last issue on subscriptions appears on each month's mailing label. For change of address notify us at least four weeks in advance. Please send your old address (if convenient, a mailing label of a recent issue) as well as the new one.

Name ____________________

New Address ____________________

Street ____________________

City ____________________

State and ZIP ____________________

Old Address ____________________

Street ____________________

City ____________________

State and ZIP ____________________

it was possible to see into most of the others.

The camp arrangement remained close and intimate until the early 1970's. Then suddenly the distance between huts increased significantly. At the same time, the circular pattern yielded to linear and other arrangements that gave families more privacy; also, in the last two camps I mapped (dating to 1981 and 1982), many of the hearths, which had been central to much social interaction, were located inside the huts instead of in front of them. The changes occurred so abruptly that the pattern of camp design can be said to have been unambiguously transformed from "close" to "distant" within a few years. By implication, such changes in camp design indicate that major changes in social norms for openness and sharing occurred as well in the early to middle 1970's.

This conclusion is consistent with other evidence. In 1976 Diane E. Gelburd, then a graduate student at George Washington University, inventoried the material possessions of individuals at Dobe and compared her data with a survey Lee had conducted in 1963. Whereas Lee found that most people could carry all their worldly belongings with ease, Gelburd found a dramatically different situation.

She showed that many !Kung owned large items, such as plows and cast iron pots, which are difficult to transport. With their newfound cash they had also purchased such goods as glass beads, clothing and extra blankets, which they hoarded in metal trunks (often locked) in their huts. Many times the items far exceeded the needs of an individual family and could best be viewed as a form of savings or investment. In other words, the !Kung were behaving in ways that were clearly antithetical to the traditional sharing system.

Yet the people still spoke of the need to share and were embarrassed to open their trunks for Gelburd. Clearly, their stated values no longer directed their activity. Although spoken beliefs and observed behavior do not coincide perfectly in any society, at Dobe in 1976 the disjunction had become extreme.

In what way did my other data set—the animal bones—clarify the causes of the social changes apparent by the 1970's? The presence of domesticated animals and cultivated fields at Dobe caused me to wonder if the changes I saw in the !Kung could be traced through some sequence of events to discontent with foraging. If the bones revealed that by the mid-1970's the !Kung derived meat almost exclusively from domesticated animals, the conclusion could then be entertained that a shift in subsistence

MARRIED COUPLE, at the request of an anthropologist at Dobe, emptied a trunk to reveal the riches they had amassed by 1976. The acquisition of goods may have catalyzed the !Kung transformation. The author suggests that once the people began to think of luxury items as "money in the bank" and to hoard them (which was antithetical to their past ethic of sharing), they lost much of their mobility. Disputes could no longer be resolved in the old way, by putting distance between the disputants; then many individuals turned to neighboring Bantu-speaking societies for arbitration. As involvement with the Bantu increased, so too did cultural change.

strategy had preceded other dramatic social changes and, hence, might have somehow given rise to them.

My data confirmed earlier impressions that through the 1950's the !Kung were almost exclusively hunters and gatherers: in sites dating from that period, the bones of domesticated animals are rare. Then, in the 1960's, the consumption of goat and cattle increased markedly; in fact, by 1974 and 1975 these animals were consumed more than any others. The frequency with which chicken was consumed also increased during that period, although the Dobe !Kung never did eat very much of this Western staple. At the same time—from 1944 to 1975—the once great popularity of certain wild animals waned, including the greater kudu (a large antelope regularly hunted in the dry season) and two smaller antelopes (the steenbok and duiker).

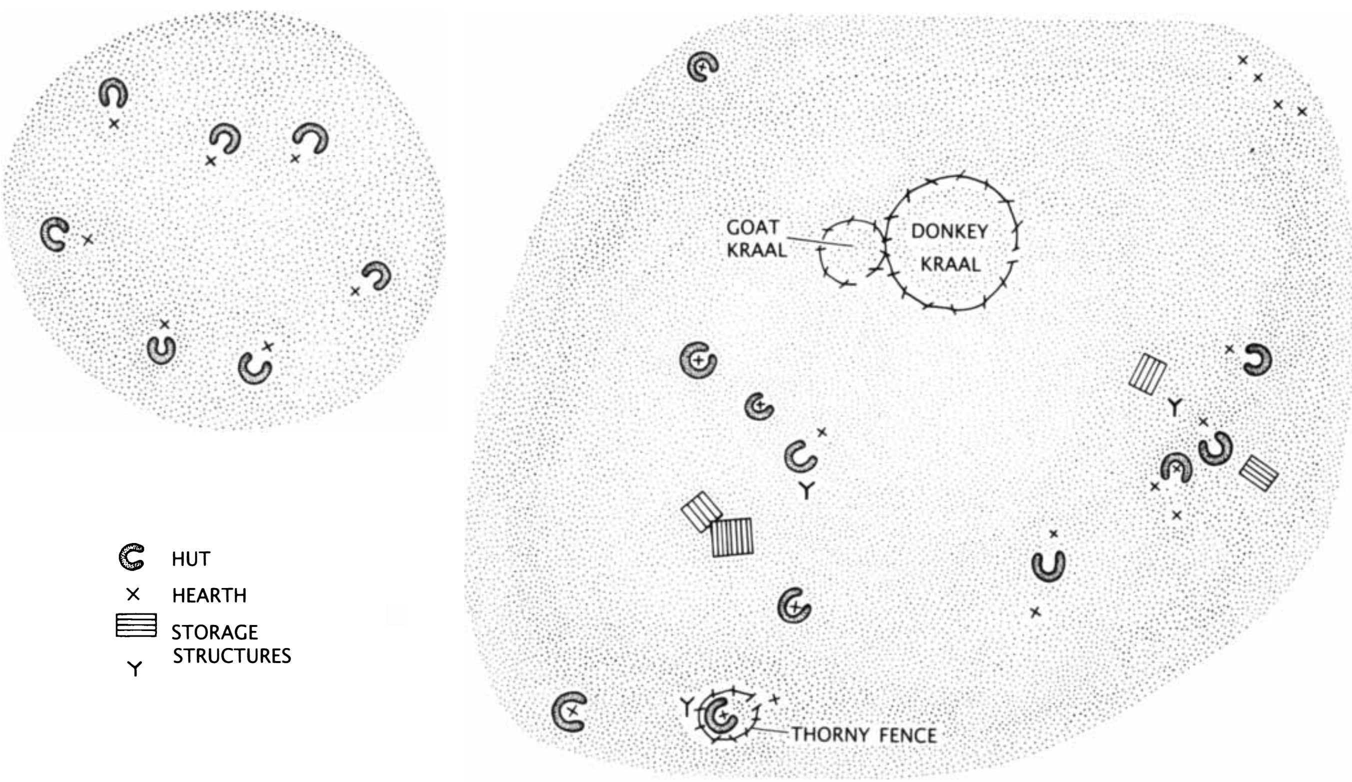
A cursory look at these data might have suggested that the !Kung were indeed abandoning hunting. Yet a closer examination revealed that cattle essentially substituted for kudu, both of which are large animals, and that goats, which approximate steenbok and duiker in size, directly replaced those animals in the diet. It also turns out that the number of species represented at each camp remained essentially the same, as did the mix of small, medium and large species. That is, if the meat diet of the !Kung in the 1940's normally consisted of 10 species, of which 50 percent were small, 30 percent medium and 20 percent large, roughly the same numbers would be found in a 1975 camp, although the species in each category might differ.

These findings show that the !Kung did not reduce the variety in their diet, as would be expected if they had abandoned the traditional, generalist strategy and had committed themselves to becoming herders, who typically are dependent on just a few animal species. Hence, I realized that although anthropologists might view "wild" and "domestic" animals as fundamentally different, the !Kung as late as 1975 did not make such a distinction. From the !Kung perspective, goats were essentially the same as any other medium-size animal (in that they provided a reasonable yield of meat and were relatively easy to carry), and cows were the same as other large creatures. If an animal was easy to obtain, the !Kung ate it, but they apparently did not come to depend on their herd animals to the exclusion of all others.

Anecdotal information supports the assessment that the !Kung of 1975 did not view themselves as herders. For instance, whereas Bantu groups, who depend on their herds for food and prestige, would quickly kill a hyena that preyed on their animals, many !Kung men would not bother to do so. I believe the !Kung would have been less indifferent if they had considered their herds to be all-important sources of meat. Similarly, they seemed to conceive of agriculture and wage labor undertaken for the Bantu and anthropologists—activities they pursued on a part-time, short-term basis—much as they perceived herding: as foraging resources just like any other.

Thus, well into the 1970's, the !Kung retained their generalist strategy, limiting their reliance on any one type of resource. Obviously that approach was adaptable enough to permit the transition from a foraging to a more mixed economy without disrupting social functioning.

If neither empty bellies nor coercion initiated the !Kung's transformation,



ARRANGEMENTS of camps changed markedly between 1944 and 1982. The changes, revealed by a series of maps much like these of dry-season camps, seem to reflect a decline in the cohesion of !Kung society. Until the early 1970's the traditional !Kung camp (*left*) was intimate: closely spaced huts roughly described a circle, and the entryways faced inward so that from a single vantage one could see into many huts. Then the arrangements changed abruptly (*center*): the average distance between huts increased, and the circular arrangement yielded to linear and other private arrays. The dwellings—which in the past were made of branches and grass (*a*) and now resembled the semipermanent mud-walled huts of the Bantu (*b*)—were

what did? The impetus may well have come largely from internal stresses generated by the desire to have the material goods that had become readily accessible. The following scenario—based in part on my map data, Gelburd's work and my interactions with the !Kung over the years—is one plausible sequence of events that may have occurred. The scenario does not attempt to be a comprehensive description of how and why the !Kung culture has changed, but it does describe some of the major processes that seem to be driving the society's transformation.

Once the !Kung had ready access to wealth, they chose to acquire objects that had never before been available to them. Soon they started hoarding instead of depending on others to give them gifts, and they retreated from their past interdependence. At the same time, perhaps in part because they were ashamed of not sharing, they sought privacy. Where once social norms called for intimacy, now there was a disjunction between word and action. Huts faced away from one another and were separated, and some hearths were moved inside, making the whole range of social activities that had occurred around them more private. As the old rules began to lose their relevance, boys became less interested in living as their fathers had. They no longer wished to hunt and so no longer tried to learn the traditional skills; instead they preferred the easier task of herding.

Meanwhile the acquisition of goods limited mobility, a change that came to be reflected in the erection of semipermanent mud-walled huts. The lack of mobility fueled still more change, in part because the people could no longer resolve serious arguments in the traditional manner, by joining relatives elsewhere in !Kung territories.

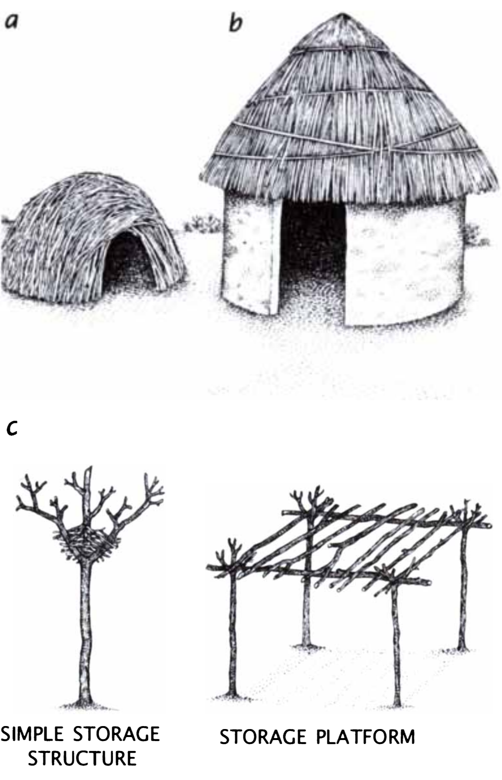
With the traditional means of settling disputes now gone, the !Kung turned to local Bantu chiefs for arbitration. In the process they sacrificed autonomy and, like other San groups, increased their reliance on, and incorporation into, Bantu society. In fact, many !Kung families currently have close relationships with individual Bantu and look on them as protectors.

For their part, the Bantu have accepted the role, often speaking of "my Bushmen." Marriage of !Kung women to Bantu men is now fairly common, an ominous sign for the cohesion of !Kung society. The children of these unions obtain full rights within the Bantu system, including the right to inherit livestock, and are more likely to think of themselves as Bantu than as !Kung.

Genetic studies of many Bantu-speaking peoples in southern Africa show that Khoisan speakers have been melding into Bantu societies for centuries. Very possibly some of those Khoisan groups and similar ones elsewhere in the world followed a course something like the modern !Kung at Dobe are following now; that is, the acquisition of goods led to a lack of mobility and to societal stresses fatal to the group's cohesion.

Today the issue of whether the !Kung experience is applicable to foraging societies of the past can best be resolved by comparing the forces acting on the !Kung with those acting on the remnants of other foraging societies in Africa, Asia and South America. These groups merit intense and immediate scrutiny. If they are ignored, an important opportunity to understand more about the ways of past foraging groups and about the forces leading to their demise will soon pass forever.

!KUNG FAMILY lives in a Bantu community. As the !Kung become more like their neighbors, the culture that was once treated in anthropology textbooks as a living exemplar of the hunting and gathering way of life may be on its way to extinction.



sometimes isolated and fenced. Hearths, formerly the focal point of social exchange, were moved inside many huts. Kraals (pens for animals) gained a central place in camp, and private food-storage structures (*c*) joined the landscape.

FURTHER READING

KALAHARI HUNTER-GATHERERS: STUDIES OF THE !KUNG SAN AND THEIR NEIGHBOURS. Edited by Richard B. Lee and Irven DeVore. Harvard University Press, 1976.

THE !KUNG OF NYAE NYAE. Lorna Marshall. Harvard University Press, 1976.

THE !KUNG SAN: MEN, WOMEN AND WORK IN A FORAGING SOCIETY. Richard B. Lee. Cambridge University Press, 1979.

OPTIMIZATION AND RISK IN HUMAN FORAGING STRATEGIES. John E. Yellen in *Journal of Human Evolution,* Vol. 15, pages 733–750; 1986.

Ancient Glazes

The beauty of a glazed ceramic arises from the interplay of light with the complex structure of the glaze. The methods of materials science reveal the ingenuity of ancient glazing technologies

by Pamela B. Vandiver

A glaze is the glassy coating that seals and adorns the surface of a ceramic body. Modern industrial glazes are usually colorless and transparent, as on a porcelain dish, or colored and opaque, as on a bathroom tile. But a visit to any art museum reveals that ceramic glazes encompass an immensely more diverse repertoire of visual effects. Blue faience glazes on ancient Egyptian amulets glow with intensity undiminished since the day they emerged from the kiln. Scenes rendered in red and black slip glazes spring to life on the sides of Grecian vases. Bright tricolor lead glazes, luminous celadons and dazzling porcelains speak to the taste and power of the Chinese imperial court.

Long before there were synthetic dyes and plastics, ceramic glazes offered artisans an unparalleled range of permanent colors and textures, which could be manipulated to satisfy a diversity of cultural demands. In most cultures, glazed ceramics were a prestige good—something that remained beyond the reach of common people because the necessary materials, know-how and manufacturing skill were often difficult to acquire. In 16th- and 17th-century Europe, monarchs collected Chinese porcelains so avidly that they sometimes risked bankrupting their treasuries. The European appetite for this "white gold" motivated prominent scientists to try to replicate porcelain during the early stages of the industrial revolution.

PAMELA B. VANDIVER is senior ceramic scientist at the Smithsonian Institution's Conservation Analytical Laboratory. Vandiver received an M.S. in ceramic science and a Ph.D. in materials science and Near Eastern archaeology from the Massachusetts Institute of Technology in 1985. She is also trained in the crafts of glassblowing and ceramics and in 1971 earned an M.A. in art and anthropology from Pacific Lutheran University. In addition to studying a wide array of ancient ceramics, she has analyzed the pigments at the Lascaux cave.

Today ceramic research focuses mainly on exploitation of the practical rather than the artistic possibilities of ceramics [see "Advanced Ceramics," by H. Kent Bowen; SCIENTIFIC AMERICAN, October, 1986]. Yet the tools and strategies of modern materials science have also taken the study of ancient ceramics into a new era. Materials scientists study ways to process raw materials in order to create microscopic structures that will endow a ceramic with desired macroscopic properties, such as the ability to withstand extreme stress or temperature. Students of ancient ceramics approach the problem from the opposite direction: starting with a finished object, they try to determine the structures underlying its properties and to reconstruct the materials and processes that created it.

In our laboratory we have been studying the technological development of ancient glazes, which have not received as much attention as the ceramic bodies themselves. We study the archaeological context and utility of an ancient glaze, analyze its composition and structure and finally try to replicate it. We hope thus to answer a multitude of questions: What is the physical basis of the appearance of a glaze? What raw materials went into it, and how do these vary or constrain the glaze? What sequence of steps did the potters choose to develop? How were these procedures refined to produce a desired effect? The answers to such questions offer an intimate window on the thought processes of ancient potters and on the values of the societies in which they worked.

Ceramics are made of clays, which consist mainly of fine particles—less than 10 microns across—of aluminosilicates, which are among the most abundant minerals in the earth's crust. The platelike particles are stacked in layers, offering narrow interstices for water, which creates capillary forces that draw the clay particles together and allow them to slide past one another so that the clay can be plastically deformed.

When a clay is fired, the particles are sintered; that is, they melt just enough to make a glassy "glue," which binds the particles into a rock-hard mass. The ceramic remains somewhat porous, however, because the clay is not heated high enough to melt completely; if it were to melt, the structure would collapse under its own weight. The porous body is therefore sealed by a glaze. Glazes are typically composed of fine, glass-forming particles, which melt and fuse to form an impermeable glassy layer after only a short firing at a relatively low temperature that retains the shape of the clay body.

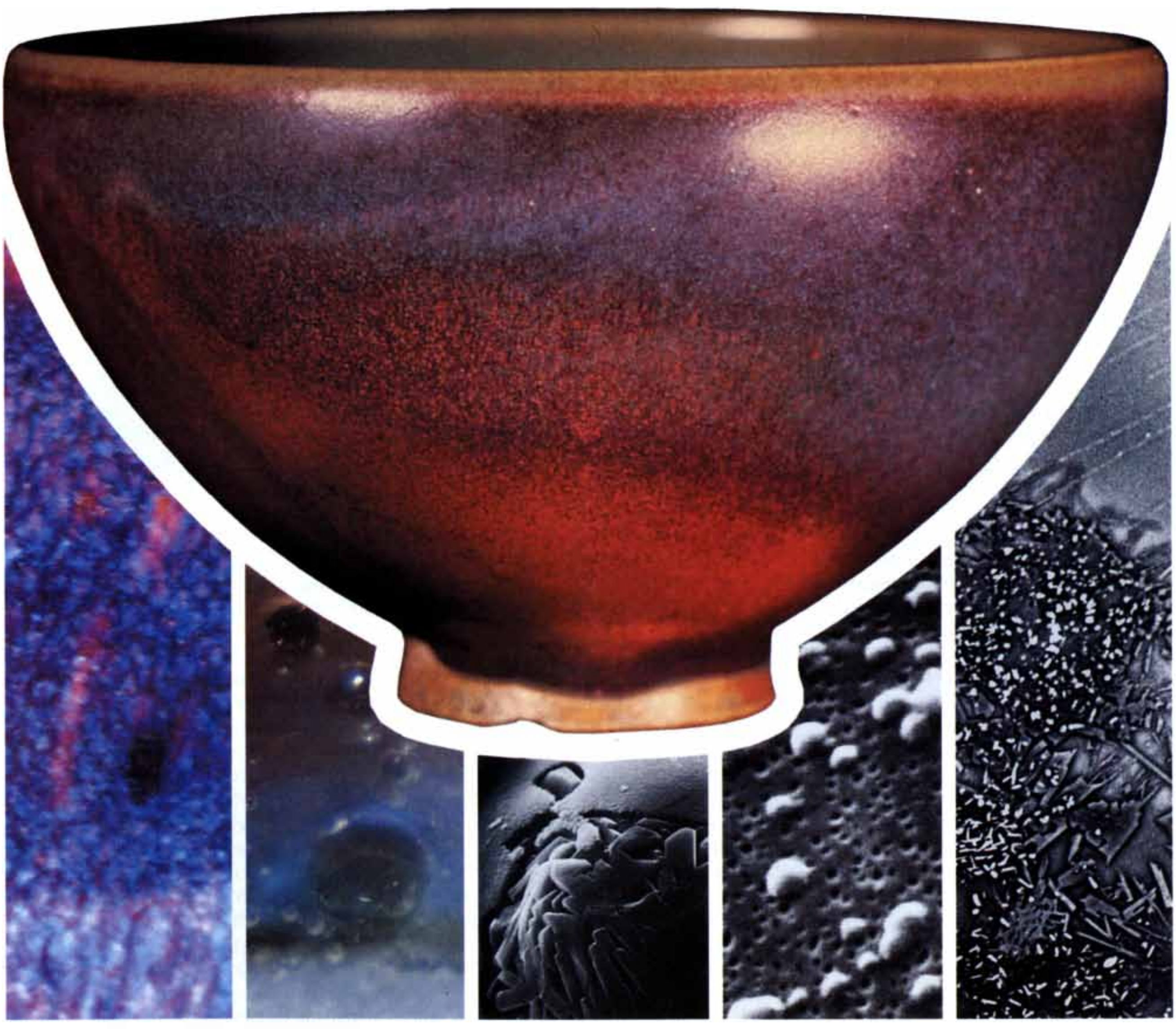
Perhaps the best way to elucidate the melting behaviors of such ceramics as a glass or glaze is to contrast them with that of a crystal. In a crystal the atoms are stacked in orderly, three-dimensional arrays. The bond strength between like atoms in a crystal is uniform. Consequently, as a crystal is heated, the atoms vibrate at fixed sites in the array until they reach the melting point, when the thermal energy makes them abruptly disassemble into a more random, liquid state. A glass, on the other hand, contains impurities, called fluxes, that disrupt the arrays, giving rise to a random network analogous to the structure of a liquid. Because of the random arrangement and the variable bond strengths, glasses do not have a fixed melting point. When a glass is heated, it dissolves gradually, changing from the consistency of a solid to that of butter, then honey and finally syrup. By incorporating fluxes that have different bond strengths, one can manipulate the melting point and stability of a glass or glaze.

Although glasses and glazes are made of the same raw materials, such as silica sand, lime, potash, borax and lead oxides, there are important differences between them. A glass in the making must be kept above the melting point for a long time, sometimes days, to ensure that all the raw materials melt, that no new crystals form and that trapped air escapes as bubbles or is reacted into the glass. A glaze, in contrast, cannot be held at peak temperature for long, because the clay body might also begin to melt.

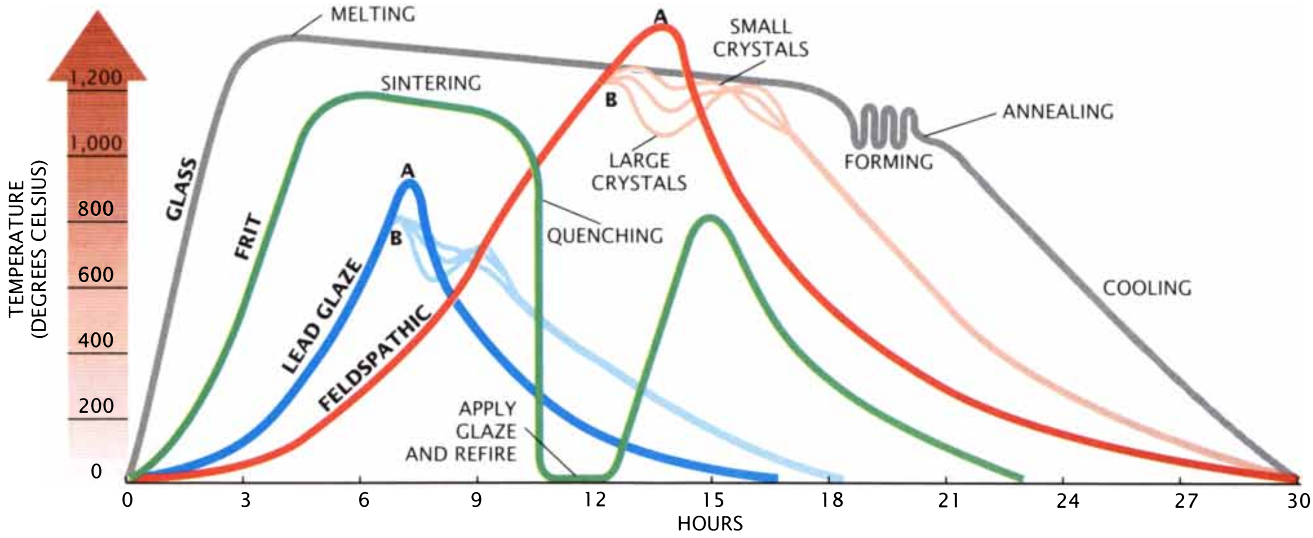
Because glazes spend so little time at or near a high temperature, they often retain impurities, such as unmelted crystalline raw materials or gas bubbles. It is these impurities that dramatically alter the appearance of a glaze, making a clear glaze appear translucent or opaque, causing an inherently smooth surface to take on a satin, velvet or even oatmeal texture, and changing an optically flat surface to one having the same illusion of depth as jade and agate.

The appearance of a glaze results from the interactions of light with the outer surface of the glaze, with the interface between the glaze and the underlying body and with the substance of the glaze itself. The microstructure of the glaze is particularly important. Structures smaller than 100 microns lie below the limit of resolution of the human eye and so contribute to subtle textural effects. Structures of between .1 and one micron in diameter are near the visible wavelengths of light and so most strongly influence the optical properties and appearance of a glaze.

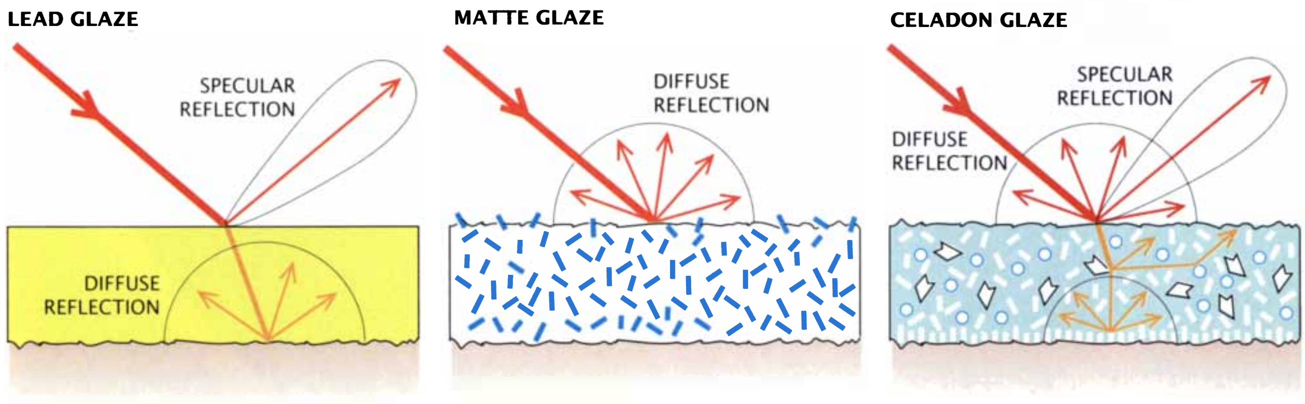
Light bounces cleanly off a smooth glaze surface (specular reflection), whereas it scatters off a rough surface. A matte glaze contains crystals that roughen the surface and scatter light in all directions. In Song dynasty Longquan celadons, the surface is slightly rough, embedded with quartz crystals measuring between 10 and 100 mi-



SOFT, OPALESCENT GLAZE on this Jun ware bowl exemplifies the sophistication of the Song dynasty (A.D. 960–1279). In a modern replica of a Jun glaze (*left*), fields of copper crystals form a myriad of reddish pinpoints within a blue ground. A cross section enlarged 60 diameters shows bubbles, which reflect light (*second from left*). Cristobalite crystals indicate that the glaze was heated for a long time and cooled slowly (*center*). Tiny droplets in the glaze emulsion were etched away by acid, leaving a pitted surface; the white balls are pseudowollastonite crystals, which make the glaze cloudy (*second from right*). Needles of anorthite grew between the glaze and the body, forming a white layer that reflects light (*right*).



THERMAL HISTORY of glass, frit and two glaze types are shown here. Glass is held at a high temperature to melt impurities and allow bubbles to escape; while being formed, the glass may be reheated many times. A frit is sintered (partially melted) at a high temperature, cooled rapidly, ground into powder, applied as a glaze and refired at a lower temperature. Clear lead and feldspathic glazes are held at peak temperature briefly (*A*). To achieve translucency, the glaze is cooled slightly (*B*) to nucleate crystals and then held at a higher temperature for some time to allow the crystals to grow.



LIGHT interacts in different ways with different glazes. A lead glaze (*left*) is transparent and highly reflective. A matte glaze (*center*) has crystalline protrusions at the surface, which scatter light. A celadon glaze (*right*) contains quartz particles, bubbles and micron-size crystals, which bend and scatter light. A crystal layer at the glaze-body interface also reflects light.

crons, and reflects a combination of specular and diffuse light. Different glaze compositions reflect and absorb varying amounts of light. A soda-lime glaze reflects about 4 percent of the incoming light, whereas a lead glaze reflects about 8 percent, which gives a brighter appearance. Heavier elements increase both the refractive index and reflectivity of a glaze.

In a transparent or translucent glaze, light is absorbed, scattered and reflected at the interface between the glaze and the body. A smooth, white clay body will reflect a large amount of light and brighten the colors of the glaze, whereas a darker body will absorb more light and cause the glaze to appear darker. In celadons, crystals of anorthite, a calcia-alumina-silicate, grow at the interface during firing; the white crystals mask the gray clay body and cause the blue-green glaze to appear brighter and more intense.

A variety of complex optical effects are generated within the glaze itself. Most obvious are the interactions that color a glaze. The simplest and most reliable coloring technique is to add a pigment, such as magnetite black and copper oxide red. More complex are colloidal colors, such as microscopic particles of gold, silver or copper; the particles produce colors by absorption, scattering and refraction.

The most subtle and interesting colors are often produced by yet a third means, a solution of transition-metal ions, which have outer electrons that absorb all but certain wavelengths of light. These ions include iron (which can produce colors ranging from yellow and green to brown and black), manganese (purple to brown), chromium (pink to green), cobalt (blue) and copper (green to blue) depending on concentration and oxidation state. These colorants can be tricky to use, since the energy level of their outer electrons is strongly influenced by the surroundings. Hence, copper is blue in an alkaline glaze but green in a lead glaze. If .5 percent of iron oxide is added to an alkaline glaze or glass, each iron ion becomes surrounded by oxygen atoms, and the resulting absorption pattern gives a Coke-bottle-green color. If a sulfur or carbon ion replaces one or more of the oxygens, the resulting color is a beer-bottle brown, because an iron-sulfur or iron-carbon pair absorbs more light across the entire spectrum.

Air bubbles also interfere with the path of light through the glaze. Bubbles may form because trapped air pockets expand with heat or aggregate as melting particles clump together or because salts in the raw materials decompose and release gases. If the glaze is allowed to melt to a fairly fluid state, most of the air bubbles rise to the surface and escape. But if the glaze remains viscous, the bubbles are trapped, much like bubbles in pumice. Large numbers of bubbles brighten a glaze, because the smooth interfaces between air and glass offer many reflective surfaces.

In addition to colorants and air bubbles, another important feature of the glaze is the presence or absence of crystalline particles. If a glaze or glass consists of .5 percent by volume of fine particles of less than a micron in size, it appears translucent, not transparent. If the concentration exceeds about 10 percent by volume, the glass or glaze appears opaque. The crystals may have a higher index of refraction than the surrounding glass, in which case light bends as it passes into a crystal and traverses a longer path through the glaze, creating an illusion of greater depth. If the index of refraction of the crystal is much greater than that of the glass, light bends by such a large angle that the glaze becomes quite opaque.

Artisans over the millennia have exploited these structural variations in glazes to create an amazing array of visual effects, including those that mimicked precious metals and stones. Ceramics could be plastically formed, usually with less effort than more precious materials, and could assume a greater diversity of shapes and sizes than metal or stone. These possibilities elevated glazed ceramics above mundane functionality to the status of luxury goods. To satisfy their patrons, artisans exercised considerable ingenuity in manipulating glazes. Ceramic glazes, then, offer an exceptional glimpse of ancient technology and its interplay with cultural values.

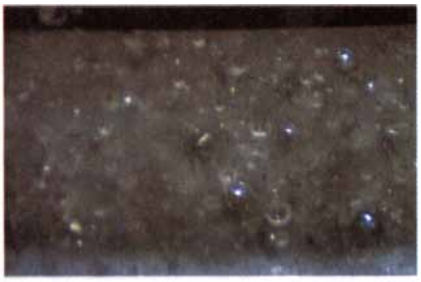
The earliest types of glazes in the New World, Southwest Asia and China were slips: suspensions of very fine clay particles in water, often with an added flux of a salt or ash. Slip glazes are fired for a short time at a low temperature, between 600 and 1,000 degrees Celsius. They melt only partially and so remain permeable.

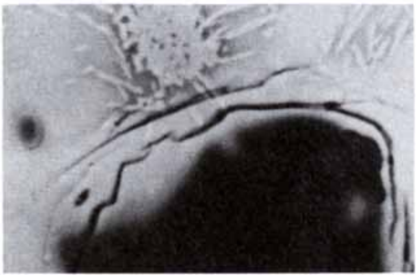
As early as 5500 B.C. potters in northern Mesopotamia had discovered that the color of a fired clay could be controlled by adjusting the furnace atmosphere. The variable agent is iron oxide, an abundant impurity in earthenware clays. An oxygen-rich kiln atmosphere maintains the iron oxide in its red, hematite form (Fe_2O_3), whereas a smoky, oxygen-depleted atmosphere reduces iron oxide to black magnetite (Fe_3O_4). Both conditions arise in ordinary hearth and kiln fires, and so people would have observed the coloring effects and then learned to manipulate the firing to achieve them on purpose.

Slip-decorated wares reached their apogee in the Attic vases of the sixth to second centuries B.C. Illitic clays mixed with water were left standing so that coarse particles would settle out. The finest clay and hematite particles remained suspended and were ladled out. This slip was painted on the areas to be rendered black. The wares were fired in a smoky atmosphere so that the iron oxide over the entire surface was reduced to a black color. In the process the fine-particled slip glaze sintered into a glossy surface, whereas the coarser clay body remained largely unsintered. At the end of the firing and during cooling, more oxygen was vented into the kiln, causing the exposed, porous clay areas to reoxidize to a red color. On the other hand, very little oxygen could diffuse into the sintered slip, and so the glazed areas remained black.

Slip glazes everywhere were variants of this general style and shared the drawback of being too permeable. We recently identified a notable exception to this rule in a ware recovered from the site of Godin Tepe in western Iran. Known as Seh Gabi painted ware and dated to 3500 B.C., it was made from a calcareous, cream-colored clay and was decorated with a glassy black glaze. Our analysis showed that the glaze has the composition of a slip but had been fired at a higher temperature (1,050 degrees C) and for a longer time than the body, which was fired at 800 degrees C. This could have been done only if the glaze had been fritted, or prefired, at the high temperature and then ground up, mixed with water, applied to the clay body and refired at the lower temperature. This technology was short-lived; it was transmitted neither to subsequent generations nor to other areas. Seh Gabi remains the sole known example of a high-temperature clay-based glaze before 1500 B.C., when stoneware glazes appeared in China.

The first truly glassy glazes in the Near East developed from the stone-





LONGQUAN CELADON VASE from the 12th century is among the finest ceramics ever made. Bubbles and quartz particles, seen in cross section (*top right*), give the glaze luminosity and depth; the sea-green hue comes from a slightly reduced state of iron in solution. A scanning electron micrograph (*bottom left*) shows a quartz particle and needles of anorthite, which scatter light.

worker's art rather than from the potter's craft. These were the turquoise- and lapis-like blue faiences of Egypt. Around 4000 B.C. stoneworkers discovered they could create a stonelike material by shaping a gritty paste of crushed quartz debris, a sodium or sodium-potassium flux, some crushed limestone or malachite and water and then drying and heating it. A copper salt was added to the paste, which, together with the other fluxes, would effloresce onto the surface; when fired, the efflorescence would melt and form a translucent, blue, glassy coating. A second method, found from around 2300 B.C., involves firing quartz-based objects buried in a powder of copper salts, quicklime and charcoal; the salts vaporized and fused onto the surface to form the glaze. This technique may still be used in Iran to make beads. Faience artisans also developed a third group of techniques in which glazes were fritted, powdered and applied to quartz-based bodies.

The technological history of this important craft is complicated in Egypt, especially during the early stages. As the technology spread to—or was reinvented in—other regions of the Near East, differences in technique arose between regions and even between objects ascribed to a single workshop.

Around 1500 B.C., about the time Egyptian artisans first began making large quantities of glass beads and vessels, they also started incorporating powdered glass into quartz-based bodies to make decorative inlays for furniture and architecture. This development expanded the palette of colors from the traditional blue-green, dark purples, brown and black to include yellow, lime green, cobalt blue, violet and orange. Artifacts from Roman times indicate that Egyptians applied these frit glazes to tan-colored clay objects; these glazes appear glassier and duller and lack the gemlike quality or function of faience.

It is of interest to note that prefritted glazes have been reinvented in modern times as a way to eliminate defects such as bubbles and unmelted particles. Such glazes also solve the problem of possible toxicity from lead glazes, which, when incompletely melted, contain unreacted lead oxides that can dissolve into mildly acidic foods such as orange juice. A prefritted glaze ensures that the lead particles are completely melted and locked into a silicate glass structure.

The stoneworking and faience-making traditions in the Near East, then, developed a glaze technology that differed from that of slip glazes, one offering a wider range of colors and a high gloss. For historical and cultural reasons, faience, frit, glass and glaze were reserved for imitating precious materials and for architectural decoration—artifacts that supported the ruling class. Glassy glazes could have also been developed to decorate and seal clay pottery, but they were not needed for these purposes, because other methods, such as slip decoration and impermeable resinous coatings already sufficed.

One reason slip glazes were firmly entrenched in the Near East was because clays in the region are of a pyrophyllite type suitable only for low-fired earthenware. These clays are difficult to work and are often found in seabed or riverine deposits that vary in composition. Clays adulterated with enough limestone have poor refractory properties: if fired in the range of 800 to 850 degrees C, they crumble easily; if fired at above 1,000 degrees C, they will be more stable but will then warp, bloat and eventually melt. What is more, kiln structures made from the poor clay cannot withstand repeated high firings. The quality of raw materials, together with the social and historical setting, constrained the technology and fixed the mindset of artisans. Hence, glossy pottery glazes made from such materials as limestone did not appear until Roman times.

In China, in contrast, potters enjoyed an abundance of excellent refractory clays and materials that enabled them to elevate ceramic technology to an art equal to painting and sculpture in the West. The analysis of Chinese glazes by our laboratory and by Chinese scientists reveals a gradual but continuous pattern of innovation in which new materials were employed for glazing and in which processes were invented or refined to produce new visual effects.

The chief driving forces of innovation in Chinese glaze technology were two: the early achievement of firing temperatures above 1,000 degrees C and the discovery of glaze stones—nonclay minerals such as limestone and "China stone" (a partially weathered mix of sericitic clay and quartz) that can be used to concoct a high-firing glaze. The first achievement can be credited to the thick loess deposits, in some places hundreds of meters deep, that blanket much of China. The loess consists primarily of quartz and so has a high melting point, which makes it a superb material for high-temperature kilns. To build a kiln in China, one had only to dig a chamber in a hillside, tamp the walls and create a vent to the surface. By 1500 B.C. pottery was often high fired at between 1,100 and 1,200 degrees C.

The idea of glaze stones might have come from observing the results of overfiring slip glazes containing clay, limestone and mica; deposits of ash from the kiln fire might have served as a flux for embedded particles of glaze stone. By the second millennium B.C. Chinese potters were using two types of glazing stones: limestone and China stone. Ground China stone formed the clay body itself, a process that culminated around A.D. 600 with the invention of porcelain, which was then predominantly of China stone. Kaolin, now the most common ingredient, was added only later.

By the 11th century A.D. artisans in northern China were supplying the imperial court with extraordinary wares such as the lavender-and-blue Jun ware and Ru ware, whose color was described as "blue sky after rain." Jun ware illustrates an unusual way to create an unctuous, stonelike texture. The glaze was allowed to sit at a temperature of about 1,250 degrees C and then cool slowly, forming an emulsion of two glassy liquids. Like oil and vinegar, the two liquids are more stable by themselves than they are in a mixture, so that they form a myriad of tiny globules some tens or hundreds of nanometers in diameter. The two glass phases have different indexes of refraction, causing light to bend as it crosses from one phase to another; the longer, more indirect light path creates the illusion of translucence and depth. The same effect appears in a vigorously shaken bottle of oil-and-vinegar salad dressing; the individual liquids are clear, but the emulsion is translucent.

In 1127 the Song court fled south and established the Southern Song dynasty in Hangzhou. Potters began a crash research program to adapt the local Yue folkware to the court's taste for lustrous, jadelike wares. The results of this effort were the celebrated Guan ("imperial") celadons and Longquan celadons. Longquan was a market town southwest of Hangzhou, where celadons were made primarily for export. The remarkable accomplishment of celadon glazes is their translucent, silky texture ranging in hue from pale sea green to gray-green.

The starting point for the celadon artisans were the Yue glazes, which are homogeneous, thinly applied clear glazes colored by iron oxide in shades

1

2

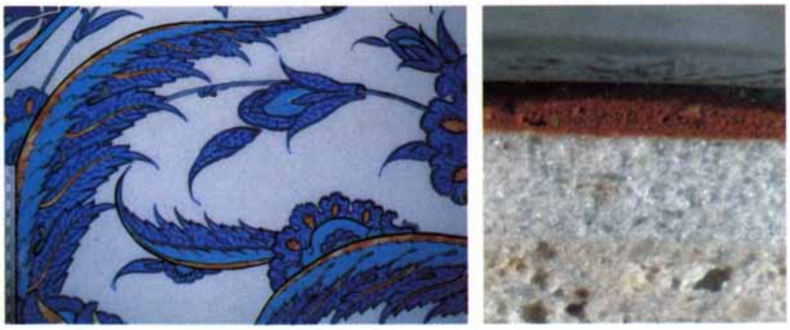
3

4

5

DIVERSITY OF GLAZES reflects variation in materials and processes. Fifth-century B.C. Attic wine cup (*1*) was decorated with a black slip over a red ground; a scanning electron micrograph shows sintered slip particles above a coarse clay base. Egyptian faience (*2*) is a glassy glaze over a ground-quartz body. Shown here are a superb chalice (800–900 B.C.) glazed by melting copper salts that had effloresced from the body; beads replicated in the author's laboratory by firing them while buried in copper salts, quicklime and charcoal; and an inlaid design (about 1450 B.C.) made by mixing colored glass with the quartz paste. A two-foot-long lion (*3*) from the Temple of Ishtar in Nuzi, Iraq, dates to about 1500 B.C. and is the largest extant ancient ceramic with a glassy glaze. Bright lead glazes decorate an eighth-century Chinese jar (*4*). Lustrous "oil spots" on a Song dynasty tea bowl (*5*) are made by red hematite and yellow magnetite crystals. Intensely colored overglaze enamels (*6*) from the *famille verte* palette adorn an early 18th-century plate; the scanning electron micrograph shows the ultrafine hematite particles in red enamel. A Turkish Iznik tile (*7*) is decorated with vivid colors under a clear lead-alkali glaze.

6



7

ranging from brown to green and yellow. Yue glazes are fluid because of their high calcia content. The celadon glazes were as much as 10 times thicker; to prevent the glaze from flowing, the calcia content was reduced, and many quartz particles and bubbles were suspended in the glaze. In addition, celadon glaze, unlike Yue, contains numerous crystalline particles, mainly needles of anorthite several microns in length and spherical particles of pseudowollastonite.

When we analyzed the chemical composition across the thickness of the glaze, we found that areas high in potassia and alumina but low in lime contain mainly anorthite, whereas areas rich in lime contain mainly pseudowollastonite. These local variations were caused by combining coarsely ground raw materials, such as limestone, with China stone and ash and mixing them incompletely. Since in Yue ware the same materials were finely ground and well mixed, celadon was clearly the result not of careless workmanship but of an intentional technology.

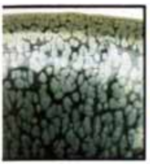
The firing process also underwent a change. Undissolved quartz particles in celadon are surrounded by a halo of molten silica, which indicates the glaze was kept at a high temperature for a long time—but not so long that the molten silica could recrystallize into cristobalite (one of the crystalline phases of silica). Based on replication experiments, we conclude that celadons were fired at a temperature of between 1,200 and 1,250 degrees C and then cooled over many days. This process allowed anorthite and pseudowollastonite crystals to form in the glaze. Nearly nine centuries after the invention of celadon, engineers at Corning Glass Works developed a high-tech analogue in Corelle Ware. The ware is formed as a clear glass and then placed in a controlled furnace that allows crystals to precipitate and grow. The process strengthens the glass and turns it an opaque white.

Another famous glaze in which crystals are induced is the "oil spot" Jian or Temmoku glaze. This glaze contains about 10 percent iron oxide; when it is kept in a viscous but molten state for a considerable time, "snowflakes" of gold-colored hematite and

		IRAN 3500 B.C. SEH GABI PAINTED WARE	EGYPT 3500 B.C. BLUE FAIENCE BEADS FROM BADARI	IRAQ 1500 B.C. WALL NAIL FROM NUZI	GREECE 500 B.C. ATTIC RED-AND-BLACK SHARD	SOUTH CHINA A.D.1000 YUE WARE JAR	NORTH CHINA A.D.1100 JUN WARE BOWL	SOUTH CHINA A.D.1200 LONGQUAN CELADON
GLAZE		OPAQUE GLASSY BLACK PREFRITTED SLIP	TRANSLUCENT GLASSY BLUE HIGH-COPPER ALKALINE	OPAQUE GLASSY BLUE HIGH-CLAY ALKALINE	OPAQUE GLOSSY BLACK REDUCED HIGH-IRON SLIP	TRANSPARENT GLASSY OLIVE HIGH-LIME CHINA STONE	TRANSLUCENT GLASSY RED-AND-BLUE CHINA STONE	TRANSLUCENT GLOSSY BLUE-GREEN CHINA STONE
BODY		CALCAREOUS EARTHENWARE CLAY	QUARTZ	HIGH-IRON CLAY	HIGH-IRON EARTHENWARE CLAY	CHINA STONE	CHINA STONE	CHINA STONE
GLAZE FIRING TEMPERATURE (DEGREES CELSIUS)		1,000 900	1,000 800	1,000 900	1,050 850	1,200 1,100	1,300 1,200	1,250 1,200
OXIDES (PERCENT BY WEIGHT) REFRACTORY COMPONENTS	SiO_2	59.00	78.97	51.46	45.63	60.40	71.70	67.00
	Al_2O_3	14.54	.51	10.10	32.52	12.80	10.40	13.60 – 16.70
FLUXES	Fe_2O_3	13.59	.19	4.35	13.61	1.25	1.74	1.60
	CaO	4.98	.70	8.93	.45	17.20	8.53	7.20 – 10.13
	MgO	1.17	0	4.23	2.40	2.30	1.60	.65
	Na_2O	1.60	9.25	14.37	.64	.95	1.30	.50
	K_2O	3.20	.24	3.31	4.18	1.60	3.92	5.60
	PbO	0	0	.15	0	0	0	0
IMPURITIES AND COLORANTS	TiO_2	.49	—	.80	.51	.70	.21	.08
	P_2O_5	.10	—	—	—	1.71	.42	.04
	Cl	—	.59	—	—	0	0	0
	CuO	.20	9.46	1.37	—	—	.10	—
	MnO	.31	.06	.01	—	—	—	.55
TOTAL		99.18	99.97	99.08	99.94	98.91	99.92	96.82

silvery magnetite grow. If the glaze temperature is raised, the glaze flows, and so the crystals may dissolve and run in downward streaks, forming a "hare's fur" texture.

Scholars think glazes such as celadon, Jun and oil spot emerged as a consequence of the large-scale production of high-fired ceramics in China. As kilns grew in size to accommodate the increased output, the time at peak temperature and the cooling period lengthened, and one accidental result was crystal formation. Potters noticed this and, when their patrons desired such textures, figured out how to control the effects. Modern ceramicists are so attuned to efficiency that the "inefficient" Song dynasty techniques—coarse milling, poor mixing



SOUTH CHINA A.D. 1200	TURKEY A.D. 1500	CHINA (K'ANG HSI REIGN) A.D. 1713
JIAN OR TEMMOKU BOWL	IZNIK WALL TILE	JINGDEZHEN PLATE
OPAQUE GLOSSY HIGH-IRON OIL-SPOT CHINA STONE	PREFRITTED COLORS UNDER TRANSPARENT GLASSY LOW-LEAD	HIGH-LEAD RED ENAMEL OVER TRANSLUCENT CHINA STONE
HIGH-IRON STONEWARE	SILICEOUS FRITWARE	PORCELAIN
1,200 1,100	1,100 1,000	900 700
60.11	63.10	24.90
19.33	.93	4.78
7.89	.40	23.76
5.99	1.20	2.39
1.75	.98	.11
.11	16.81	.28
2.55	1.32	.25
0	13.70	42.90
.65	.02	0
1.05	—	.10
0	—	—
—	—	—
—	—	—
99.43	98.46	99.47

CHEMICAL COMPOSITIONS typical of the glazes on the objects shown here were determined in the author's laboratory.

and long firing—are difficult to understand and duplicate.

After the discovery of glaze stones, the next major technological advance was the development of lead glazes. Archaeologists have not yet determined whether these first appeared in the Mediterranean region, Southwest Asia or China, but they agree that lead glazes existed in all three regions by about 2,000 years ago—a time famous for the consolidation of the Roman, Parthian and Han empires, which allowed extensive mutual trade and communication.

Lead glazes were formulated from a new class of lead-containing glaze stones. These glazes can be ground up, applied directly, fired at a low temperature (from 800 to 1,000 degrees C), support a wide variety of bright colors and produce a brilliant surface because of the high refractive index of lead glass. Lead glazes could be applied very thinly and so were used on delicate porcelain bodies to make an imitation silverware. They had the disadvantage, however, of being too fluid and flowing easily.

In the Song and Ming dynasties, lead glazes provided the basis for overglaze enamels: intensely colored glazes that are fritted, pulverized and then painted or sprayed onto a glazed body that has already been fired. The decorated ware is then fired a second time at a lower temperature to sinter and bond the enamels to the glaze. The enamels are formulated to all melt at the same temperature, usually around 800 degrees C. Overglaze enamels are particularly important for decorating porcelain, which must be fired at such a high temperature that most colors would decompose and diffuse into the glaze. The enamels placed a wide range of compositions, colors and textures at the disposal of artisans.

In 12th-century Persia, ceramic artisans developed overglaze enamels from lead glazes for a ware called minai. In general, however, Near Eastern artisans preferred to paint designs in colors that were applied under the glaze (similar to the Chinese blue-and-white tradition). That technology reached a peak with the Turkish Iznik tiles created in the reign of Süleyman the Magnificent (A.D. 1520-1566). The tiles are decorated with exuberant floral designs under a brilliant, clear lead glaze. To sharpen the contrast between the colors and the background, the tan clay tile was first covered with a layer of white quartz paste. Great care went into formulating the underglazes. The artisans made blue, green and turquoise frits using cobalt and copper; some of the frits have higher indexes of refraction than the clear glaze and so produce strong, saturated hues. Reds, gray-greens and black were obtained from stable pigments such as iron oxide, chrome ores and chromite. Iznik tile was a tour de force, a culmination of glazing artistry that evolved out of the faience tradition of the Near East.

Modern scientists rarely think of technological progress as something that could have taken place outside the context of the scientific and industrial revolutions of the West. Yet the artisans of antiquity were able to invent and refine complex and diverse technologies—many of which have direct analogues in high-tech ceramics—without the benefit of modern analytical methods, instruments or communications. It is only with the help of those modern tools, however, that we now know about their achievements.

New knowledge of ancient glazes benefits craftsman, connoisseur and culture historian alike. For example, in the early 1980's Chinese ceramic technologists applied knowledge gained from the scientific study of ancient celadon production to make exact replicas of Longquan celadon for the first time since the 13th century. The new celadons could fool even experts. For the lover of ceramic art, an understanding of the interplay among physics, geology, culture and history that underlies the visual properties cannot but enhance appreciation of these ancient masterpieces.

FURTHER READING

CERAMIC GLAZES. Third Edition. Cullen W. Parmelee and Cameron G. Harman. Cahners Books, 1973.

THE PHYSICS AND CHEMISTRY OF COLOR. Kurt Nassau. John Wiley & Sons, Inc., 1983.

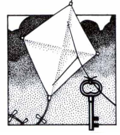
CERAMIC MASTERPIECES: ART, STRUCTURE, TECHNOLOGY. W. David Kingery and Pamela B. Vandiver. The Free Press, 1986.

INTRODUCTION TO FINE CERAMICS: APPLICATIONS IN ENGINEERING. Edited by Noboru Ichinose et al. John Wiley & Sons, Inc., 1987.

MATERIALS ISSUES IN ART AND ARCHAEOLOGY. Symposium Proceedings of the Materials Research Society, Vol. 123. Edited by E. V. Sayre et al. Materials Research Society, 1988.

THE AMATEUR SCIENTIST

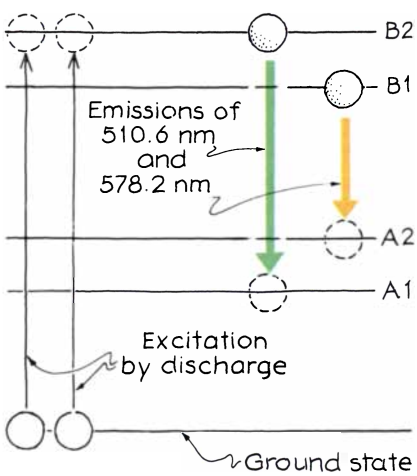
A homemade copper chloride laser emits powerful bursts of green and yellow light



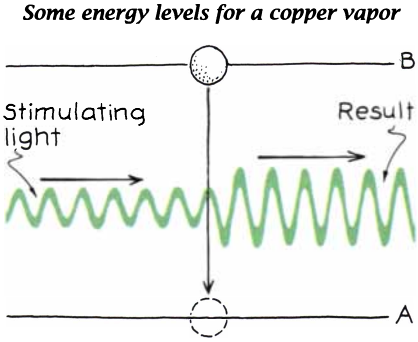
by Jearl Walker

Anyone who has tried to construct a homemade gas laser that operates in the visible spectrum knows how difficult and expensive it can be to carry out the ambitious project. Even the popular helium-neon laser can easily be beyond the reach of an amateur. The laser tube requires precise glasswork, for example, and the dielectric mirror that reflects light back through the tube to maintain lasing is often beyond the budget of an amateur.

Now Martin Gosnell of Charmhaven, near Sydney in Australia, has sent me plans and instructions that will enable an adept amateur to build a copper chloride laser without great expertise and at a reasonable cost. The lasing element is copper vapor, which is generated when the first of a pair of electrical discharges is sent through the laser; the second discharge, about 150 microseconds later, causes the vapor to lase with a pulse of green and yellow light. Gosnell is able to fire the laser at a rate of up to 50 times per second, so that the output appears to be continuous.



Some energy levels for a copper vapor



Stimulated emission

Although the project is certainly quite challenging, little specialized glasswork is required, and the light production is so strong that an aluminum-coated microscope slide can be substituted for the dielectric mirror. (A strong word of caution: the high-voltage discharges involved in the lasing are lethal, and so no one should consider building the laser who has not had substantial experience with high-voltage circuits.)

Before I get to the details of construction, I should explain how it is that a copper vapor can be made to lase. The simplified chart in the upper illustration at the left indicates some of the energy levels that a copper atom is permitted by quantum mechanics to occupy. The atom is initially in its "ground state," or at the lowest energy level. When a discharge runs through the vapor, an electron in the current collides with the atom and transfers enough energy for the atom to jump to a higher energy level—to be in an "excited state." Two pairs of excited states are shown on the chart. (The levels in each pair differ slightly in energy, for reasons that I shall not consider here; I have labeled each pair arbitrarily rather than with reference to their true spectroscopic designations.)

An excited atom can "de-excite" to a lower level by spontaneously emitting a photon, which carries away the energy the atom loses in making the jump. Because the energy levels that are allowed an atom are preset, the energy of the photon is restricted to certain fixed values. The light that is emitted is often pictured as being a wave instead of photon, in which case the wavelength associated with the wave is restricted to certain fixed values. For example, if a copper atom jumps from B2 to A1 in the chart, it emits light with a wavelength of 510.6 nanometers, which is green. A jump from B1 to A2 releases light with a wavelength of 578.2 nanometers, which is yellow.

An excited atom can also de-excite in a process known as stimulated emission. Suppose that a copper atom in the B2 state is passed by a wave with a wavelength of 510.6 nanometers—just the wavelength the atom would emit with a spontaneous jump to A1 [*see lower illustration on this page*]. The passing wave interacts with the atom, forcing it into the jump as if on command. In this process the passing light is said to stimulate the emission, preventing the atom from making some other jump, such as to the ground state, or even from making the same jump later on. The wave emitted by the atom reinforces the original wave because they are identical in wavelength, travel in the same direction and are "coherent," or locked in step, and so the resultant light is brighter than the original light. Similarly, an atom that is initially in the B1 state can be stimulated into jumping to the A2 state by light with a wavelength of 578.2 nanometers. In each case the pair of levels involved in the stimulated jump is referred to as the laser pair, because they are the basis of the laser's emission of light.

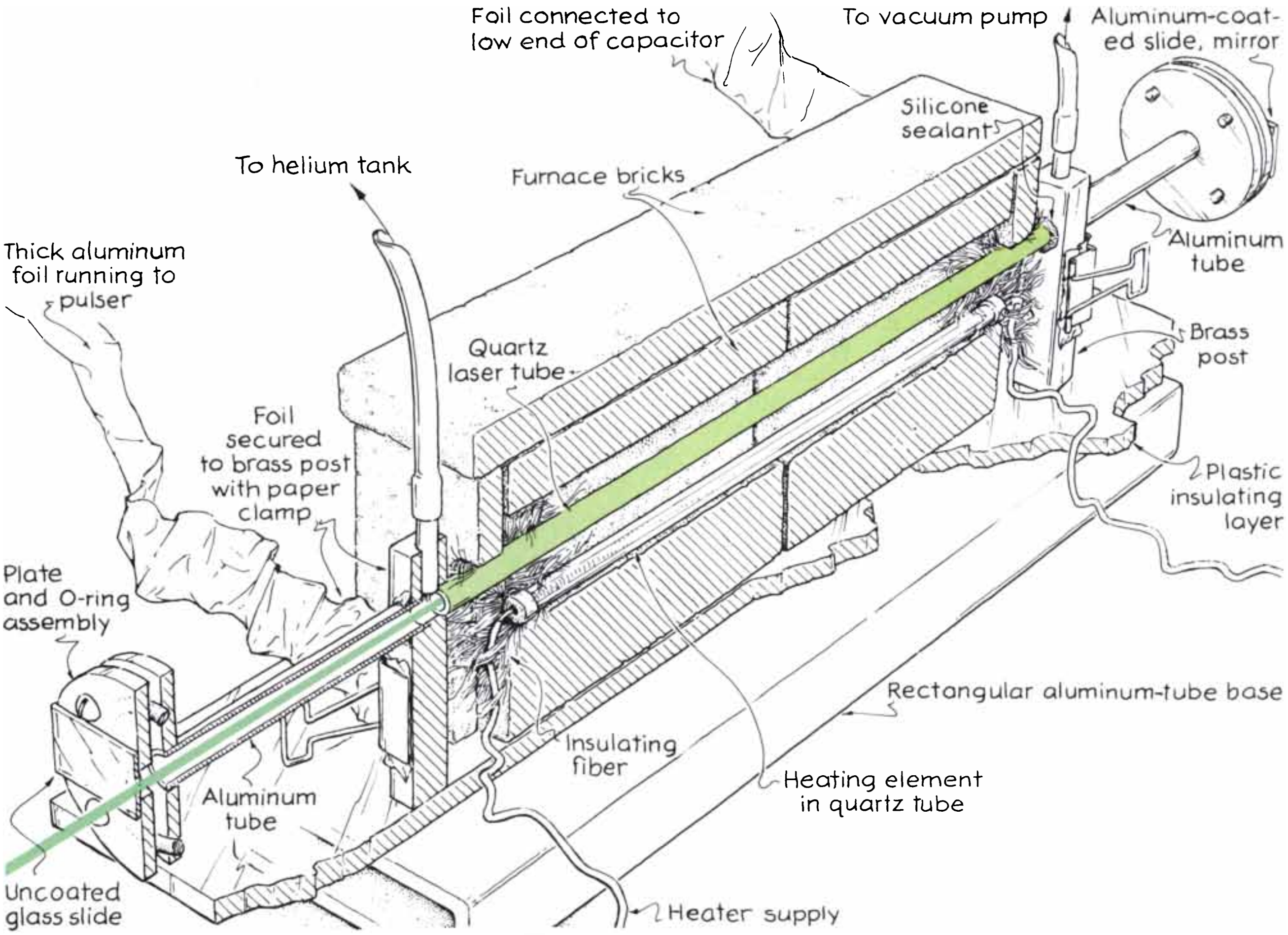
In a copper-vapor laser the idea is to excite the atoms into the B states by bombarding them with the electrons in an electrical discharge. Some of the excited atoms will then happen to jump spontaneously down to the A states and fortuitously emit waves (or photons) along the tube. That emitted light is quickly reinforced by the chain reaction of stimulated emission it sets up when it passes other atoms in the B state. In the wave picture of light, the wave grows stronger; in the photon picture, the number of photons increases. The light that reaches a mirror at one end of the tube is largely reflected back through the atoms for another go at stimulating jumps. The light that reaches the opposite end of the tube, which lacks a mirror, escapes as the laser beam.

When gas lasers first appeared in the early 1960's, a copper-vapor laser was an intriguing prospect because it promised to be more efficient than other gas lasers. One reason for the expected efficiency is the fact that what I have called the B levels are not high above the ground state, so that not much energy should be required for atoms to reach them. Research soon revealed a major drawback, however: a copper vapor had to be heated to a temperature of about 1,500 degrees Celsius if it was to lase.

In 1973 researchers discovered that when a copper halide such as copper chloride was substituted for the pure copper in the original design and a series of double-pulse discharges was sent through the tube, the required temperature was a more attainable 400 degrees C. The success was attributable to the role of the pair of discharges. The first discharge dissociates the molecules while also exciting and ionizing some of the released atoms. If the second discharge is delayed long enough for the copper atoms to settle back to the ground state but not long enough for them to recombine with the chloride, it excites the copper atoms into the B states, just as in the original design with pure copper as the source of the vapor. The double-discharge technique was an excellent idea, but it required expensive switches, a double-pulse signal generator and other costly electronics. What Gosnell managed to do was to build a copper chloride laser with inexpensive and more readily available parts.

The core of Gosnell's laser is a quartz tube that passes through a furnace fashioned from alumina-silica furnace bricks [*see illustration below*]. The tube, 55 centimeters long and with a one-centimeter bore, extends into brass posts at each end of the furnace. The posts, each standing about four centimeters beyond the oven, are 25 millimeters square and 12 centimeters high. The tube is sealed into the holes in the posts with a silicone sealant.

On the opposite side of each post, the laser is extended with a 13-centimeter-long aluminum tube; it cools and condenses the internal vapors so that they do not reach the optical elements, which are at the far ends of each aluminum tube. At one end of the laser an aluminum-coated microscope slide serves as a mirror. (The mirror is mounted with its reflecting surface on the exterior.) At the other end an uncoated glass slide allows the laser beam to escape from the tube. (Were the vapors to condense on either slide, the lasing action would be eliminated.) To mount each slide so that it could later be adjusted, Gosnell devised an assembly of plates that are separated by an O ring and held together by three screws. He bored a hole through



Martin Gosnell's copper chloride laser tube and its furnace the assembly and then, with the sealant, glued the assembly up against the end of the aluminum tube and fixed the slide over the exterior of the hole.

When the laser is fired, the discharge through the tube runs between the brass posts, which are electrically connected to the power supply. Current is delivered to each post by a strip of thick aluminum foil, which is connected to the post by a large, spring-loaded clamp of the kind that normally clips together sheets of paper. Vertical holes drilled into the posts serve as ports through which one end of the tube is linked to a vacuum pump and the other end to a tank of helium. The pump is needed to remove air from the tube and to draw in the helium. The helium has two functions. In the segment of the tube between the oven and a post, where the vapor may not be abundant, the helium helps to conduct the electrical discharge. The helium also promotes the condensation of the vapors in the outer regions of the tube by colliding with vapor atoms and removing their energy.

Gosnell favors the alumina-silica bricks because they are easy to cut, but he suggests that other high-temperature confinement materials might be tested. The essential characteristic of any confinement is that when the furnace is heated, the temperature along the confined length of the tube should be as uniform as possible. Gosnell heats the furnace with a common electric heating element that can be purchased from a supplier of home electrical parts. The heating element is inserted into a quartz tube that runs through the oven, parallel to the lasing tube and about 25 millimeters away from it. (Closer spacing would invite arcing between the two tubes.) Insulating fiber is used to seal off the holes in the bricks where the tubes enter the oven.

The helium is pulled into the laser by an ordinary single-stage vacuum pump, but refrigerator compressors working in tandem might be adequate substitutes. A valve inserted into the hose connecting the laser and pump allows one to close off the pump. A mercury manometer is also inserted to monitor the helium pressure, which was kept at about two torr.

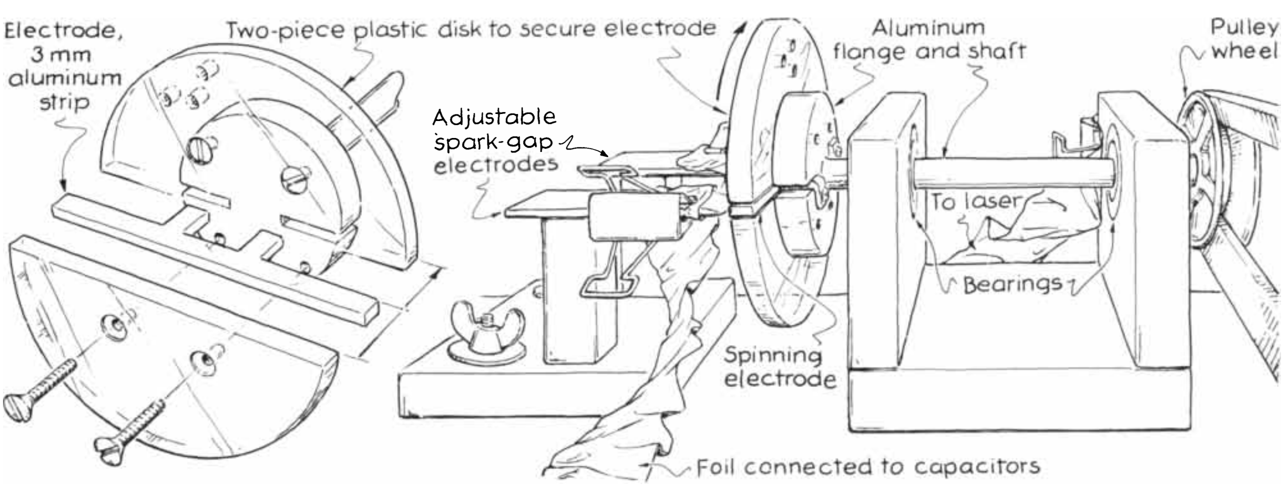
The reader will recall that the lasing action depends on a pair of closely timed discharges through the vapor. Rather than purchase expensive switches and a pulse generator, Gosnell built a mechanical switching device that he calls a pulser [*see illustration below*]. On one side of the device two bars serve as electrodes. They are separated by a short gap from a plastic disk, into which an aluminum strip is sunk. The disk is mounted on an aluminum shaft, which is rotated at about 6,000 revolutions per minute by a belt and motor at the opposite side of the device. Each bar electrode is connected to a charged capacitor, which is also connected to one of the brass posts supporting the laser tube. The aluminum strip and the shaft and its mount are electrically connected to the other brass support post.

As the disk rotates and the sunken strip approaches the tip of one of the bars, the capacitor connected to that bar discharges across the gap—and so also through the laser tube. Another discharge takes place when the strip approaches the other bar, which is connected to the other capacitor. The time between the discharges is set by the relative location of the bars and the rotational speed of the disk. The time should be about 150 microseconds, but the optimum value depends on the temperature of the copper vapor and related parameters.

Gosnell suggests that the strip should be flush with the disk face to eliminate the possibility of its catching on a bar's tip during rotation. He also advises that the pulley belt that connects the motor to the shaft should not be conducting, as many common belts are. (They are designed that way so that they bleed off any electrostatic buildup, but such a belt will short out the pulser.) The spacing between the tips of the bars and the disk is usually a few millimeters, but the optimum spacing can be determined only experimentally.

The power-supply circuit for the laser is shown in the illustration on the opposite page. At the left, plugged directly into the house electrical supply, are two identical neon-sign transformers rated at 15 kilovolts AC at 60 milliamps. (One transformer may well provide enough current.) The current from the transformers is rectified by a series of high-voltage diodes and fed to two storage capacitors, C1 and C2. Positioned along the way are two optional capacitors that smooth the current supply and allow current to be drawn regardless of the particular phase of the alternating current. Gosnell says the laser will fire well enough without these extra capacitors.

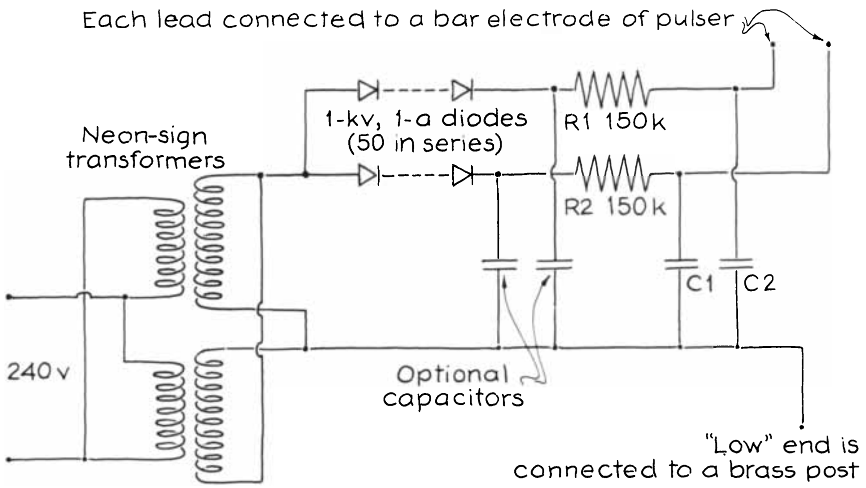
The resistance represented by R1 and R2 in the illustration consists of 1,500-ohm resistors rated at two watts and connected in a series-parallel combination. The C1 and C2 capacitors are made of flat, alternat-



Construction of the pulser, with details of the rotating electrode at the left

ing layers of aluminum foil and plastic sheets and have a face area of about 1,500 square centimeters; Gosnell used polyester or polyethylene for the plastic component to attain a capacitance of about 15 nanofarads.

In Gosnell's setup the capacitors are placed on a support just above the laser in order to minimize electrical problems that greater distance would create. (Each of the optional capacitors was similarly constructed of foil and plastic sheets but then was rolled up and inserted into PVC pipe, one meter long and 10 centimeters in diameter, that was mounted below the laser.) The pulser sits on a rigid, insulating piece of thick plastic just above the capacitors. The strips of thick aluminum foil that connect the capacitors, pulser and laser posts are all about five centimeters wide. The inductance of the capacitors and of the circuit between them and the laser tube must be low so that the current in the discharge increases sharply, dissociating the copper chloride molecules and exciting the atoms abruptly.



A circuit diagram of the power supply

To guard against accidental rupture of the rotating elements of the pulser, Gosnell erected a thick plastic shield in front of the pulser. To decrease the danger of electric shock, he connected high-voltage "bleeder" resistors across each capacitor to drain their charge when the system was turned off. (Here again I must warn of the danger implicit in the lethal currents that run through the power supply and laser, which can cause trouble if a charged capacitor is touched even after the system is turned off.)

The current for the heating element was controlled by a Variac, a variable transformer. For an alternative control, Gosnell connected a second heating element outside the furnace to the internal one and then attached one of the leads from the electrical source to the second element with an alligator clamp. By varying the location of the lead along the second heating element, he could control how much of the second element was in the circuit, thereby varying the resistance in the circuit and consequently the heat within the oven. With either technique of control, he usually heated the furnace to 390 degrees C, as read with a thermocouple he placed within it.

To align the mirror on the laser tube so that it reflected directly back along the tube, Gosnell sighted through the opposite end from a distance of about a meter while an assistant adjusted the mounting screws on the mirror platter. When Gosnell spotted a reflection of his eye at the center of the mirror, the alignment was correct. (It should go without saying that he never looked into the laser when it was firing. Laser bursts can cause severe damage to the retina.)

To check for an air leak in the tube, Gosnell disconnected the electric circuit and then connected a neon-sign transformer between the brass posts. After pumping down the tube and flushing it with helium several times, he filled it with helium to a pressure of about 10 torr and plugged in the transformer. When the system was free of leaks, the discharge through the tube was whitish gray; a pink tint indicated a leak.

Gosnell prepared his copper chloride by heating about 1/4 teaspoon (roughly one milliliter) of the crystals in a chemical hood to yield a greenish-brown liquid. (The hood is mandatory because breathing the vapor is harmful.) After the material cooled and solidified, he hammered it into a fine powder, which he sealed in a desiccating container. Both the heating and the desiccation serve to remove water and excess halide.

When he was ready to operate the laser, he put the powder into the central part of the tube through an opened end with a long, thin "spoon" he had fashioned. (He points out that the transfer would be easier if the tube were outfitted with a vertical section through which the powder could be poured. The extra section would extend out of the furnace and could be sealed with a rubber bung.)

The furnace was heated for about an hour before laser operation to stabilize the temperature. During that period Gosnell pumped down the system and flushed it with helium several times. After a final check on the pressure and the spacing between the bar electrodes and the plastic disk in the pulser, he turned on the motor that drives the pulser. In his initial trials the bursts of laser light were not continuous, but some experimentation with the gas pressure, the furnace temperature and the locations of the bar electrodes in the pulser eventually produced a more reliable output of bright green and yellow light on a card placed in the beam.

Gosnell, ever modest, suggests that someone who is particularly skilled in the construction of homemade lasers might well improve on his design.

FURTHER READING

EFFICIENT PULSED GAS DISCHARGE LASERS. W. T. Walter, N. Solimene, M. Piltch and G. Gould in *IEEE Journal of Quantum Electronics*, Vol. QE-2, No. 9, pages 474–479; September, 1966.

DOUBLE-DISCHARGE COPPER VAPOR LASER WITH COPPER CHLORIDE AS A LASANT. C. J. Chen, N. M. Nerheim and G. R. Russell in *Applied Physics Letters*, Vol. 23, No. 9, pages 514–515; November 1, 1973.

A PARAMETRIC STUDY OF THE COPPER CHLORIDE LASER. Noble M. Nerheim in *Journal of Applied Physics*, Vol. 48, No. 3, pages 1186–1190; March, 1977.

RESONANCE RADIATION TRAPPING EFFECTS IN COPPER AND MANGANESE LASERS. K. Srigouri, S. Ramaprabhu and T. A. Prasada Rao in *Journal of Applied Physics*, Vol. 61, No. 3, pages 859–863; February 1, 1987.

BOOKS

Exposé treatment confounds understanding of a serious public-health issue

by M. Granger Morgan

CURRENTS OF DEATH: POWER LINES, COMPUTER TERMINALS, AND THE ATTEMPT TO COVER UP THEIR THREAT TO YOUR HEALTH, by Paul Brodeur. Simon and Schuster, 1989 ($19.95).

The dust-jacket notes provide the following summary of *Currents of Death:* "Told in the form of a riveting medical detective story," the book "is a meticulously researched and dramatically written warning about the most pervasive—and covered up—public health hazard Americans face: the pernicious effect of our continuous exposure to low-level alternating-current electromagnetic fields."

The blurb is correct on at least one score. *New Yorker* staff writer Paul Brodeur has produced a dramatic and riveting story, one that should sell lots of books. The assertions about cover-up, pernicious effects and meticulous research stand on shakier ground. Brodeur tends to impute bad faith and an effort to cover up to any individual or institution that disagrees with his point of view. Reporting on very complex science, he cites findings selectively; in his effort to establish the pernicious effects of low-frequency electromagnetic fields, he thus gets far out ahead of the evidence. And his research, while exhaustive, is not always meticulous.

Before I consider these transgressions in detail, let me summarize Brodeur's basic argument. A handful of intrepid scientists discover that exposure to low-frequency electric and magnetic fields poses a grave risk to public health. The evidence is clear and unambiguous to the few honest scientists who will take the trouble to look. These honest few try to warn the American people of the grave dangers they face, but their efforts are blocked by a massive conspiracy of silence, denial and obfuscation.

M. GRANGER MORGAN is professor of engineering and public policy and of electrical and computer engineering at Carnegie-Mellon University, where he heads the department of engineering and public policy.

At one time or another, participants in this cover-up include a variety of branches of the U.S. government, including the Environmental Protection Agency, the National Institute for Occupational Safety and Health, the Occupational Safety and Health Administration, the Food and Drug Administration, the Office of Management and Budget, the Army, the Navy and the Air Force; several of the Department of Energy's national laboratories; health, safety and regulatory agencies in New York, Florida and other states; provincial and federal agencies in Canada; professional groups such as the American Medical Association; the publishers and editors of many of the nation's newspapers; *Scientific American;* most of the power companies of North America; IBM, Bell Laboratories and many other private firms; scores of bought-off university scientists; and various international groups such as the World Health Organization.

For 20 years the battle rages. At last the evidence becomes overwhelming. Chinks in the official stone wall open too wide to be plugged. A few valiant reporters get the story out. The conspiracy begins to disintegrate.

Will anyone take Brodeur's book seriously? Yes, they have and they will. There is rapidly growing public concern in the U.S. and a number of other developed countries about possible health risks from the weak low-frequency electric and magnetic fields associated with high-voltage transmission lines, lower-voltage distribution lines, radar and communications systems that involve low-frequency modulation, appliances such as electric blankets and workplace equipment such as video display terminals (VDT's). Although it has been out only a few months, Brodeur's book is already shaping the way many concerned, technologically untrained citizens, including some decision makers, frame and think about these issues.

The truth is that over the past two decades scientists *have* learned that low-frequency electric and magnetic fields can produce a variety of effects in living systems, including people. Whether one or several of these biological effects can in turn lead to adverse health consequences such as cancer, birth defects or depression—and in fact do so with great enough frequency to constitute a significant risk to public health—is not clear. There is indeed evidence, however, suggesting that they might.

Unfortunately, *Currents of Death* deliberately oversimplifies and misrepresents the complexity of the scientific process and the evidence it has produced. The book cites portions of this evidence in a highly selective fashion. It mixes and confuses adversarial legal and regulatory proceedings with honest scientific debate and portrays the bulk of the scientific community as bought-off hacks worried about preserving their funding. As a result, the book will considerably complicate and confuse public understanding, needlessly divert and confound the development of appropriate public policy, and provide exactly the evidence hard-line industry lawyers and publicists need to dismiss Brodeur's entire story out of hand as ravings from an extremist fringe—thus further polarizing the argument.

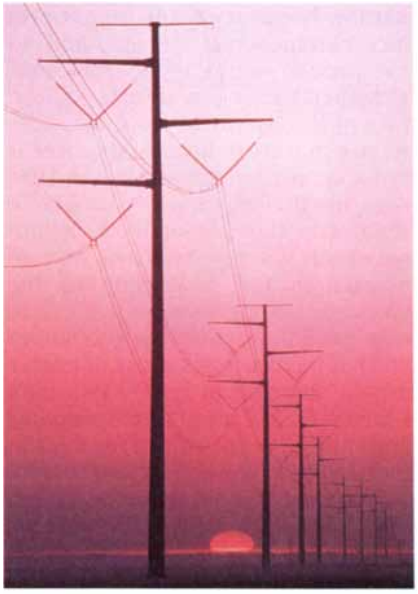
Perhaps it is best to start with fundamentals. In his classic 1962 book-length essay *The Structure of Scientific Revolutions,* Thomas S. Kuhn observed that science does not always progress smoothly through the continuous incremental accretion of new knowledge. Rather, he argued, major changes in scientific thinking typically involve a revolutionary shift in paradigm, or scientific "worldview." Such revolutions in thinking are not confined to epic events, such as the transition from classical to quantum mechanics. Kuhn argues that smaller-scale revolutions in thought occur regularly, if somewhat unpredictably, as a natural part of building scientific understanding in any field.

The past two decades have witnessed such a local paradigm shift in scientific thinking about the interaction of low-frequency electric and magnetic fields with biological systems. Twenty years ago all but a handful of scientists would have argued that weak low-frequency fields could not possibly produce significant effects in living systems. The thermal energy such fields deposit in tissue through what is called joule heating is much smaller than the thermal energy released by metabolic processes. Unlike ionizing radiation, the quantum energy contained in low-frequency fields is much too small to break chemical or nuclear bonds. Finally, indigenous fields in biological systems are often much larger than induced fields from external sources such as 60-hertz (cycles per second) electric power. For all these reasons, the conventional wisdom maintained, there should be no significant effects.

It was in the face of this dominant "paradigm" that evidence suggesting there might be effects from mechanisms other than the discounted ones began slowly to accumulate. Not surprisingly, this evidence met with considerable skepticism in established scientific circles. The skepticism was reinforced by the fact that many early studies were done on a shoestring—often with small numbers of experimental subjects, so that questions about statistical significance arose, and often with experimental designs that controlled inadequately for potential confounding effects.

Most environmental agents, such as ionizing radiation and chemical pollutants, produce effects whose magnitude increases in some steady (although perhaps nonlinear) way with their concentration in the environment. For virtually all such agents—air pollutants, toxins in drinking water and so on—scientists can safely assume that "if it's bad, more is worse." Not unreasonably, in evaluating results from experiments on the possible biological effects of weak low-frequency electric and magnetic fields, investigators looked for similar patterns. Often they did not find them. Instead effects appeared and disappeared as what appeared to be very small—many said irrelevant—changes were made in experimental conditions. This, too, provoked considerable skepticism in established circles. Long experience has led scientists to think experimental results that are marginally significant and that come and go in apparently random ways are usually not real.

Today, although we still do not fully understand the mechanisms of interaction, a variety of laboratory experiments have shown unequivocally that cells, and in particular their membranes or the complex receptor molecules incorporated in those membranes, can be sensitive to even fairly weak coherent low-frequency fields. As the strength of the evidence for



Health hazard?

these effects has grown, so too has the number of first-class experimenters investigating them. Effects that have now been clearly demonstrated include modulation of the flow of ions across cell membranes, changes in DNA synthesis and RNA transcription, effects on the response of normal cells to signaling molecules (including hormones, neurotransmitters and growth factors) and effects on the kinetics of some cellular biochemical reactions.

Many findings display considerable complexity. Some responses are observed only at discrete "windows" in frequency and amplitude; other responses depend on the strength and orientation of the ambient DC magnetic field. There are transient responses that appear only briefly after exposure has begun or ended, and there are threshold responses that appear only after exposure to fields of a given magnitude, or exposure for a given duration, but do not appear to depend on field strength.

Most animal "screening studies" have failed to find statistically significant differences between animals exposed to fields and unexposed animals. There are important exceptions, however. For example, in rodents chronically exposed to fields, the usually elevated nocturnal levels of pineal melatonin, a hormone with a variety of roles, including one in the regulation of circadian rhythms, are depressed. Considerable attention has also been paid to a search for developmental abnormalities. Despite some suggestive findings in rodents and miniature swine, nothing definitive has been found for sinusoidal 60-hertz exposure (Brodeur's discussion of these studies fails to mention their limitations and problems). Yet exposure to weak low-frequency pulsed magnetic fields with a rapid "rise time" has been reported (by a number of laboratories, all using the same experimental protocol) to induce severe developmental abnormalities in chick embryos—although apparently not in all breeds.

Studies of people exposed to fairly strong fields in a special exposure room have reported effects on heart rate and on reaction time. There is evidence that some individuals are more sensitive than others to such effects. Similar individual variability has been seen in studies with nonhuman primates. Some of the effects appear to be more pronounced when the fields are cycled on and off rather than when exposure is continuous. A study of people sleeping under electric blankets has reported transient changes in the level of the hormone melatonin. A large study of the possible adverse effects of electric blankets on human pregnancy outcomes is now in progress. Earlier studies of a connection have been suggestive but have involved too few women to allow one to draw reliable conclusions (again, Brodeur reports these findings but not their limitations).

Two kinds of epidemiologic studies have looked for an association between cancer and exposure to 60-hertz fields. The first set examined death rates from various diseases among people employed in "electrically related" occupations and compared them with the death rates from those diseases in other people. The second set compared estimates of the magnetic-field exposures received by people diagnosed as having specific cancers with the estimated exposures received by similar people who did not have cancer. Several of the latter studies have involved residential exposures from power distribution lines.

A surprising number of studies of both kinds show statistically significant associations between chronic exposure to weak AC magnetic fields (several milligauss or more) and the incidence of a variety of cancers, including leukemia, breast cancer and brain cancer. Most relative risks reported have been in the range of 1.5 to three, but a few have been significantly higher. Although none of these studies is perfect, several have been carefully done. The results are inconclusive, but the persistent pattern they display is clearly troubling.

There is fairly persuasive evidence

that low-frequency fields cannot produce chromosomal damage, and so few people would argue that low-frequency fields can initiate cancer. If there is a causative association between exposure to fields and cancer, it probably involves promotion—a later stage in carcinogenesis. A variety of findings at the cellular level, including effects on gap junctions and an increase in the production of the enzyme ornithine decarboxylase, have been cited to support the promotion effect. If fields present a risk of cancer but the increase in risk is less than a factor of five to 10, epidemiology alone may never be able to resolve the uncertainty; large and expensive animal studies will probably be necessary. A few such studies are now beginning, but it will be some years before results are available.

This is as much background on the state of the science as space will allow. Readers interested in more detail may want to start with the review *Biological Effects of Power Frequency Electric and Magnetic Fields* (Office of Technology Assessment, U.S. Congress, report number OTA-BP-E-53, U.S. Government Printing Office, May, 1989), of which I was one of the authors.

With this background, let me return to explore several aspects of Brodeur's book, paying particular attention to his discussions of power-frequency fields. The dust-jacket notes describe the book as "meticulously researched." At one level that is correct. For example, the extended and very engaging description of Nancy Wertheimer's self-financed, careful and pathbreaking epidemiologic studies in Denver, which threads through the first several chapters of the book, is well done. It corresponds precisely with my own understanding of the events, based on both a reading of her papers and several extended conversations I have had with her over the years. Several of Brodeur's other reporting jobs are similarly accurate and well written, although some specific statements—such as the explanation of why Richard Phillips left Battelle to go to the Environmental Protection Agency—are factually wrong. Yet the big problems in this book arise not from errors of fact but from problems of interpretation.

The scientific community did *not* conspire to discredit Nancy Wertheimer's work. They simply did not believe it. The idea that chronic exposure to 60-hertz magnetic fields of a few milligauss could double a child's risk of leukemia struck most of them as too improbable to warrant serious attention. It is nonetheless true that a number of utility people and their expert witnesses, convinced that her findings could not be true and faced with getting transmission lines built, did take steps to discredit the work, particularly in such adversarial forums as court cases and site hearings.

Brodeur reports in great detail on testimony by expert witnesses in a variety of adversarial regulatory hearings and court cases. What he does not report is the vigorous intellectual exchanges that took place twice a year within the scientific research community, once in the fall at the annual public meeting of investigators whose research was sponsored by the Department of Energy and the Electric Power Research Institute (EPRI) and once in the spring at the annual public meeting of the Bioelectromagnetic Society (BEMS). Although the positions of major stake holders and the adversarial encounters in which they were engaged have occasionally colored aspects of these debates, rarely have they figured in a central way.

During the period covered by Brodeur's book, BEMS grew into a substantial professional society with a first-rate quarterly scientific journal that publishes approximately 400 pages of carefully refereed scientific papers every year. Although Brodeur usually cites secondary sources, many of the positive research findings he cites were published in that journal. So too were a large number of negative and contradictory findings that—presumably in the interest of sustaining his cover-up theory—he chooses not to mention. Indeed, Brodeur's information sources are remarkably selective. For example, he draws repeatedly on secondary reports of research results published in the monthly newsletter *Microwave News* but never mentions *Transmission/Distribution Health & Safety Report,* the other newsletter published in this field.

Reading *Currents of Death* is a peculiar experience for someone who has studied the scientific literature on the biological effects of low-frequency fields carefully. Like many literatures, this one is uneven. There are outstanding studies done with great care by very careful investigators. There are also poorly conceived or poorly executed studies done by people who do not exercise appropriate care. Brodeur's book treats them all alike. If they found a positive effect, they must be right. If they did not find a positive effect, usually they go unmentioned.

But those unmentioned negative

studies and the vigorous scientific debate at those annual scientific meetings are critically important parts of the story. They contributed substantially to the continuing doubts of the scientific community. They help to explain the remarkable complexity of the scientific picture that has begun to emerge—a complexity Brodeur does not portray. And they have played an important role in the paradigm shift that most investigators would now agree has occurred.

Yet another troubling aspect of the book is Brodeur's conviction that an investigator's funding source is critical in determining the reliability of his findings. He makes a particular point of casting doubt on the motives and reliability of the research programs sponsored by the utility-funded Electric Power Research Institute. In spite of some uneven early performances, which Brodeur reports in detail, in recent years EPRI has done a remarkably good job of walking its institutional tightrope to support an important, high-quality program of research. Today, with a research budget of approximately $6 million per year, EPRI is the largest single source of research support in the field. This is not how EPRI, or most investigators, would like it. Several years ago federal research budgets were seriously cut; support from the Environmental Protection Agency was eliminated. The cuts did not result from a conspiracy. They reflected the squeeze of the Reagan administration's civilian-budget trimming. Federal support through the Department of Energy has begun to recover, but it still has a long way to go.

Brodeur assumes throughout his discussion that strong fields are more hazardous than weak ones. If fields do pose health risks, this may not be a safe assumption across the range of field strengths that people commonly encounter. We do not yet know how to define dose. Evidence of resonant, nonlinear, threshold and transient responses suggests that reductions in field strengths by factors of two or three might not always result in lower risk. If "more may not be worse," then the safest bet is to look for strategies—such as wider transmission-line rights-of-way—that get people out of fields. Strategies that instead simply change the distribution of field strengths across exposed populations (by, for example, increasing the height of transmission lines above the ground) might not provide protection.

Given the ambiguity of the science,

what should be our current policy toward exposure reduction? Brodeur is apparently clearly convinced that fields are the "most pervasive... public health hazard Americans face." Since he does not seriously consider ambiguity about the existence of the risk or the unsolved problem of defining dose and since he adopts a traditional "more is worse" formulation, for him the question of what to do is a nonissue: the strength of fields to which people are exposed should be dramatically reduced. The strong implication is that this should be done more or less without regard to cost.

Public and scientific debate on such issues is complicated by the different starting assumptions of participants. Scientists tend to assume that something is not so until it is clearly demonstrated to be so. They often carry this view into discussions outside the realm of science. Public-health officials adopt a different standard, often taking action on the basis of suggestive but still inconclusive evidence. A few people, including Brodeur, adopt the view that given any suggestion of hazard we should assume an agent is risky until it is proved safe. Since science can not demonstrate absolute safety, this creates grounds for conflict and disagreement.

Despite these differences, most people who are grappling seriously with the public-policy question of what to do about fields find the choice to be less obvious than does Brodeur. The reason is that the complex and puzzling scientific evidence will not support firm conclusions about either the existence or the magnitude of risk. Anything we say about policy responses, beyond a plea for more research, must involve judgments and values.

In other writings I have sketched three possible alternative policy approaches:

- Conclude that there is not yet enough evidence to warrant any action.
- Conclude that there is some basis for concern. Adopt a position of "prudent avoidance," which means limiting exposures that can be avoided with small investments of money and effort while for the moment forgoing others.
- Conclude that we face a serious health problem and commit substantial time and money to an aggressive program of limiting field exposures while acknowledging that because of current scientific ambiguities, some or all of this effort may turn out to be wasted.

In choosing among these strategies, people must consider both the available evidence and their own values. Different attitudes toward risk can suggest different actions even to people who agree about the evidence.

Some honest people have concluded that the current scientific ambiguities about possible health risks are so large that no action, beyond more research, can be justified. They argue that the country should limit its investments in safety to dealing with demonstrated hazards whose reduction will surely yield health benefits.

Suppose, however, a decision maker is concerned, thinks the evidence points toward the possibility of some risk and feels something should be done. When we face such situations in our private lives, we exercise prudence. In public decision making we have more trouble being "prudent" about uncertainty. Public-risk management tends to treat things as either dangerous or safe, with no middle ground. The traditions of adversarial regulatory and legal decision making exacerbate this problem. It takes real guts for a regulator, lawyer, judge or legislator to adopt a "prudent avoidance" strategy—but a number have begun to do so.

How might prudent people manage their possible risks from low-frequency fields if they wished to? Not by tearing all the wiring out of their house or moving out to the end of the distribution line; in most cases such measures would be extreme. But surely people could put away their electrically heated bedding and go back to traditional beds and blankets. Small electric motors produce strong highly localized magnetic fields, and people could rearrange their home and work environments to avoid them. For example, a motor-driven electric clock could be moved from a bedside table to a dresser across the room or be replaced with a digital or travel clock. Similarly, a short-sighted person who works at a word processor with a video display terminal might get reading glasses and push the machine to the back of the desk.

If someone is buying a new home, it might be prudent to consider the location of distribution and transmission lines as one of many things to be considered. Still, it is important to remember that even if fields do pose a health risk, things such as traffic patterns in the streets and radon levels in the house are likely to be more important in determining the occupants' overall safety. Selling an existing home solely in order to move away from power lines or other sources of field exposure would, in most circumstances, go beyond prudence.

Until more is known, state regulators who wish to exercise prudence should limit most of their concern to new facilities. The reason is that even under the most pessimistic assumptions, it is hard in most cases to justify the costs of modifying old facilities. Regulators who want to exercise prudence should site new facilities in a way that tends to keep people out of electromagnetic fields.

Because only a few people live near transmission lines and many people live near distribution lines, the latter probably deserve more attention. Not enough has been done to learn how to reduce the fields associated with such lines. Putting them underground removes electric fields and, if it is done right, can reduce magnetic fields dramatically. Modified grounding practices can also reduce magnetic fields. In the interest of prudence, utilities and the home-building industry should begin now to explore and evaluate ways to reduce or eliminate fields stemming from distribution lines and from house wiring.

EPRI has recently decided to increase its support in this area. At least one electric-blanket manufacturer is already marketing a low-field blanket, and several computer manufacturers have taken steps to reduce the fields associated with video display terminals. Unless anticipatory engineering and economic studies are undertaken now, a lot of wasteful and ineffective effort can be expected in a scramble to take protective measures if and when experimental evidence someday clearly links fields to health risks.

Taking more drastic action than what is indicated by "prudent avoidance" would cost a lot of money, a good deal of which might be wasted. In order to justify aggressive action today, someone must be persuaded that fields pose a risk to human health and that the level of that risk is pretty high compared with the large number of other risks people face. In contrast to most decision makers, Brodeur clearly believes that both of these assumptions are valid.

In simplifying a complex problem by sweeping all complexity under the rug of cover-up, and by failing to discuss the possibility that in the face of present evidence reasonable and honest people may reach very different conclusions, Paul Brodeur's *Currents of Death* has done a disservice to the public interest he presumes to champion.

ESSAY

A new initiative to save the planet

by Al Gore

The world is in the midst of an environmental crisis beyond anything yet experienced; unless radical steps are taken to safeguard the planet, our future on it cannot be secured. In response to the crisis, I propose that the U.S. launch a Strategic Environment Initiative (SEI), similar in scope and funding to the Strategic Defense Initiative (SDI).

The need for such an initiative is urgent: no longer is the threat of nuclear war at the top of the world agenda. Instead global deterioration has become the primary concern of many of the world's leaders. With that in mind, it is imperative that we approach environmental protection and sustainable economic growth with the same intensity as we approached SDI—and with comparable or greater levels of funding.

Recognition of the global environmental crisis now on us is not new. It was 20 years ago, on April 22, 1970, that millions of Americans turned their attention to the first Earth Day. In the end, the energy and enthusiasm generated on that day had a lasting effect on our nation. Among other things, Earth Day was the catalyst that led to the creation of the Environmental Protection Agency and the passage of the Clean Air Act and Clean Water Act. Today, as we enter a new decade, the American people are again coming together to demonstrate their concern for the environment. And the U.S. government is again responding to its citizens' demands: the EPA is about to become a cabinet-level agency and the U.S. will soon have an Environment Secretary. More effective environmental legislation is pending. Yet the world is facing novel environmental problems that were unforeseen 20 years ago and serve to underscore the urgency of SEI.

The world's forests are being destroyed at the rate of one acre per second. An enormous hole has appeared in the stratospheric ozone layer. Species are dying at an unprecedented rate: 1,000 times faster than at any time in the past 65 million years. Chemical wastes are seeping down to poison groundwater, while huge quantities of carbon dioxide, methane and chlorofluorocarbons are entering the atmosphere to threaten the planet with climate change.

From the time of Christopher Columbus to the beginning of this century, world population tripled—to 1.6 billion. In only the past 75 years it tripled again—to 5.2 billion. We are told that in the next 75 years it will double and perhaps even triple again. Nearly every aspect of modern industrialized society follows the same pattern of sudden, unprecedented acceleration.

Americans consumed about 40 billion kilowatt-hours of electricity in 1914; last year we consumed 60 times that amount. The production of organic chemicals has gone from almost nothing to 225 billion pounds per year. The consumption of fossil fuels has increased by a factor of 10, producing a corresponding flux of carbon dioxide into the atmosphere. In the U.S., emissions of oxides of nitrogen have increased by 800 percent since 1914 and methane concentrations have nearly doubled. The world's emissions of chlorofluorocarbons (substances not invented until the 1930's) have multiplied 80 times since World War II and are jeopardizing the earth's protective ozone shield.

These dramatic changes are taking place not only because the human population is surging and the environmental impact of our economic activities has increased but also because we have been willing to tolerate environmental vandalism on a global scale. To avert disaster, we must recognize that our survival is inextricably tied to the fate of the earth. Indeed, the environment has become a question of national security: it directly affects the interests of all nations and the welfare of all peoples.

To secure our planet, we need to deal comprehensively with global warming, stratospheric ozone depletion, species loss, deforestation, ocean pollution, acid rain, air and water pollution and other threats. In every major sector of the economy, we must find and disseminate new technologies.

Developed countries need to balance the imperatives of growth with the need for environmental management. Developing countries, where the problems of reconciling economic growth with environmental protection far exceed those of the developed world, face even greater challenges. Yet the cooperation of developing nations, whose emissions of greenhouse gases may increase to from 20 to 60 percent of the world total by 2050, is crucial in dealing with global climate change and stratospheric ozone depletion.

To promote such cooperation requires the mobilization of talent and resources on a scale ordinarily reserved for national defense. Such mobilization is possible only if we are willing to commit ourselves to a Strategic Environment Initiative. By exporting new technologies to developing countries, where 95 percent of the world's population growth is expected to take place in the next century, SEI could have a major impact on moving the planet closer to sustainability. Here in the U.S., SEI could have a significant influence on every sector of the economy, promoting such technological advances as more fuel-efficient cars and less damaging methods of agriculture.

A Strategic Environment Initiative is far from all that is needed. We must also learn once again to see ourselves as a part of the vast, interconnecting ecological system that sustains us and regain our sense of place in the natural world. In short, we must regain our eco-librium. To do so, we must transform global politics, shifting our vision from short-term concerns to long-term goals, from conflict to cooperation.

Slowly we are beginning to take the right steps toward those goals. Earlier this year scientists, policymakers and religious leaders gathered in Moscow to address global environmental concerns. Soviet President Mikhail S. Gorbachev delivered a message of environmental awareness to the gathering and in so doing demonstrated an understanding and sensitivity rarely seen among world leaders. Later this month I will chair the first Interparliamentary Conference on the Global Environment in Washington, D.C., which will gather together representatives from around the world to craft the policies necessary to slow our planet's destruction.

In thinking about Earth Day 1990, it will not be enough to change our laws, policies and programs. The solutions we seek must stem from a new faith in the future, a faith that justifies sacrifices in the present, and from a new courage to choose higher values in the conduct of human affairs. We must also display new reverence for our place in the natural world. In the words of the environmental philosopher Ivan Illich, it is time to stop running from the "shadow our future throws."

AL GORE is U.S. senator from Tennessee and chairman of the Subcommittee on Science, Technology and Space of the U.S. Senate Commerce Committee.

This will help you spend money overseas.

This will help you save money overseas.

Calling home from overseas can be a lot less expensive when you use *AT&T USADirect®* Service. Just dial the ***USADirect*** number for the country you're calling from, and you'll be connected to an *AT&T Operator* in the U.S. within seconds. It's fast. It's easy. It's available in over 50 countries. And it helps you minimize hotel surcharges. You save with AT&T's economical international rates, whether you use your *AT&T Card* or call collect. For your information card, call **1-800-874-4000, Ext. 301.**

AT&T
The right choice.

Earth Day 1990: General Motors Marks 20 Years of Environmental Progress

At General Motors, we recognize the effects that cars and their manufacture have on the environment. We understand the relationship better than any other carmaker in the world.

Science, technology, and engineering are the tools of our business. In recent decades, GM has brought all these resources to bear on the environment, to the benefit of the planet and our current and our future customers.

- **Since clean air became a national goal, we have substantially reduced exhaust emissions from our cars and trucks—hydrocarbons and carbon monoxide are down by 96%, nitrogen oxides by 76%.**
- **The fuel economy of GM cars has increased 130% since the mid-'70s—average weight has been reduced by 27%, aerodynamic efficiency improved by one-third, and the rolling resistance of tires cut by half.**
- **GM's Impact concept car is the highest-performing electric vehicle in existence, and it is the first with the potential to meet the safety standards for a production vehicle.**
- **Our solar-powered Sunraycer won the 1,950-mile World Solar Challenge in Australia, and we are sponsoring Sunrayce USA to further academic investigation into solar energy.**
- **GM sponsored the Methanol Marathon to promote study of alternative fuels in colleges and universities in the U.S. and Canada; we are currently operating a test fleet of methanol-powered cars in California.**

Today's GM cars and trucks are far more compatible with the environment than their predecessors were twenty years ago. Throughout the corporation, research is under way to make our next generation of vehicles even better: to increase battery power and endurance, to improve catalysts, to reduce weight and raise fuel efficiency.

Our dedication to environmental improvement has been applied with similar success to the process of building GM cars and trucks. We have reduced plant emissions and made more efficient use of energy in all of our facilities.

- **We are using electrical energy more efficiently in our plants—since 1972, we've reduced usage by 44% per vehicle produced.**
- **The use of new paints has lowered hydrocarbon emissions at our assembly plants by 70% since 1977.**
- **GM is committed to using non-CFC air conditioners in all our vehicles by 1996, and we are eliminating CFCs from manufacturing processes as quickly as possible.**
- **GM scientists have developed a way to eliminate a hazardous solid waste from the making of a type of high-strength iron—we've already applied this process to 70% of our production.**
- **Our new Saturn plant in Spring Hill, Tennessee, has been acknowledged as an exemplary development of a greenfield site.**

Throughout General Motors, people are using their talents to advance our understanding of the complex interaction between man and nature, to identify problems and develop solutions.

In 1989, GM scientists were honored for their work on air quality by the American Meteorological Society. This year, people throughout the corporation—in our plants and offices as well as in our labs—are showing their support for Earth Day 1990. It is a day for gaining some measure of how far we have come, and how far we have yet to go.

GM

MARK OF EXCELLENCE

Chevrolet Pontiac Oldsmobile Buick Cadillac GMC Truck