MARCH 1990
$2.95

SCIENTIFIC AMERICAN

TRENDS IN COMMUNICATIONS: *The Road to the Global Village.*

Earthquakes: should the Midwest brace for a big one?

Physicists predict—and find—new radioactivities.

Observatories on the moon: *a bold proposal for a practical lunar scientific outpost.*

Lumina Sedan. What's Become of

You're about to see a big difference in running the family chauffeur service. It's a smooth and easy, more comfortable experience than what you're probably used to. And you feel more confidence. You could even call it a big pleasure, now that Lumina's here.

Easier.

Lumina's front-wheel drive and power steering make driving easy come, easy go. The rear doors open wide, so getting in and out's easier. And as a bonus, only Chevrolet gives you standard Scotchgard™ Fabric Protector on seats and doors,* so spills are easy as pie to clean up.

Smooth.

Four-wheel independent suspension helps keep a family running with smooth stability. (Which pleases everyone. Including the dog.) Four-wheel power disc brakes and all-season radials add to that feeling of security, especially in bad weather.

Bigger, More Comfortable.

Six members of the family can enjoy Lumina's large capacity for comfort. And its full-suspension seats provide even more comfort. Its large 15.7-cu.-ft. trunk, that's exactly the height of a grocery bag, is a very convenient feature. It holds enough to keep one family — or teen — fed for a week. Creature comforts include a quad-speaker stereo and dual vanity visor mirrors.

GM® *All Chevrolet cars except Geo Metro.™ †Except tires which are covered by their manufacturer. See your Chevrolet dealer for terms of this limited warranty. A deductible will apply after 12 months/12,000 miles. Chevrolet, the Chevrolet emblem and Lumina are registered trademarks of the GM Corp. Let's get it together...buckle up.

the Family Car.

More Confidence.

Each Lumina has such high quality, we back everything with a 3-year/50,000-mile Bumper to Bumper Plus Warranty.† And its corrosion protection (using two-sided galvanized steel and base-coat/clear-coat paint) is so effective, we back it for 6 years or 100,000 miles against body-rust-through.

Now that you know what the family car has become, we invite you to test drive a new Lumina Sedan. Because, when it comes to new ideas for your family, nobody's winning like The Heartbeat of America.

THE Heartbeat OF AMERICA IS WINNING.
TODAY'S CHEVROLET

Every age has sought perfect sound.
In the age of telecommunications, we're achieving it with our fiber optics network.

CONTEL® We go out of our way.℠

SCIENTIFIC AMERICAN

March 1990 Volume 262 Number 3

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 yearfor 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

THE MING DYNASTY AND THE SPACE STATION

In 1492 China discovered Europe.

This is the way history might have been written but for a sudden and fateful change of policy on the part of the Ming Emperors.

In the early 15th century China was the most advanced power on Earth. The Middle Kingdom led the world in such decisive technologies as gunpowder, printing, metallurgy, engineering, and medicine.

Navigation was also in China's strong suit. From 1405 to 1433 vast Chinese fleets, which would dwarf much later European expeditions, were already exploring and opening up trade routes westward. By 1431

View of the Imperial Court, Beijing, Anon. 17th Cent., Bibliotheque Nationale, Paris

a fleet of sixty-two ships and 28,000 men had reached the east coast of Africa. It was only a matter of time before the Chinese discovered Europe.

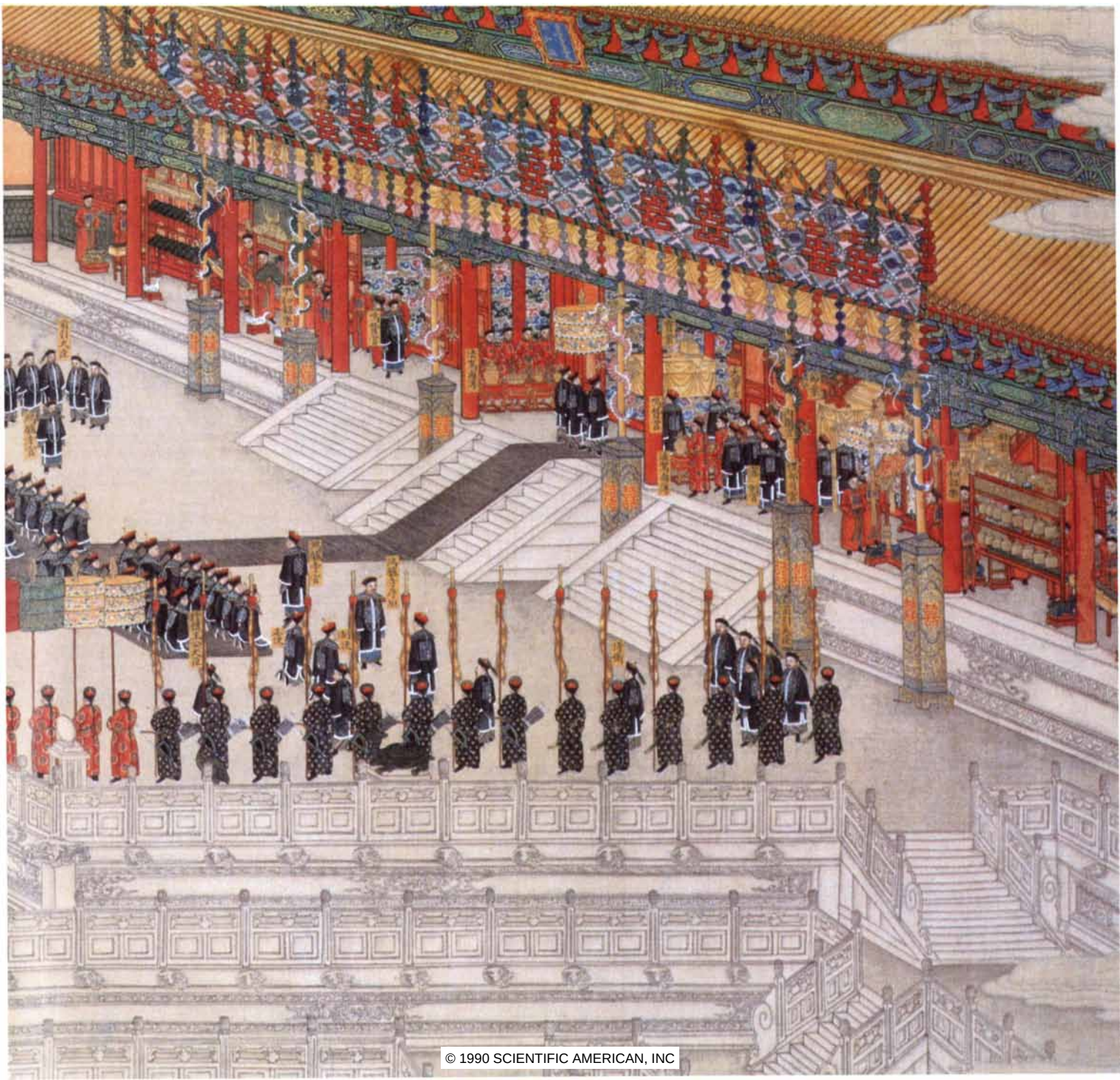
Then, in a sudden policy reversal, the Ming Emperors halted all further voyages and began to foster an atmosphere of xenophobic conservatism. Science and technology decayed. Trade became passive. And the next five centuries saw China become one of the world's exploited nations.

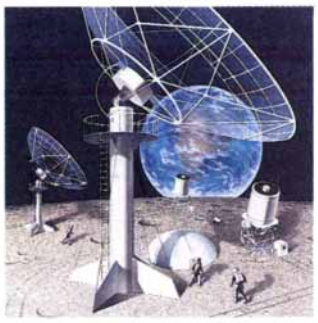
The United States now also finds itself at a critical time in history. So many domestic and defense needs make investment in space exploration seem, by comparison, of lower priority. The funding of our orbiting space station, indeed, our whole space program, seems to be politically in doubt.

One thing that is inevitable, though, is that humans will continue to reach out into space. It is in our nature. And whether space is explored by the United States or the Soviet Union or Europe, it will be explored.

Had the Ming Emperors not hamstrung China five hundred years ago, history undoubtedly would have taken a dramatically different turn. But their moment was lost.

The direction the United States now takes into — or away from — space will have implications for our fate during the next five hundred years. Our moment is now.





THE COVER painting is an imagined view of a manned astronomical observatory on the moon, an almost ideal site from which to examine the universe (see "Observatories on the Moon," by Jack O. Burns, Nebojsa Duric, G. Jeffrey Taylor and Stewart W. Johnson, page 42). The blue earth hovers over a scene dominated by two large, linked radio telescopes; to the right are two 1.5-meter optical telescopes, part of a high-resolution array. The dome at the center protects sensitive instruments from the impact of fine particles.

THE ILLUSTRATIONS

Cover painting by George V. Kelvin

Page	Source
43	George V. Kelvin
44–45	Andrew Christie
46	George V. Kelvin
47–48	Andrew Christie
51–56	George Retseck
58–60	Hank Iken
62–63	Michael Goodman
66	Walter Greiner
69	J. Roger Bowman, Australian National University
70–71	Ian Worpole
72	Ian Worpole (*top*), Arch C. Johnston (*bottom*)
73	Ian Worpole
74	Edward Bell (*top*), Ian Worpole (*bottom*)
75	Robert Gelinas and David Amick, Ebasco Service
83–84	Joe Lertola
85	The Bettmann Archive (*top*), *Huxley in Hollywood,* Harper & Row, Publishers, Inc., © 1989 (*middle*), Wayne Miller, Magnum Photos, Inc. (*bottom*)
88	Joe Lertola and Edward Bell
89	Jim Marchese and Edward Bell
90	Jon Brenneis
92	AT&T (*top*), Quesada/Burke (*bottom*)
93	AT&T
94	Joe Lertola and Edward Bell
97–101	Tom Prentiss
104–108	Patricia J. Wynne and Johnny Johnson
109	Johnny Johnson
111–112	Thomas C. Moore
113	Edward Bell
114	Nancy Field
115	Thomas C. Moore
118–119	Andrew Christie
120	Edward Bell
126	*Herreshoff of Bristol,* WoodenBoat Publications, © 1989

SCIENTIFIC AMERICAN®

Established 1845

EDITOR: Jonathan Piel

BOARD OF EDITORS: Armand Schwab, Jr., *Managing Editor;* Timothy Appenzeller, Laurie Burnham, *Associate Editors;* Timothy M. Beardsley; Elizabeth Corcoran; Deborah Erickson; John Horgan; June Kinoshita; Philip Morrison, *Book Editor;* Corey S. Powell; John Rennie; Philip E. Ross; Ricki L. Rusting; Russell Ruthen; Paul Wallich; Karen Wright

ART: Samuel L. Howard, *Art Director;* Edward Bell, Joan Starwood, *Associate Art Directors;* Johnny Johnson

COPY: Maria-Christina Keller, *Copy Chief;* Nancy L. Freireich; Michele S. Moise; Philip M. Yam

PRODUCTION: Richard Sasso, *Vice President Production and Distribution; Managers:* Carol Eisler, *Manufacturing and Distribution;* Carol Hansen, *Electronic Composition;* Madelyn Keyes, *Systems;* Leo J. Petruzzi, *Manufacturing and Makeup;* Carol Albert; William Sherman; Debra Townes

CIRCULATION: Bob Bruno, *Circulation Director;* Lorraine Terlecki, *Business Manager;* Cary Zel, *Promotion Manager*

ADVERTISING: Robert F. Gregory, *Advertising Director.* OFFICES: NEW YORK: Peter Fisch; John Grant; Meryle Lowenthal; William Lieberman, Inc. CHICAGO: 333 N. Michigan Avenue, Chicago, IL 60601; Patrick Bachler, *Advertising Manager;* Litt Clark, *Midwest Manager.* DETROIT: 3000 Town Center, Suite 1435, Southfield, MI 48075; William F. Moore, *Advertising Manager;* Edward A. Bartley, *Detroit Manager.* WEST COAST: 1650 Veteran Avenue, Suite 101, Los Angeles, CA 90024; Kate Dobson, *Advertising Manager;* Joan Berend, San Francisco. ATLANTA, BOCA RATON: Quenzer/Stites. CANADA: Fenn Company, Inc. DALLAS: Griffith Group.

ADVERTISING SERVICES: Laura Salant, *Sales Services Director;* Diane Greenberg, *Promotion Manager;* Ethel D. Little, *Advertising Coordinator*

INTERNATIONAL: EUROPE: Roy Edwards, *International Advertising Manager,* London; GWP, Düsseldorf. HONG KONG/SOUTHEAST ASIA: C. Cheney & Associates. SEOUL: Biscom, Inc. SINGAPORE: Cheney Tan Associates. TOKYO: Nikkei International Ltd.

BUSINESS MANAGER: Marie D'Alessandro

PUBLISHER: John J. Moeling, Jr.

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

PRESIDENT AND CHIEF EXECUTIVE OFFICER: Claus-Gerhard Firchow

EXECUTIVE COMMITTEE: Claus-G. Firchow; *Executive Vice President and Chief Financial Officer,* R. Vincent Barger; *Vice Presidents:* Linda Chaput, Jonathan Piel, Carol Snow

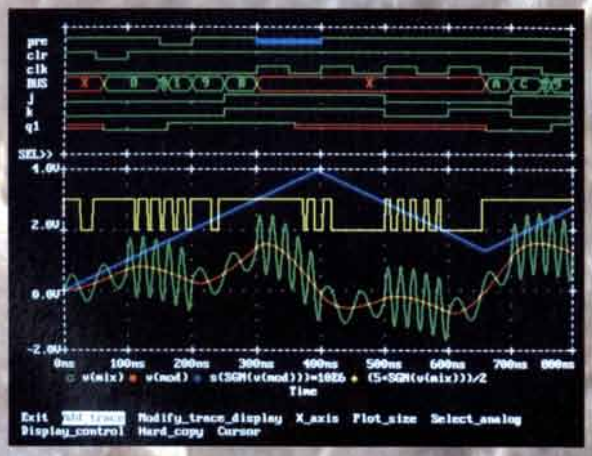
CHAIRMAN OF THE BOARD: Georg-Dieter von Holtzbrinck

CHAIRMAN EMERITUS: Gerard Piel

PSpice

MicroSim Corporation

The Standard for Circuit Simulation



Analog and digital waveforms with a common time axis

For further information call *toll free* at (800) 245-3022. Find out for yourself why PSpice has become the standard for circuit simulation.

PSpice and its options form an integrated package for the analysis of electronic and electrical circuits. The software enables you to create a "computer breadboard" of your circuit for testing and refinement *without* building hardware.

A wide variety of circuits can be analyzed using PSpice. These range from precision instrumentation amps, to microwave mixers, to switching power supplies, to digital control circuits. Our capability to analyze mixed analog/digital circuits is unique within the industry.

Introduced over six years ago, MicroSim's PSpice has more copies sold than all other SPICE based programs combined. PSpice has a reputation for extensive capabilities, accurate results, diverse options, and capabilities across a range of computer platforms. PSpice includes an extensive device library of 2500 analog parts and 1200 digital parts at no extra charge.

The PSpice options include:

Digital Simulation, which allows the simulation of mixed analog/digital circuits with feedback between the analog and digital sections.

Monte Carlo Analysis, which calculates the variations in a circuit's performance allowing for component tolerances. This option performs statistical analyses: Monte Carlo, Sensitivity, and Worst Case.

Analog Behavioral Modeling, which allows you to describe analog components, or entire circuit blocks, using a formula or a look-up table.

Probe, which acts as a "software oscilloscope" to provide an interactive viewing environment for simulation results (see photo).

Parts, which is a parameter extraction program allowing the extraction of device model parameters from data sheet information.

PSpice is available on these computer systems:

- The PC family, including the PS/2, running DOS, Protected Mode DOS, or OS/2.
- The Macintosh II family.
- The Sun 3, Sun 4, and SPARCstation families.
- The HP/Apollo DS2000, DS3000, and DS4000 series workstations.
- The DECstation 2100 and 3100.
- The VAX/VMS family, including the MicroVAX.

Each copy of PSpice comes with our extensive product support with over 100 years of experience in CAD/CAE. MicroSim believes that the finest support for the software occurs by utilizing the engineers who wrote it.

20 Fairbanks • Irvine, CA 92718 USA • FAX (714) 455-0554

3:32 PM

Jamie Conlon's first step proves NEC's laser technology can help solve infertility.

Jamie Elizabeth Conlon is living proof of the great strides medicine is making towards reducing America's rising infertility rate.

Twelve years before Jamie's birth, Claire Conlon experienced a series of medical difficulties that convinced her she would never bear children. But in October of 1987 she was told that Contact Laser™ surgery might correct her problem. Two months later she was pregnant.

This miracle was made possible through the help of an NEC laser. The laser's superior quality and high precision contributed greatly to the success of Mrs. Conlon's surgery. And its easy adaptability to non-medical applications has made it the most widely-used laser on the market today.

NEC's advanced laser technology is just one of the ways we're improving people's lives. You'll find the quality products of NEC in corporations large and small. In public and private institutions. In homes across the country. Helping people at every step along the way.

For more information, call us at 1-800-338-9549.

NEC

Quality That Proves Itself.

Communications • Computers • Electron Devices • Home Electronics

Contact Laser is a trademark of Surgical Laser Technologies, Inc.

LETTERS

To the Editors:

Contrary to Tim Beardsley's "Gaia" ["Science and the Citizen," SCIENTIFIC AMERICAN, December, 1989], Gaian science is entirely consistent with Darwinian (not necessarily neo-Darwinian) evolution. Beardsley misrepresents it. The Gaia hypothesis states: The surface sediments and the lower atmosphere are actively regulated by the biota (the sum of the living organisms) with respect to the abundance of reactive elements (for example, hydrogen, carbon, nitrogen, oxygen and sulfur), the acidity (hydrogen, hydroxyl, carbonate and bicarbonate ions), the oxidation-reduction state and the temperature. Beardsley didn't finish the thought: Gaian regulation, like the physiology of an embryo, is more homeorhetic than homeostatic in that the internally organized system regulates itself around moving, rather than externally fixed, set points. (That is, the values of surface temperature, atmospheric composition and other points around which there is Gaian regulation do change with time.) Gaia, a single enormous system deriving from common ancestors at least 3,500 million years old, is connected through time (by ancestry) and space (through atmospheric chemical signals, ocean currents and the like); the Gaian system persists in the face of changes (population expansions and extinction, sea-level changes and so forth). Variations in the ability of organisms to inject their genes into the next generation are one such change.

We suspect that even plate-tectonic movements and ocean salinity are biotically modulated. Consider this: Mars and the earth probably have the same internal heat sources, but whereas the surface of Mars is marked by immense canyons and mountains, the earth's surface is broken up into thick plates. At least since the beginning of the Phanerozoic eon, 600 million years ago, the earth has shed heat through lateral plate movements, while the ancient mountains and canyons of Mars have simply grown higher and deeper. The earth's huge carbonate reserves and oxygen-rich atmosphere—the products of life—may have played a role in the crustal movements. Furthermore, the ocean is undersaturated with respect to many potentially toxic ions (for example, sodium, manganese and magnesium). Thus, changes in the composition of seawater may also be intimately related to biotic regulation. These Gaian suspicions suggest programs for scientific research that would never be entertained under a strictly neo-Darwinian worldview such as that espoused by Tim Beardsley.

LYNN MARGULIS

JOHN F. STOLZ

University of Massachusetts, Amherst

Tim Beardsley responds:

First, Dr. Margulis did not mention "homeorhesis" during our several hours of discussion about Gaia, and the term does not appear in either of the books by James E. Lovelock, the originator of the Gaia hypothesis. The term homeorhesis was coined by the British biologist Conrad H. Waddington to describe the phenomenon underlying the alternative developmental pathways (resulting in stereotyped aberrant forms) that are seen in embryos exposed to various experimental treatments. He was able to infer the existence of homeorhesis only by manipulating many embryos. Given that there is only one earth that we are manipulating (and no control), I wonder at Dr. Margulis's confidence that the earth's biota constitutes a homeorhetic system. Second, biogeochemical dynamics has emerged as an important area of research without having to appeal to Gaia or homeorhesis: see, for example, "Our Changing Planet: The FY 1990 Research Plan," a multiagency document that describes the U.S. global-change research program.

To the Editors:

"The First Color Photographs," by Grant B. Romer and Jeannette Delamoir [SCIENTIFIC AMERICAN, December, 1989], is as delightful as it is instructive. It is particularly nostalgic for those of us who had a passion for color photography in the 1940's. At that time, it seems to me, one had far more choice than now, for all sorts of additive color processes were still available: one could use Lumière color, with its brilliantly dyed starch grains, or Dufay, with its brilliantly colored grids (it was a special delight to look at these films under the microscope and *see* the red, green and violet starch grains or grids). Now, nearly half a century later, these old additive films have retained all their brilliance of color, while the tripacks of that era have faded.

The registration problems of a separate du Hauron color grid and film, which Romer and Delamoir speak of, were ingeniously solved in the then popular Finlay color. Lining up black-and-white photographs with a Finlay screen, or putting the three separation positives in a color camera or a chromoscope (which were common and easy to get in antique shops), one got a sudden and intense sense of the *creation* of color—as also when one tilted the Lippmann plate to the right angle. It was precisely such a magical creation of color in a photograph, albeit of a very special kind, that stunned Edwin Land when he saw it and led him to his radical theory of color vision [see "The Retinex Theory of Color Vision," by Edwin H. Land; SCIENTIFIC AMERICAN, December, 1977].

In the 1940's a passion for color photography *demanded* a knowledge of its history and a delight not only in its results but also in all its concepts and mechanisms. Is there some way of bringing back this sense of fun, of discovery, to color photography today?

OLIVER SACKS

City Island, N.Y.

To the Editors:

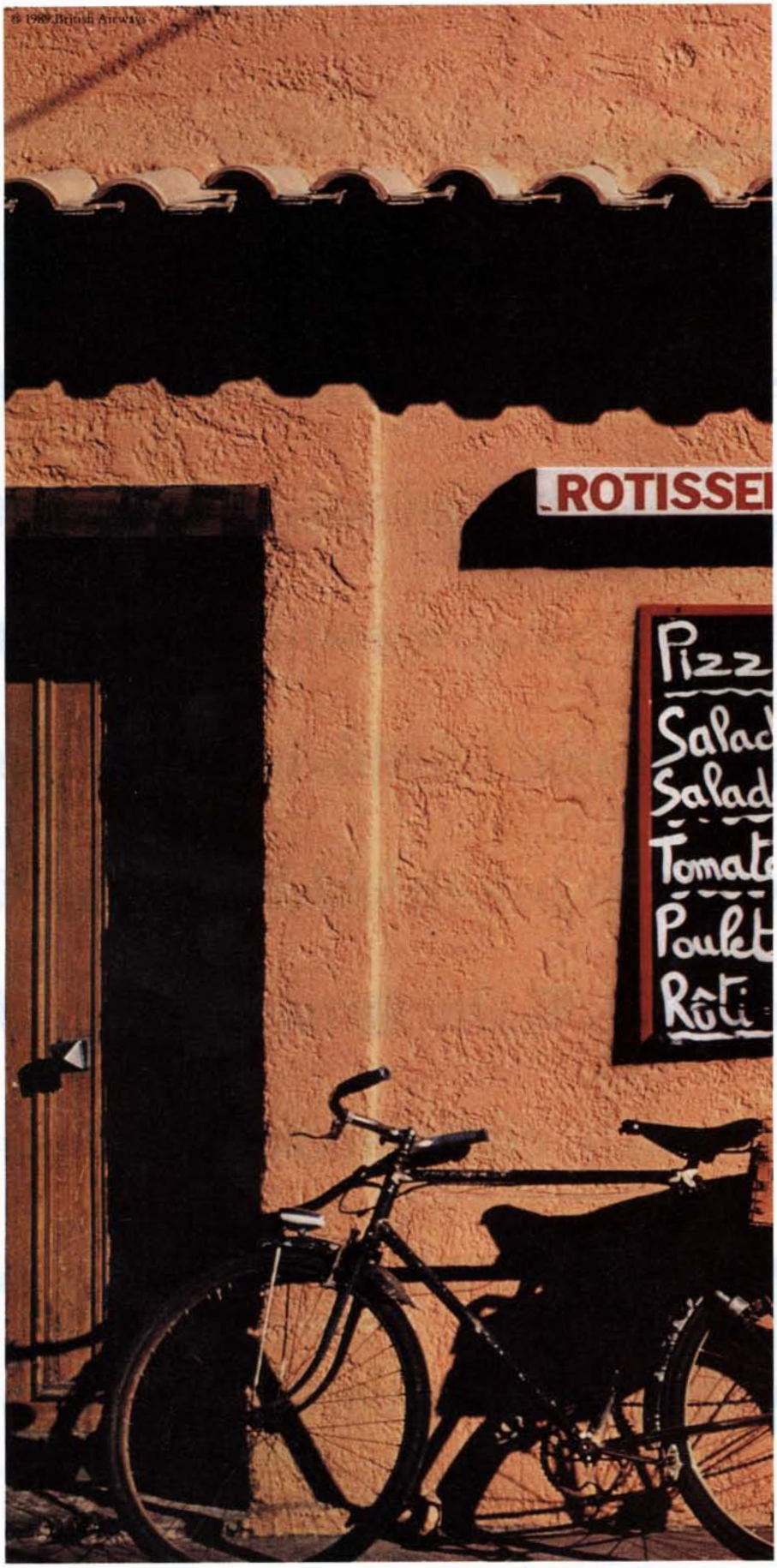
As a result of last-minute editorial changes, the map of Mithraic archaeological sites on page 82 of my article, "The Mithraic Mysteries" [SCIENTIFIC AMERICAN, December, 1989], highlighted several locations where the evidence for Mithraism is in fact dubious. For the record, the altar found at Thessalonike is probably not Mithraic, and the evidence for Mithraism at Vichy, Eleusis, Philippi and Alexandria is unconfirmed or disputed; the evidence for Mithraism at Athens consists solely of a few brief inscriptions. In addition, some of the other highlighted locations are of only minor importance. The interested reader will find complete information in Maarten J. Vermaseren's *Corpus Inscriptionum et Monumentorum Religionis Mithriacae* (two volumes, The Hague: Martinus Nijhoff, 1956, 1960).

DAVID ULANSEY

To the Editors:

My recent article, "Microclusters" [SCIENTIFIC AMERICAN, December, 1989], mistakenly attributed the structure proposed for the 45-atom silicon cluster to Richard E. Smalley of Rice University. Actually this structure was proposed by Efthimios E. Kaxiras of the Naval Research Laboratories in Washington, D.C.

MICHAEL A. DUNCAN



COME BACK TO EUROPE AND SEE IT FOR THE FIRST TIME.

British Airways Holidays introduces Preferred Vacations.™ Uncommon holidays for experienced travellers.

You'll do things that, until now, only insiders got to do. Go to backstage parties at some of England's most exciting theatres. Visit tiny antique shops that most people don't know exist. Relax at one of Britain's outstanding country house hotels. Take an exotic excursion to Istanbul.

And to ensure that your vacation runs smoothly, we provide you with certain amenities. Like our London Help-Line, which answers any and all questions you may have along the way, a comprehensive insurance package and a toll-free number that puts you in touch with legal or medical experts 24 hours a day.

Send for your Preferred Vacations Brochure by clipping this coupon, call British Airways at 1-800-876-2200, or see your travel agent. And look forward to seeing Europe in a whole new light.

SCTA 03/0090

British Airways Holidays, P.O. Box 7707, Woodside, NY 11377

Please send me British Airways Holidays Preferred Vacations™ Brochure.

☐ I would also like information about your Privileged Traveller program for those age 60 and over.

Do you currently have a valid U.S. passport?
☐ Yes ☐ No

(Circle one) Mr./Mrs./Ms./Miss

Name _____ (Please Print)

Address _____

City _____ State _____

Zip _____ Phone (___) _____

Business Phone (___) _____

BRITISH AIRWAYS

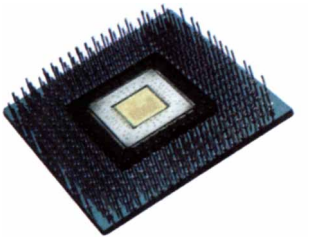
The world's favourite airline.®

THE POWER OF BELIEF: #1 in a series

When you aim for perfection, you discover it's a moving target.

Every advance in quality raises new expectations. The company that is satisfied with its progress will soon find its customers are not. It is this belief that has spurred Motorola to a 100-fold improvement in quality since 1981. ■ Our formula is a simple one: First, banish complacency. Second, set heroic goals that *compel* new thinking. Finally, "raise the bar"

Over 100,000 Motorolans carry this card at all times...a constant reminder of everyone's overriding responsibility.



Motorola semiconductor defects, once measured in parts per thousand, are now measured in parts per million and even parts per billion.

Winner 1988 Malcolm Baldrige National Quality Award

as you near each goal. Set it out of reach all over again. ■ Today, Motorola's standard is Six Sigma quality in all we do by 1992. In statistical terms: 99.9997 percent perfect. Our progress has been swift. Recently, Motorola shared the first Malcolm Baldrige National Quality Award, given by the President. ■ Total customer satisfaction, our goal, is now on the horizon. We dare not rest in its pursuit.

In 1988 Motorola was a winner of the Malcolm Baldrige National Quality Award.

Building On Beliefs

MOTOROLA

Ⓜ and Motorola are registered trademarks of Motorola Inc., ©1989 Motorola, Inc.

50 AND 100 YEARS AGO

MARCH, 1940: "Above the audible range of sound is the region of the ultrasonics. These soundless sound waves have recently become the focus for a flattering share of scientific attention. Ultrasonic vibration is, among other things: A producer of dispersions of solids in liquids but a destroyer of dispersions of solids in gases; in liquids, a disperser of solids but a coagulator of gases; in electrolysis, a promoter of desirable but a suppressor of undesirable gas evolution; with pathogenic bacteria, in some cases an augmentor and in others a diminisher of virulence; and, when applied to a human limb, a heater of the marrow but a non-heater of the bone!"

"After several months of excavations into the long-hidden depths of the Roman Colosseum, Prof. Guiseppe Cozzo has uncovered the *hypogeum* or 'back' stage, 20 feet below the arena, where the stage was set for gladiatorial combats with wild beasts, also where the stage settings and the cages for the beasts were kept. The *hypogeum* contained 32 cells for lifting the wild beasts simultaneously to the arena."

"The lowly milkweed can be made to yield a rare derivative that aids digestion and makes tough steaks tender, it has been revealed. Asclepain, the new derivative, is a proteolytic enzyme, which breaks down the protein in foods into more simple, easily digestible substances, it was reported."

"Man is a walking hotel in which more than 100 species of parasitic worms reside. These parasites lodge in his skin, in his muscles, in his lungs, in his intestines and in his brain, while several kinds swim through his blood. Every country in the world is plagued by worm diseases and the United States is not excepted. For example, about one million Southerners still have hookworm in spite of a 30-year campaign. Even more common than hookworm is trichinosis."

SCIENTIFIC AMERICAN

MARCH, 1890: "Aluminum, whether pure or in combination, deserves to rank with the noble metals. Prior to 1887, the entire amount of aluminum manufactured annually was only 10,000 pounds, and it sold that year at $10 a pound. The Hall process, on patents of Charles M. Hall, is now being carried on by the Pittsburgh Reduction Company, who are now selling pure aluminum at a rate cheaper than nickel. Briefly, Hall's process is this: A flux being discovered that, at a moderate temperature, takes the aluminum ore into solution, that is of lighter specific gravity, and that also is unaffected by the passage of an electric current, he fills a series of carbon-lined steel pots with the flux, which is kept molten. Carbon electrodes are plunged into these baths, through which passes the electric current, which acts to send the aluminum to the sides and bottom of each pot."

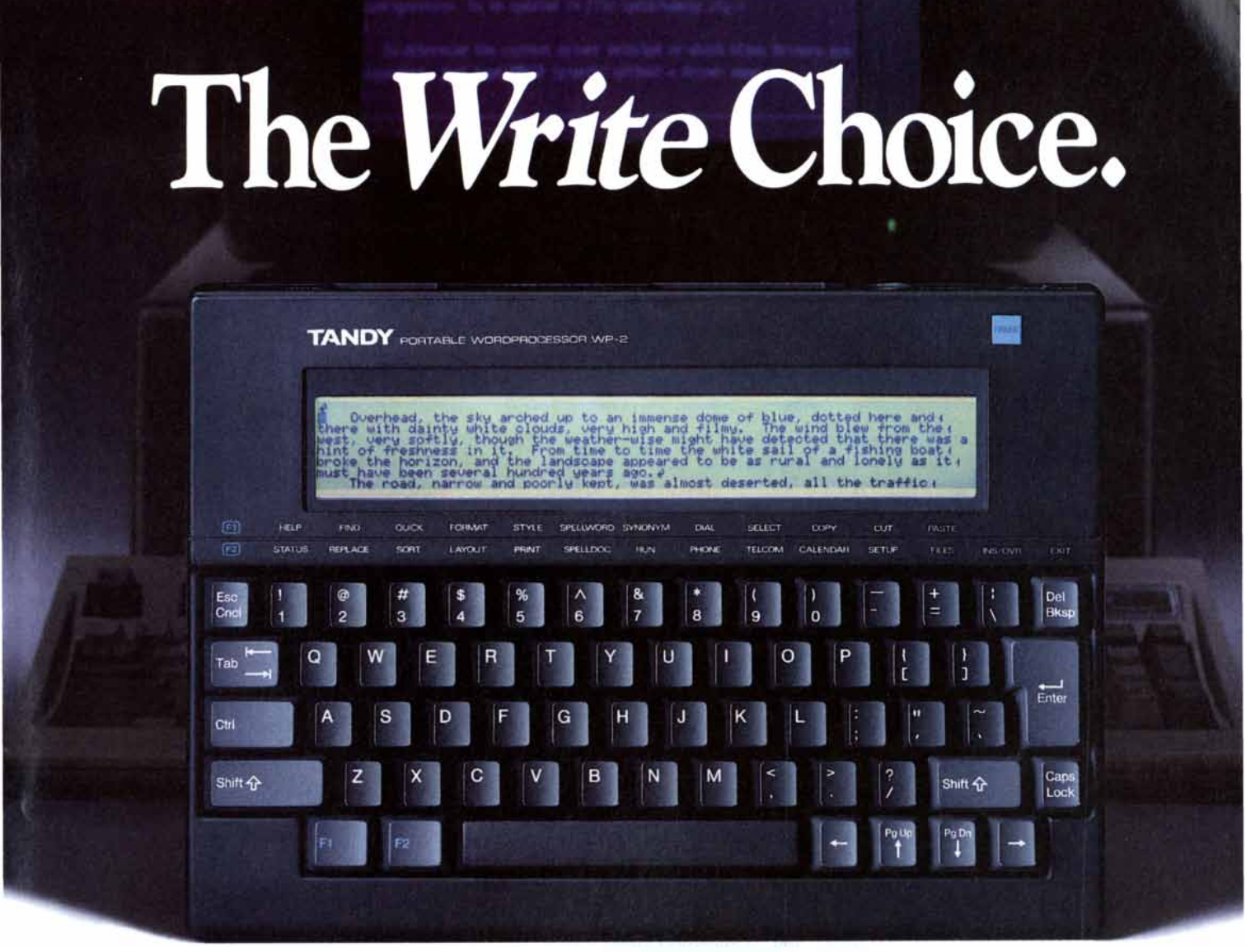
"A writer in the *British Medical Journal* makes a capital suggestion. It is, in short, a pencil or 'stick' for use on the chafed and irritated skin, or on skins very susceptible to insect bites, etc. He says that an addition of two per cent of cocaine to the ordinary cacao butter pencils converts the latter into a remedy, which gives almost instant relief when rubbed over the irritated spot."

"At a recent meeting of the Physiological Society, Berlin, Prof. Fritsch spoke on the anatomy of the torpedo fish. In opposition to the views of many recent investigators, who deny the nervous nature of the ganglion cells, he laid great stress upon the extremely close relationship which exists between the ganglia and end organs and is so strikingly shown in the torpedo. A thick nerve fiber, running from each ganglion cell to the electrical organ, divides into twelve to twenty-three fibrils, each of which is connected up with some one special plate of the organ."

"A system of artificially watering the parched, arid plains of the rainless region of the West has recently redeemed much waste land. This system of overcoming the defects of nature by artificial means will be extensively used in the future, when the population in those regions has become dense enough to create a greater demand for farm land than exists at present. One of the most perfect and extensive works of this nature is the San Diego Flume System, which has recently been completed, and which is designed to supply water to the city of San Diego and to irrigate the surrounding *mesas*."

Floating down a trestle on the San Diego flume

The *Write* Choice.



The Affordable, Portable Tandy® WP-2 Wordprocessor

When you're away from the office, you may only need a computer for one thing—writing. So why pay for computer functions you'll never use?

With the Tandy WP-2 there's no need! Everything you require to compose, view and edit text is combined in one small package. And it's so easy to use—if you can use a typewriter, you can operate the WP-2. Just turn it on and take notes, write letters and prepare reports.

The one-inch-thin, three-pound WP-2 is so compact it fits easily into your briefcase for traveling.

Take notes, prepare reports, write letters, and more. And because it isn't always convenient to lug around a heavy dictionary, the WP-2 has a 100,000-word spelling checker built in. And with its built-in 200,000-word thesaurus, you're never at a loss for words.

Back at the office, the WP-2 easily connects to a printer for as many copies as you like. Or, add a modem or adapter, and transfer files to your personal computer by phone.

Why buy more than you need? At just $349.95, the Tandy WP-2 Portable Wordprocessor is the *write* choice.

Price applies at Radio Shack Computer Centers and participating stores and dealers.

Radio Shack®
The Technology Store℠
A DIVISION OF TANDY CORPORATION

SCIENCE AND THE CITIZEN

Test Negative

A look at the "evidence" justifying illicit-drug tests

More than eight million working Americans had their urine tested for illegal drugs in 1989, and as many as 15 million will undergo such testing this year, according to the National Institute on Drug Abuse (NIDA). The fraction of companies that subject employees or job applicants to testing has jumped from 21 percent in 1986 to more than 50 percent last year, according to the American Management Association. The trend seems likely to continue: a majority of the respondents to a recent Gallup poll favored random drug testing of all workers.

What underlies the broad acceptance of a practice that conservative Supreme Court Justice Antonin Scalia has called a "needless indignity"? One factor may be the alarming statistics cited by testing advocates to demonstrate the high costs of drug abuse. Examination of some of these claims suggests that they do not always accurately reflect the research on which they are based. In fact, some of the data could be used to "prove" that drug use has negligible or even beneficial effects. Consider these examples.

• Last year President George Bush declared that "drug abuse among American workers costs businesses anywhere from $60 billion to $100 billion a year in lost productivity, absenteeism, drug-related accidents, medical claims and theft." Variants of this statistic abound in discussions about drug abuse and are commonly repeated without qualification by the media. Yet all such claims derive from a single study, one that "was based upon assumptions which need additional validation," according to an assessment last year by NIDA, the chief federal agency sponsoring research on substance abuse.

The study grew out of a survey of some 3,700 households by the Research Triangle Institute (RTI) in 1982. The RTI group found that the average reported income of households with at least one person who admitted to having *ever* used marijuana daily (20 days or more in a 30-day period) was 28 percent lower than the average reported income of otherwise similar households. The RTI researchers defined that difference in income as "loss due to marijuana use"; the total loss, when extrapolated to the general population, came to $26 billion. The researchers then added on the estimated costs of drug-related crime, health problems and accidents to arrive at a grand total of $47 billion for "costs to society of drug abuse." This figure—"adjusted" to account for inflation and population increase—represents the basis of Bush's statement, according to Henrick J. Harwood, who headed the RTI study and is now in the White House drug-policy office.

The RTI survey included questions on current drug use (at least once within the past month). Yet according to Harwood there was no significant difference between the income of households with current users of any illegal drug—including marijuana, cocaine and heroin—and the income of otherwise similar households. Does this mean that current use of even hard drugs—as opposed to perhaps a single marijuana binge in the distant past—does not lead to any "loss"? "You would be on safe ground saying that," Harwood replies.

• Officials of the U.S. Chamber of Commerce have testified before Congress and at national conferences on drug abuse that employees who use drugs are "3.6 times more likely to injure themselves or another person in a workplace accident . . . [and] five times more likely to file a workers' compensation claim." The pharmaceutical giant Hoffmann-La Roche, which is leading an antidrug campaign among businesses (and has a big share of the drug-testing market), also promulgates this claim in "educational" literature.

In fact, the study on which the claim is based has "nothing to do with [illegal] drug users," according to a 1988 article in the *University of Kansas Law Review* by John P. Morgan of the City University of New York Medical School. Morgan, an authority on drug testing, has traced the Chamber of Commerce claim to an informal study by the Firestone Tire and Rubber Company of employees undergoing treatment for alcoholism.

• In an interview with SCIENTIFIC AMERICAN, J. Michael Walsh, who heads NIDA's applied research division and is a strong supporter of workplace testing, singled out two studies that he said showed drug users are more likely to cause accidents, miss work and use health benefits. The studies were done at two utilities: the Utah Power and Light Company and the Georgia Power Company. The 12 workers in Utah and the 116 in Georgia who served as the primary research subjects were tested "for cause": they had either been involved in accidents, exhibited other "problem" behavior (commonly, high absenteeism) or submitted to treatment for alcoholism or drug abuse. Critics point out that it should not be terribly surprising if these subjects exhibited the cited traits at a higher-than-average rate.

What may be surprising is that, according to a report published by NIDA last year, Utah Power and Light actually "spent $215 per employee per year less on the drug abusers in health insurance benefits than on the control group." Those who tested positive at Georgia Power had a higher promotion rate than the company average. Moreover, Georgia workers testing positive only for marijuana (about 35 percent of all the positives) exhibited absenteeism some 30 percent lower than average. Nationwide, Morgan says, marijuana accounts for up to 90 percent of all positive findings, both because it is by far the most widely used illegal drug and because it persists in urine for up to a month (compared with two days for most other drugs).

• Perhaps the study most publicized of late by testing proponents involves employees of the U.S. Postal Service. The service tested 4,396 new hirees in 1987 and 1988 and—keeping the test results confidential—tracked the performance of positives (9 percent of the total) and negatives. By last September, the service reported, 15.4 percent of the positives and 10.5 of the

"Canon Brings The Copier Home."

"**Introducing the new PC-1.** It's really something—the ideal copier for the home.

First of all, because it's a Canon Personal Copier. That means it's so reliable, it's just about maintenance free. And so easy to use, a little kid can do it. That's all because Canon Personal Copiers are made with a special single cartridge system. Plus, the new PC-1's got a price tag even a mother—or father—can love.

And here's another terrific thing about the PC-1. You don't even have to wait around for it to warm up. Just turn it on, and it's ready to work. Or play. Thanks to Canon's breakthrough RAPID Fusing System.™

And the results? From tax forms to great party invitations, Canon Personal Copiers really deliver. You'll love what you see.

Finally, the wait is over. America's favorite copiers are coming home."

Canon U.S.A., Inc., Home Office Products Division, One Canon Plaza, Lake Success, N.Y. 11042. Regional phone numbers: New York (516) 488-6700, Chicago (708) 250-6200, Atlanta (404) 448-1430, Dallas (214) 830-9600, Los Angeles (714) 979-6000.

Canon
PC
PERSONAL COPIERS

No other system of keeping up can compare with SCIENTIFIC AMERICAN *Medicine.*

YOUR SYSTEM: a time-consuming, futile struggle to keep up with the information explosion.

The classic texts are convenient references—but the information they contain is obsolete before publication.

Like many physicians, you probably rely on the texts you first used in medical school. But even using the most recent editions, you find material that no longer reflects current clinical thinking—and they lack the latest information on such topics as herpes, oncogenes, AIDS, and photon imaging.

Reading stacks of journals alerts you to recent developments—but can't give you quick answers on patient management.

Struggling through the hundreds of journal pages published each month—even on only the really significant advances in the field—is arduous and memory-taxing. And it's a task that costs physicians valuable time—their most precious resource.

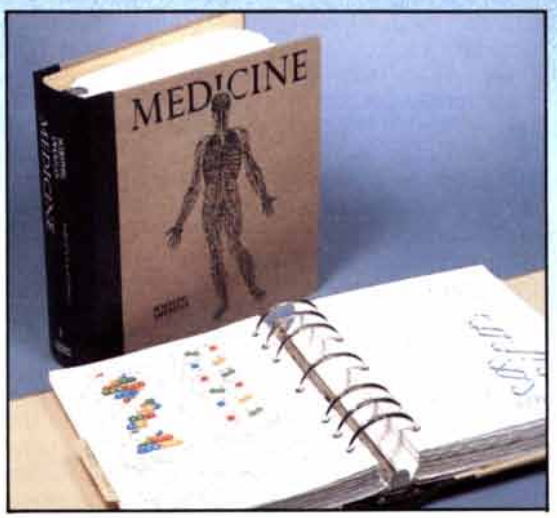
Review courses cover clinical advances—but, months later, do you recall the details of a new procedure or unfamiliar drug?

Seminars can also be costly and make you lose valuable time away from your practice—expenses that may amount to several thousand dollars. And, the speaker's skill often determines how much you learn.

SCIENTIFIC AMERICAN

MEDI

OUR SYSTEM: a rewarding, efficient way to keep yourself up-to-date—and save hundreds of hours of your time for patient care.



A comprehensive, 2,300-page text in two loose-leaf volumes, incorporating the latest advances in medical practice as of the month you subscribe.

This superbly designed, heavily illustrated resource, called "the best written of all [the internal medicine] books" by *JAMA* (251:807, 1984), provides a practical, comprehensive description of patient care in 15 subspecialties. And, because the text is updated each month, the clinical recommendations reflect all the current findings. A practice-oriented index and bibliography of recent articles further enhance the efficiency of the text.

Each month, six to nine replacement chapters to update your text *plus* new references, an eight-page news bulletin, and a completely new index.

You'd have to read hundreds of journal pages each month—and memorize the contents—to get the same information SCIENTIFIC AMERICAN *Medicine* contains. With our updated text, you read only the information you really need. Our authors, largely from Harvard and Stanford, sort through the literature and monitor developments, incorporating the significant advances into our chapters.

At no additional cost, a 32-credit CME program, to save you valuable patient-care time and the expense of attending review courses.

Earn 32 Category 1 or prescribed credits per year with our convenient self-study patient management problems; each simulates a real-life clinical situation. Choose either the complimentary printed version or, at a modest extra charge, the disk version for your IBM® PC or PS/2™, Macintosh™, or Apple® (or compatible).

Your subscription includes:
the two-volume, 2,300-page loose-leaf text
and, each month, six to nine replacement chapters,
a newsletter, new references,
and a completely revised index.

Or call toll-free: 1-800-345-8112

CINE

SCIENTIFIC AMERICAN MEDICINE
415 MADISON AVENUE, NEW YORK, N.Y. 10017

☐ **Yes,** I'd like to try the SCIENTIFIC AMERICAN *Medicine* system. Please enter my subscription at a first-year price of US$265* plus $7 shipping for a total of US$272.
☐ Also enroll me in the CME program. I'd prefer:
☐ the printed version at **no additional charge.**
☐ the disk version at an additional $US97.†
Computer type or compatible:
☐ IBM® PC (5¼") 256K ☐ IBM® PS/2 (3½")
☐ Macintosh™ (512K)
☐ Apple® IIc/IIe/IIGS
☐ Apple® II + with 80-col. card by: ☐ Apple® ☐ Videx®
☐ Enroll me in the disk CME **only** at US$152.†
(Note model above.)
☐ Check enclosed* ☐ Bill me
Signature _____
☐ VISA ☐ MasterCard Exp. Date _____
Account No. _____
Name _____
Specialty _____
Address _____
City _____ State _____ Zip Code _____

*Add sales tax for IA, Ill, Mich., or N.Y. Allow 8 weeks for delivery. Add US$10 for shipping to Canada. IBM is a registered trademark of International Business Machines Corporation. Apple is a registered trademark of Apple Computer, Inc. Videx is a registered trademark of Videx, Inc. Macintosh™ is a trademark of McIntosh Laboratory, Inc. and is used by Apple Computer, Inc. with its express permission.
†Please add sales tax for MI or NY.

negatives had been fired; the positives had also taken an average of six more sick days a year.

This study may be distorted by more subtle biases—related to race, age or gender—than those displayed by the utility studies, according to Theodore H. Rosen, a psychologist and a consultant on drug testing. Indeed, Jacques L. Normand, who headed the study, acknowledges that minority postal workers tested positive at a much higher rate than nonminority workers and that previous studies have shown minorities to have higher absenteeism.

Morgan points out, moreover, that the Postal Service study (like all those cited above) has not been published in a peer-reviewed journal. In fact, he says, only one study comparing the work of drug-test positives and negatives has passed peer review. Last year, in the *Journal of General Internal Medicine,* David C. Parish of the Mercer University School of Medicine in Georgia reported on a study of 180 hospital employees, 22 of whom had tested positive after being hired. Parish examined supervisor evaluations and other indexes and found "no difference between drug-positive and drug-negative employees" at the end of one year. He noted, however, that 11 of the negatives had been fired during that period and none of the positives.

• Proponents of testing often imply that drug use among workers is growing. A Hoffmann-La Roche brochure, for example, quotes Walsh pronouncing that "the problem of drug abuse has become so widespread in America that every company must assume that its employees will eventually be faced with a substance abuse problem." Yet in 1989 NIDA reported that illegal drug use has been decreasing for 10 years and that the decline has accelerated over the last five years. From 1985 to 1988 the number of current users (at least once in the last month) of marijuana and cocaine dropped by 33 and 50 percent, respectively.

To be sure, a subset of this group of current users is increasing: NIDA estimated that from 1985 to 1988 the number of people using cocaine at least once a week rose from 647,000 to 862,000 and daily users increased from 246,000 to 292,000. NIDA found that addiction to cocaine (including "crack") is particularly severe among the unemployed—who are beyond the reach of workplace testing.

Clearly, the U.S. has a drug-abuse problem. Could it be that neither indiscriminate testing of workers—which could cost upward of $500 million this year—nor the dissemination of alarmist information by testing advocates is helping to resolve that problem? —*John Horgan*

Polar Heat

The argument continues over an explorer's good name

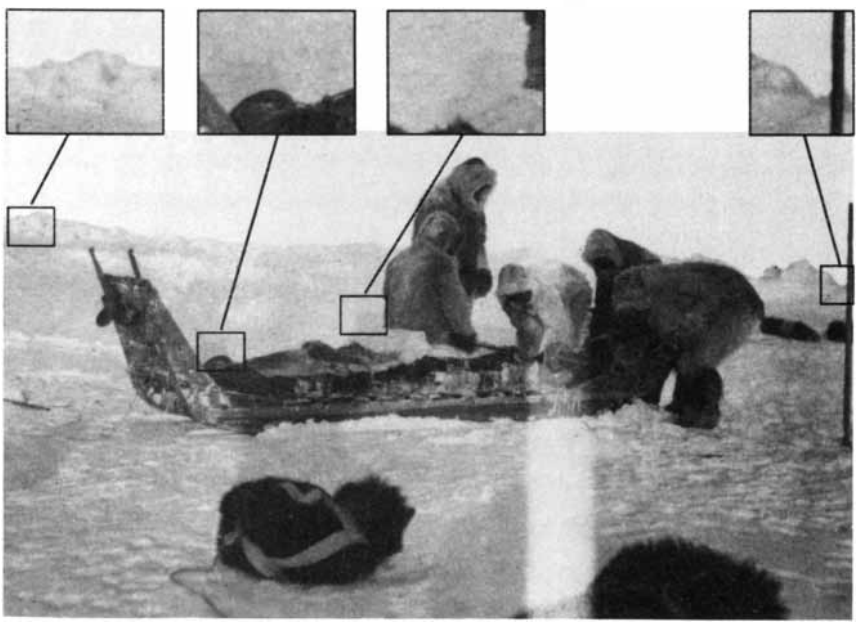
Did Robert E. Peary reach the North Pole in 1909, or was he a fraud who knew that he had not even come close? This question has resurfaced periodically ever since the admiral returned to civilization with a loose page of his diary proclaiming "the Pole at last!!!" Can Peary's claim be proved or disproved from the data available?

The most significant piece of new evidence adduced by the National Geographic Society (which sponsored Peary's expedition) and its independent commission of navigators is photogrammetric analysis of pictures taken during Peary's trip. The analysis reveals the angle of the sun above the horizon when the photographs were made; the experts say the angle is precisely the elevation of the sun at the pole when Peary claimed to have been there. Case closed.

Skeptics dismiss the photogrammetry and say Peary's supporters gloss over many questions about whether he reached the pole. Although deriving position from photographs "could work in principle," Charles T. Kowal of the Space Telescope Science Institute says, the technique is "useless in practice." Many factors could degrade the technique's margin of error from the claimed eight-to-25-kilometer (five-to-15-mile) range to 100 kilometers or more, he says. Among them he cites the difficulty of determining the exact positions on the old photographs of shadows and the objects that cast them, the uncertain angle of the horizon in a landscape full of ice hillocks and pressure ridges, unknown distortions introduced by the lens and its position relative to the film, and uncertainties about the time each photograph was made. Kowal thinks pinning down the angle of the sun to one part in 1,000 under such circumstances is an unlikely tour de force.

Other evidence presented by Peary's supporters is not inconsistent with the possibility that he reached the pole—but also not inconsistent with the possibility that he did not. For example, maps of the sea bottom

Analysis of Arctic shadows underpins a new claim that Peary reached the pole



***VANISHING POINT** determined by relative positions of objects and their shadows helps to mark the latitude of this 1909 photograph said to have been taken by Peary near the pole. The enlarged insets show the quality of the shadows and of the horizon line, both required to determine position. The original negatives were "improved" by either Peary or his publisher and no longer exist.*

"BOLD, BRILLIANT, GROUNDBREAKING

...Ranks among the most innovative and exciting science books to have been published in the last 40 years.
...A new perspective on the scientific landscape, and a visionary glimpse of the possible future of science.''*

THE EMPEROR'S NEW MIND

Concerning Computers, Minds, and the Laws of Physics

ROGER PENROSE

"When Mr. Penrose talks, scientists listen

...Penrose is a gifted mathematician with an impressive record of lighting lamps that have helped guide physics on its way....In 'The Emperor's New Mind'...he argues that we lack a fundamental insight into physics, without which we will never be able to comprehend the mind. Moreover, he suggests, this insight may be the same one that will be required before we can write a unified theory of everything."

—Timothy Ferris,
*The New York Times Book Review**

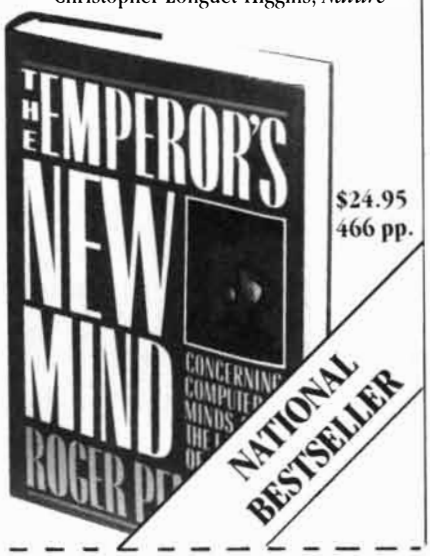
"Perhaps the most engaging and creative tour of modern physics that has ever been written."

—*The Sunday Times* of London

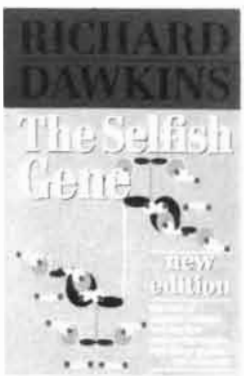
"I don't know any book which anyone, interested in the deepest problems of modern physics, will find more stimulating."
—John A. Wheeler, Princeton University

"Enchanting, tantalizing

...the testament of a brilliant man wrestling desperately—and unashamedly—with the deepest problems of metaphysics....Roger Penrose reveals himself as an eloquent protagonist, not only of the wonders of mathematics, but also of the uniqueness of people, whom he regards as mysterious, almost miraculous beings."
—Christopher Longuet-Higgins, *Nature*



Also of interest



THE SELFISH GENE
New Edition
Richard Dawkins

This is a substantial revision of one of the most innovative books on evolution in years. It features two completely new chapters—"Nice Guys Finish First" and "The Long Reach of the Gene"—plus new endnotes in which Dawkins replies to critics of the original edition or simply elaborates on a point. **$22.95, 350 pp.**

COSMOGENESIS
The Growth of Order in the Universe
David Layzer

Harvard astrophysicist David Layzer here probes the origin of order in the universe, taking readers on a sweeping tour of modern science, from spiral nebulae to neuromodulators, and from the evolution of the mind to the philosophical implications of modern physics. **$24.95, 322 pp.**

To order, or for information about other Oxford science books, simply complete the coupon below and mail to:

OXFORD UNIVERSITY PRESS
Dept WG-SCI I
200 Madison Avenue
New York, NY 10016.

☐ Please send information about other Oxford science books.

Enclosed is my check for $ _____
Please charge my ☐ VISA ☐ MasterCard

Card _____ Exp. Date _____

Bank # (M.Card only) _____

Signature (*Credit card orders not valid without a signature*) _____

Please send me the following books:

Author/Title	Price
Shipping ($1.50 first book, 75¢ each add'l.)	
TOTAL	

Name _____

Address _____

City/State/Zip _____

made during the past 40 years fix one sounding that Peary took about 270 kilometers from land within about 10 kilometers of his projected course, but soundings taken over the remaining 450 kilometers of the trek cannot be tied nearly so well to a particular chunk of bottom topography.

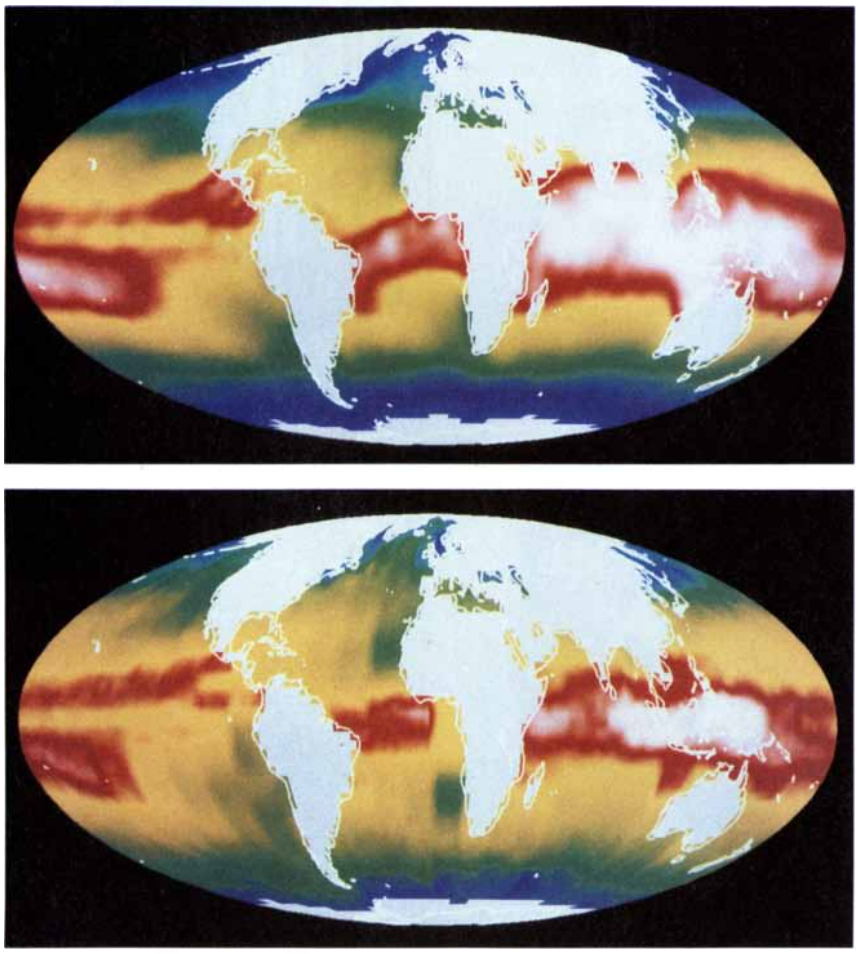
The question still remains: When Peary reported his position on April 6, 1909, as 89 degrees, 57 minutes, 11 seconds north latitude (a little more than five kilometers from the pole), was he truthful? In one obvious sense, he could not have been. That last second of arc represents a distance of just under 31 meters. Even today, with the help of satellite-borne navigation beacons, it is difficult to fix a position to within 30 meters. Both supporters and skeptics agree that Peary could at best have fixed his position to about six minutes of arc, or about 10 kilometers. Furthermore, the actual position of the pole wanders, sometimes by as much as hundreds of meters, as the earth's mass shifts and its axis of rotation wobbles; standing at the very spot is largely a symbolic act.

For many observers the record of Peary's polar accomplishments ranks him as a great explorer regardless of whether he bestrode the dimensionless point at the top of the world. In any case, if the minutes figure of Peary's April 6 record is correct, or even within two or three minutes of his actual position, then his loose diary entry would be justified. Given the evidence that remains today, however, it seems unlikely the question will ever be settled. —*Paul Wallich*

Vapor Trail

Observations support a key aspect of warming forecasts

Observations from ships, buoys and satellites show that water vapor may amplify a global warming



***SEA-SURFACE TEMPERATURES* (top) *correlate closely with measurements of the strength of the greenhouse effect* (bottom) *in images from the University of Chicago. The correlation seems to confirm that rising temperatures produce more water vapor and thereby an enhanced greenhouse effect. In both images the scale ranges from blue* (lowest) *through green, yellow and red to white* (highest).**

Lately a small but loud group of scientists has challenged predictions that the earth's climate will undergo significant warming in the coming decades. The skeptics point out, quite rightly, that computer models projecting a global warming incorporate assumptions that lack firm observational support. Modelers who forecast a warming retort that their calculations reflect these uncertainties. They are nonetheless gratified by a report that one of their key assumptions—which has been specifically questioned by a prominent skeptic—is borne out by observations.

The finding concerns the role that water vapor might play in a putative warming. Discussions of climatic change generally focus on carbon dioxide and other gases whose atmospheric levels are increasing as a result of human activity. The gases warm the earth through a greenhouse effect: they let incoming sunlight pass through but block the infrared radiation, or heat, emitted by the earth.

Water vapor is also a greenhouse gas, the most abundant of them all. In 1963 Fritz J. Möller proposed that water vapor would amplify any warming initiated by the other greenhouse gases through a positive-feedback effect: the initial warming would cause more water to evaporate from the oceans, leading to more water vapor in the atmosphere, an enhanced greenhouse effect and further warming.

The proposal was widely accepted and has become an integral part of modern climate models. Ameet Raval and Veerabhadran Ramanathan of the University of Chicago are the first investigators to test the hypothesis rigorously.

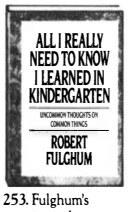
The two scientists analyzed observations of the earth's oceans, where water-vapor feedback is expected to be most pronounced. They determined the amount of heat emitted by a given region of the ocean from temperature readings by ships and buoys. They then consulted satellite observations by the Earth Radiation Budget Experiment (ERBE) to establish how much of this heat escaped into space and—by inference—how much was blocked by a greenhouse effect. These results were correlated with humidity measurements from the *Nimbus 7* satellite.

"The greenhouse effect is found to increase significantly with sea surface temperature," the workers conclude

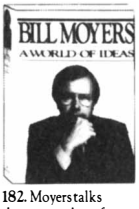
Tres Libros por Tres Dolares and a Farewell to Commitment. No Bull.

198. QPB's specially produced volume of Woody Allen's three hilarious books.
(3-in-1 vol.)
QPB: $12.95

***482.** This latest edition of the famous dictionary has 40,000 memorable quotations.
Hardcover: $45
QPB: $18.95

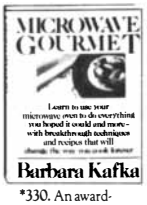


253. Fulghum's premise is that we learned the important lessons by age 5.
Hardcover: $15.95
QPB: $7.95



182. Moyers talks about everything from morality to science.
Hardcover: $29.95
QPB: $15.95

601. The evolution of beliefs throughout history. Illustrated.
Hardcover: $29.95
QPB: $12.95



***330.** An award-winning food writer offers more than 600 recipes for your microwave oven.
Hardcover: $19.95
QPB: $9.95

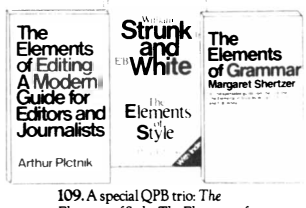
***406.** "A rollercoaster ride over a vast landscape of the imagination"—*London Guardian*.
Hardcover: $19.95
QPB: $9.95

Ernest Hemingway

***447.** A hilarious new collection from the author of *Lake Wobegon Days*.
Hardcover: $18.95
QPB: $9.50

204. Relax with Calvin & Hobbes in this new full-color collection of Sunday cartoons.
QPB: $7.95

***256.** What to write, how to write it, where to sell it.
Hardcover: $27.95
QPB: $13.75



109. A special QPB trio: *The Elements of Style, The Elements of Grammar, The Elements of Editing.*
3-volume set
QPB: $10.95

207. A beguiling novella that asks: How do you thank a man who won't be thanked?
QPB: $5.95

***321.** Stephen Hawking offers a convincing big picture of the origins of the cosmos.
Hardcover: $18.95
QPB: $9.50

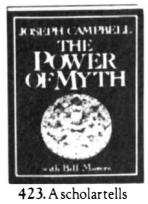
(Translation: 3 books, 3 bucks. No commitment. No kidding.)

No matter who says it, it's an amazing deal—any 3 books from this page for just $3 (plus shipping and handling). After that you don't have to buy any more books ever.

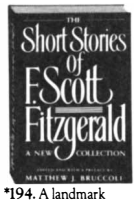
You'll receive the *QPB Review* 15 times a year (about every 3½ weeks). It lists our selection of fiction, history, science, biography plus excerpts, reviews, photos and anecdotes. And QPB softcover books are priced at up to 60% less than their hardcover counterparts. Every book you buy, after this offer, will earn you Bonus Points that can be traded in for free books—you pay just shipping and handling. So set your sights on 3 books and join QPB today.

***117.** Over 1600 problem-solving drawings with explanations on every aspect of home repair.
Hardcover: $24.95
QPB: $12.95

211. Take a trip down Larson's memory lane with over 300 color and black-and-white cartoons.
QPB: $10.95



423. A scholar tells how myths have shaped our lives. A PBS TV series.
Hardcover: $27.50
QPB: $15.95



***194.** A landmark omnibus collection of 43 of Fitzgerald's best short stories.
Hardcover: $29.95
QPB: $13.95

178. Hell hath no fury like a woman scorned—at least in this bitingly funny novel.
Hardcover: $13.95
QPB: $6.95

267. The story of civilization in a unique chart format.
Hardcover: $29.95
QPB: $15.95

***716.** From the pages of *Rolling Stone*, a documentary of the past 20 years.
Hardcover: $24.95
QPB: $12.95

297. Soothing therapeutic techniques. 150 graphic photographs.
QPB: $9.50

***QPB Exclusive:** The only softcover edition now available.

How QPB Membership Works.

Selection: Each *Review* lists a new Selection. If you want it, do nothing—it will be shipped automatically. If you want another book or no book at all, complete the Reply Form always enclosed and return it by the specified date. (A shipping and handling charge is added to each shipment.)

Return Privilege: If the *QPB Review* is delayed and you receive the Selection without having had 10 days to notify us, you may return it for credit.

Cancellations: You may cancel membership at any time by notifying QPB. We may cancel your membership if you elect not to buy at least one book in any six-month period.

OK. Send me 3 books for 3 bucks. No commitment. No bull.

Quality Paperback Book Club,® P.O. Box 8804, Camp Hill, PA 17011-8804

Please enroll me in QPB and send me the 3 choices I've listed below, billing me only $1 each, plus shipping and handling charges. I understand that I am not required to buy another book. You will send me the *QPB Review* (if my account is in good standing) for at least six months. If I have not bought at least one book in any six-month period, you may cancel my membership.

0-16

Indicate by number your 3 choices

Name *(Please print clearly.)* QB170-3-0

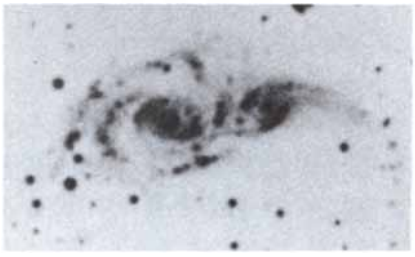
Address

City State Zip

QPB® Quality Paperback Book Club

Prices are generally higher in Canada. © 1990 Quality Paperback Book Club. All orders subject to approval.

"Eyeball" galaxies and pairs of quasars may both result from galactic close encounters



***INTERACTING GALAXIES can produce temporary "eyeball" galaxies* (left); *these may later develop a barlike structure like one recently detected in the Milky Way. More violent encounters may send matter falling into black holes at the center of galaxies and produce quasar pairs* (simulated above).**

in *Nature*. "The rate of increase gives compelling evidence for the positive feedback between surface temperature, water vapour and the greenhouse effect; the magnitude of the feedback is consistent with that predicted by climate models." Other climatologists find the report convincing. "It's a first-rate piece of work," says Michael McElroy of Harvard University.

The finding deals a blow to a proposal by Richard S. Lindzen of the Massachusetts Institute of Technology, a noted critic of global-warming predictions, that water vapor might diminish rather than amplify a global warming. In Lindzen's view, an initial warming would increase atmospheric convection in such a way that more water vapor would fall back to the earth as precipitation; the net result would be less water vapor in the atmosphere and a smaller greenhouse effect.

"We didn't find a shred of evidence for [Lindzen's] theory," Ramanathan says. He hastens to add that many other questions about climate models remain. "The big uncertainty is the cloud issue," he says. "It could be positive or negative feedback. No one knows." —*J.H.*

PHYSICAL SCIENCES

Astronomy Potpourri

Footballs, eyeballs, attractors and twins

Seventeen hundred astronomers gathered at the American Astronomical Society meeting in Washington, D.C., in January to present and hear the latest cosmic findings. Perhaps the biggest surprise of the meeting was that the *Cosmic Background Explorer* (COBE), launched last November 18, has already yielded precise measurements of the pervasive microwave and infrared emissions that are thought to be relics of the big bang that formed the universe.

The data are most remarkable for what they do not show. The spectrum of the background radiation shows no discernible deviation from the spectrum expected from a mathematically ideal radiating object; this disproves observations made in 1987 that reportedly showed intriguing irregularities. COBE also has found no signs of lumps or explosions in the early universe that might explain the clumpy structure astronomers see today. As COBE continues to collect data, these measurements will improve tenfold; if no lumps appear at that level, cosmologists will have to discard many current theories of how galaxies formed.

Meanwhile evidence that the universe is full of enormous, complex structures continues to pile up. Alan M. Dressler of the Carnegie Institution of Washington and Sandra M. Faber of the University of California at Santa Cruz have mapped the flow of galaxies around the previously announced "Great Attractor"—a tremendous conglomeration of galaxies, covering one third of the sky in the Southern Hemisphere, which is pulling the Milky Way and all nearby galaxies toward it—and demonstrated that it is indeed a real structure with an observable far side.

Ed Shaya of Columbia University and R. Brent Tully of the University of Hawaii have attempted to calculate the mass of the Great Attractor to determine if it can account for all of the Milky Way's motion in that direction. They find that the observed galaxies account for only about two thirds of the motion, and so they propose the existence of a larger, more distant gathering called, of course, the Even Greater Attractor.

A particularly intriguing, if speculative, finding was reported by T. J. Broadhurst of the University of Durham, U.K., and colleagues, who have accumulated evidence that the population density of galaxies varies regularly across most of the visible universe, with peaks every 400 million light-years. The mechanism that could generate such enormous coherent structures is entirely unknown.

Another surprising report concerned the discovery of a double quasar. Most observed pairings of quasars are thought to be optical illusions, the result of the bending of light by

distant, massive galaxies. A team led by Georges Meylan of the Space Telescope Science Institute thinks it has discovered the first known physically paired quasars. The two objects seem to lie at virtually the same distance from the earth—about 12 billion light-years—which implies that they are 100,000 light-years apart, less than the diameter of the Milky Way.

The extreme proximity of these two objects suggests that they are interacting with each other gravitationally. Current theories hold that close encounters between certain types of galaxies may switch on incipient quasars in the galaxies' cores. Meylan speculates that the two quasars may be undergoing just such an encounter and that "we may be witnessing the birth of a quasar pair."

Closer to home, the decades-old image of the Milky Way as a spherical nucleus of stars with arms trailing from it has been overturned by Leo Blitz of the University of Maryland at College Park and David N. Spergel of Princeton University. They found that the inner region is actually football-shaped, making the Milky Way not a spiral galaxy but a mild form of a "barred spiral" in which the spiral arms trail from a bar, rather than a sphere, of stars.

How could this structure come to be? One possibility was raised by Debra Meloy Elmegreen of Vassar College, who has been studying a peculiar class of interacting galaxies that she whimsically calls "ocular" because of their eyeball-like shape. These galaxies, she believes, settle down quickly in cosmic terms (perhaps a couple hundred million years), quite possibly to become barred spiral galaxies. The Milky Way has two close companion galaxies, the Magellanic Clouds, that may have interacted with it in the past, but Elmegreen says that her attempts to simulate such interactions on a computer have not been able to create a football structure like the one observed. *—Corey S. Powell*

Broadcast Warning

Did radio waves portend the Loma Prieta quake?

One of the most troubling aspects of the Loma Prieta earthquake, which rocked the San Francisco Bay Area last October, was its lack of premonitory tremors. In vain geologists have scoured seismic data recorded before the quake for signs of warning. Yet the quake may

have signaled its coming in another way—with a radio broadcast. Geologists say the serendipitous finding, if confirmed, may help them achieve a long-sought goal: short-term prediction of earthquakes.

The radio signals were discovered by a group of electrical engineers led by Antony C. Fraser-Smith of Stanford University. For years he has explored the potential of very-low-frequency (VLF) radio waves as a communications medium for submarines. VLF radio waves oscillate at less than 30,000 cycles per second, or hertz. Although conventional radio waves (such as the AM variety, which vibrate at about a million hertz) can carry more data, VLF waves pass more readily through water—and the earth.

Since 1987, in an effort to survey background noise that might affect VLF communications, the Stanford team has monitored radio waves ranging from .01 to 10 hertz. This ultra-low-frequency (ULF) margin of the VLF band has been largely neglected by other scientists, according to Fraser-Smith. He and his colleagues originally installed their antenna, a metal cylinder wrapped in wire, on the Stanford campus. Then they moved the device to Corralitos, a small town where electromagnetic emissions from cars and a mass-transit rail system would not distort the readings. Corralitos is only four miles from the epicenter of the Loma Prieta earthquake.

The quake damaged data-storage equipment at the Corralitos site, but after some delay Fraser-Smith and a graduate student, Arman Bernardi, finally retrieved the data. They immediately noticed that the amplitude of the .01-hertz signals had increased dramatically about three hours before the quake. The effect persisted until the quake cut off power to the computer recording the data. The researchers checked for signs of equipment malfunction but, somewhat to their surprise, found none.

What caused the signals? Fraser-Smith speculates that increased stress on rocks along the fault line generated a surge of piezoelectricity, or pressure-induced current, that in turn produced radio waves. Another possibility, he says, is that prequake activity altered the flow of preexisting electric currents in the earth. (Such currents are a normal by-product of electromagnetic effects in the atmosphere.) The fact that only the lowest-frequency (.01-hertz) radio waves were detected implies that the prequake activity occurred well beneath the surface.

Experts on earthquakes have reacted to the Stanford group's finding with a mixture of excitement and caution. Allan G. Lindh of the U.S. Geological Survey in Menlo Park says that after 20 years of investigating claims of surefire earthquake precursors, he is inclined toward skepticism. "But I can't find a crack in their arguments," he says, "and no one else can either."

Malcolm J. Johnston of the USGS points out that the geologic literature is full of "hints" that earthquakes may be linked to electromagnetic activity. Japanese and Soviet scientists claim to have detected strong VLF signals before and during quakes. Witnesses have also recalled lightninglike flashes emanating from the earth during quakes and poor radio reception beforehand. Johnston adds, however, that it is difficult to imagine a mechanism that could induce powerful electric currents within the earth without any simultaneous seismic activity.

Johnston and Lindh agree that only further observations will establish whether ULF signals have predictive potential. With the encouragement of the USGS, Fraser-Smith is seeking federal funds for more ULF antennas for earthquake monitoring. A likely location is Parkfield, a town on the San Andreas fault where geologists expect a magnitude-6 earthquake to strike soon.

Fraser-Smith suspects that only large quakes produce detectable signals. His equipment, he notes, did not record signals before or during a magnitude-5 quake near Corralitos this past August; the magnitude-7.1 Loma Prieta quake was more than 100 times stronger. He hopes that even if ULF radio waves do not prove to be reliable earthquake predictors, they may still provide scientists with a new source of information about these destructive events. —*J.H.*

BIOLOGICAL SCIENCES

Bee Police

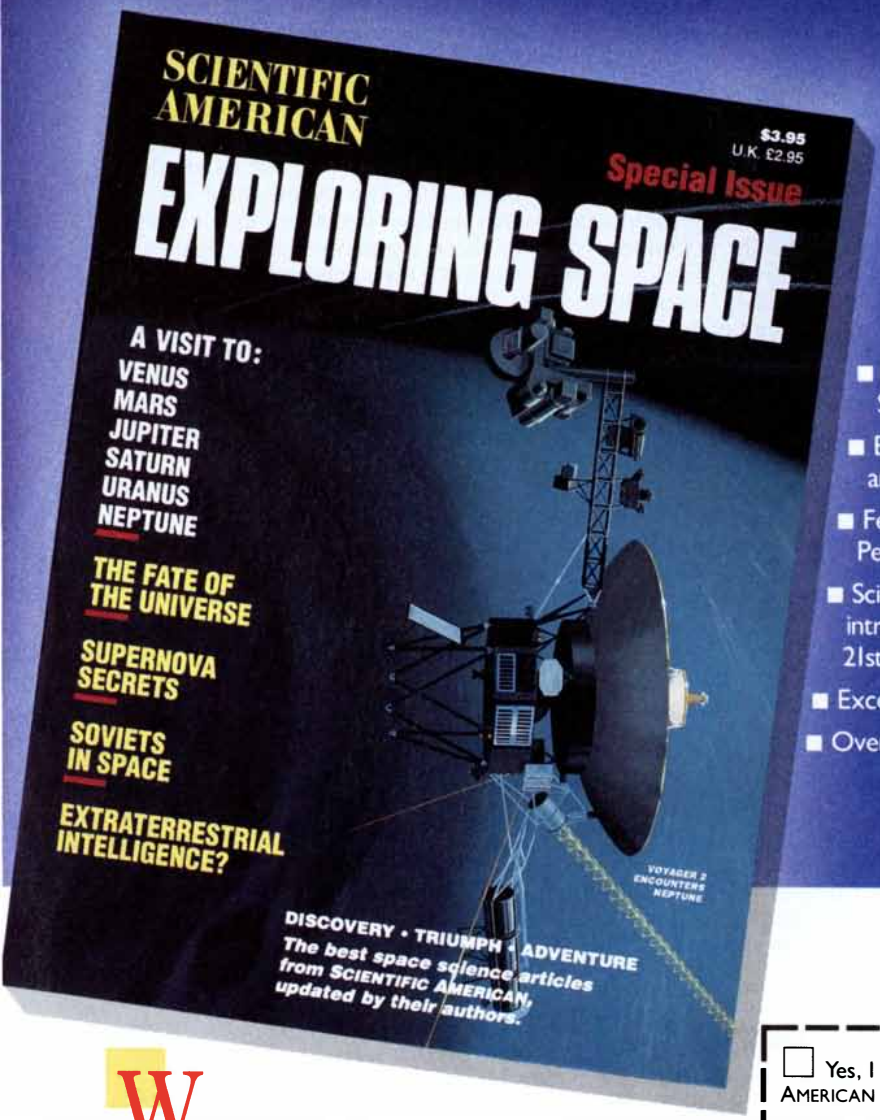
How coercion leads to cooperation among honeybees

Honeybee workers have often been admired for their selfless citizenship. Whereas the queen and the male drones live merely to breed, the female workers do the chores needed to keep the hive humming: gathering nectar, building the nest and nurturing the queen's progeny. Yet recent experiments reveal that workers are not as altruistic as

SCIENTIFIC AMERICAN

PRESENTS

A SPECIAL JOURNEY INTO SPACE



Sponsored by

Lockheed
Giving shape to imagination.

VITAL AND POPULAR ARTICLES COVERING SPACE AND ITS EXPLORATION ASSEMBLED BY THE EDITORS OF SCIENTIFIC AMERICAN.

- Introduction by Jonathan Piel, the editor of Scientific American
- Extensively illustrated: Outstanding color photography, art, and graphics
- Features by Carl Sagan, Frank Drake, Sally K. Ride, Peter M. Banks, and Ronald G. Prinn
- Scientific American's concise and informative introduction to space exploration approaching the 21st century
- Excellent classroom teaching aid
- Over 150 pages

What's been learned in recent years of the freezing and fiery heavens? Here is an essential collection of answers that only the editors of Scientific American could put together. Fourteen fundamentally important articles culled from over a decade of publishing. Articles originally written and newly updated by the men and women who know the story best. Beautifully photographed and illustrated, over 150 pages.

ON SALE AT YOUR NEWSSTAND APRIL 12TH FOR ONLY $3.95

☐ Yes, I would like to receive the special issue SCIENTIFIC AMERICAN Exploring Space as soon as it is available. Please send me _____ copies @ $3.95 per copy (U.S. addresses; $4.95 elsewhere). Add $1.00 per copy for postage and handling. Postage and handling paid by Scientific American on orders of eight or more copies.

Enclosed is my check/money order (payable to Scientific American) for $_____. SA3

Name _____

Company _____

Address _____

City/State/Zip _____

Send Order to: Scientific American, Dept. ES, [illegible] Madison Avenue, New York, NY 10017, U.S.A.

they appear: they produce the illusion of willing allegiance to their queen through brutal suppression of each other's attempts to compete with her in producing offspring.

A review of apiarian genealogy helps to illuminate this dark family secret. Typically the queen is mother to all the bees in a hive; after mating with as many as 20 drones from other colonies, she bears daughters who become workers. Those workers who share a father are full sisters, but most are only half sisters to each other. The queen also lays unfertilized eggs that become drones, who are brothers to all the workers in the colony.

Although workers cannot mate, they are not wholly sterile. When a queen dies, workers often lay unfertilized eggs that—like those of the queen—hatch into male drones. Yet workers rarely reproduce while a queen reigns. Last fall P. Kirk Visscher of the University of California at Riverside reported that workers produce about one in 1,000 of the drones in a hive.

Why don't workers produce more male offspring when the queen is present? After all, a worker's son carries more of her genes than do her brothers, the drones spawned by the queen. According to modern evolutionary theory, a worker would enhance her fitness—or ability to propagate her genes—by hatching her own eggs in addition to or in place of the queen's.

On the other hand, a typical worker's fitness would be diminished if the queen's sons were supplanted by other workers' sons, most of whom would only be half nephews to the worker. Building on this insight, Francis L. W. Ratnieks of the University of California at Berkeley proposed two years ago that workers somehow block each other's attempts to reproduce through mutual "policing."

Ratnieks and Visscher teamed up to test the hypothesis. They put unfertilized eggs laid by workers and by the queen in a hive. As suspected, workers in the hive quickly sniffed out and devoured virtually all of the workers' eggs while leaving most of the queen's eggs alone. Does such policing occur often, or are most workers deterred from laying? "We're studying that now," Visscher says.

In a comment accompanying Ratnieks and Visscher's report in *Nature,* Jon Seger of the University of Utah compares the honeybee society to the totalitarian state described by George Orwell in *Nineteen Eighty-Four.* Ratnieks suspects that coercion may underlie cooperation among other highly organized insects, such as ants. "Insects may be a little more like us than we'd thought," he says. —*J.H.*

WORKER BEE finds and then eats an egg laid by another worker in these photographs made by Francis L. W. Ratnieks.

Who's the Dealer?

What controls gene shuffling in the immune system?

Just as the dealer in a poker game can arrange the 52 cards into almost 2.6 million different hands, unidentified enzymes in the immune system variously combine a few hundred pieces of DNA into the genes for billions of different antibodies. This gene-shuffling trick is crucial to the body's natural defenses: if the cells of the immune system called lymphocytes had to carry the complete set of genes for all the antibodies needed to protect the body from invaders, they would not have enough DNA left over to encode anything else. Now two groups in the U.S. and Japan have identified the genes for two different proteins, one of which may be the recombinase, the long-sought enzymatic "dealer" in this immunologic game.

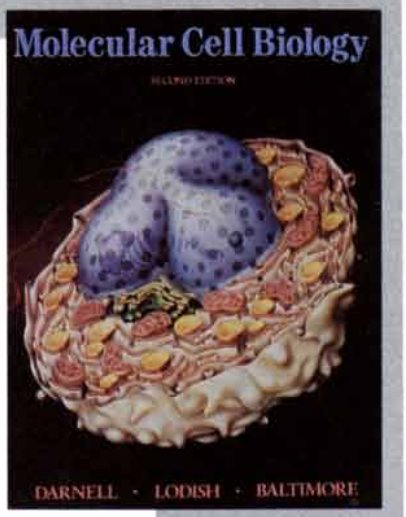
The search for this elusive entity began in 1976, when Susumu Tonegawa and Nobumichi Hozumi, who were then at the Institute for Immunology in Basel, demonstrated that genes for antibody proteins were assembled randomly from separate variable (*V*), joining (*J*) and, sometimes, diversity (*D*) gene segments [see "The Molecules of the Immune System," by Susumu Tonegawa; SCIENTIFIC AMERICAN, October, 1985]. The hypothetical enzyme that could recognize and splice together these segments became known as the *V*(*D*)*J* recombinase.

For many years the best candidate for the *V*(*D*)*J* recombinase seemed to be a product of the gene known as *scid,* for severe combined immunodeficiency. Mutant mice manifesting *scid* are incapable of assembling immunoglobulin genes properly. The *scid* gene has never been identified precisely, but it is known to sit on chromosome 16.

In the spring of 1986, David G. Schatz and David Baltimore of the Whitehead Institute for Biomedical Research began to close in on a gene that was centrally involved in *V*(*D*)*J* recombination. They demonstrated that DNA from lymphocytes could trigger *V*(*D*)*J* recombination in fibroblasts, a class of connective-tissue cells that do not ordinarily recombine the segments. More recently, in *Cell,* Schatz, Baltimore and their colleague Marjorie A. Oettinger announced they had isolated the single gene responsible for this activity, which they dubbed the recombination activating gene, RAG-1. Tests by Schatz, Oettinger and Baltimore proved that RAG-1 is not *scid:* the two genes are located on different chromosomes.

Is RAG-1 the gene for the recombinase itself, or does it marshal other, subordinate recombinase enzymes? Schatz admits readily that the available evidence is ambiguous and that "this is the critical question that will guide our future research." He does point out, however, that several weak lines of circumstantial evidence suggest RAG-1 protein is the recombinase. RAG-1 has been evolutionarily conserved among animals as unrelated as frogs, horses, chickens and humans,

KEEPING YOU CURRENT IN TODAY'S CHANGING WORLD



New for 1990!

The Eagerly Awaited Revision
of a Groundbreaking Book

MOLECULAR CELL BIOLOGY
Second Edition

James Darnell, Harvey Lodish, and David Baltimore
A Scientific American Book

The most authoritative, comprehensive guide to the molecular biology of the cell, MOLECULAR CELL BIOLOGY is the only book to fully present modern techniques and findings and tie these developments to basic principles.

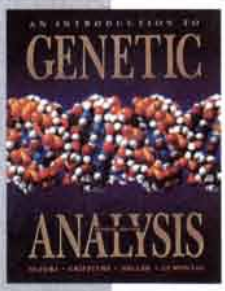
Thoroughly updated to cover the latest advances and fully revised for greater accessibility and convenience, MOLECULAR CELL BIOLOGY is an ideal reference. The Second Edition features a brand new full color art program that adds a wealth of visual detail. Don't miss it!

New for 1990! Approx. 1100 pages, 1095 illustrations, hardbound.
Order your copy today and save 20%! Offer good until April 15, 1990.

BIOCHEMISTRY, Third Edition
Lubert Stryer

"Destined to remain one of the 'bibles' of our subject for many years to come."—*Trends in Biochemical Sciences*

The world's leading biochemistry book is an indispensable desk reference for anyone in the biochemical sciences.
1988, 1098 pages, 1476 illustrations, hardbound



AN INTRODUCTION TO GENETIC ANALYSIS
Fourth Edition
David T. Suzuki, Anthony J. F. Griffiths, Jeffrey H. Miller, and Richard C. Lewontin

This profusely illustrated book is an accessible introduction to the principles, methods, and findings of genetic analysis. Including all recent advances, it presents balanced coverage of classical and modern genetics.
1989, 768 pages, 831 illustrations, hardbound

Your satisfaction matters!
Return any book in good condition within 15 days of receipt and we'll send you a prompt refund.

To order, mail this coupon or a copy to:
W. H. Freeman and Company
4419 West 1980 South
Salt Lake City, UT 84104

☐ I enclose my check or money order made payable to W. H. Freeman and Company.

☐ Please use my: ☐Visa ☐MasterCard
Exp. Date_____
Account #_____
Signature_____
(Credit card orders must be signed)

Name_____
Address_____
City/State/Zip_____

Title	Author	ISBN	Price	Total
Molecular Cell Biology, 2/e	Darnell et al.	1981-1	$42.59*	
Biochemistry, 3/e	Stryer	1843-X	$52.00	
Intro to Genetic Analysis, 4/e	Suzuki et al	1956-8	$42.95	

Subtotal _____
(Add $1.95 for shipping & handling of first book, $1.25 for each additional book.) _____
Sales Tax (NY, CA, UT) _____
TOTAL _____

* Darnell, et al. regular price $52.95.

For Canadian residents add $2.25 for postage and handling, $1.50 each additional book. Please allow four weeks for delivery. All orders should be prepaid in US dollars.

4827

W. H. Freeman and Company
The book publishing arm of Scientific American

as would be expected of so important an enzyme. It is naturally expressed only in developing lymphocytes and then only during the stages in which recombination takes place. Perhaps most important, RAG-1 does not seem to confer any lymphocytic properties on fibroblasts except for *V(D)J* recombination; a developmental or regulatory protein might be expected to trigger other changes as well.

Another possible nominee for the recombinase is the product of the gene RBP-2, which was recently reported in *Nature* by a group led by Tasuku Honjo of Kyoto University. RBP-2's protein binds specifically with the "recognition sequence" that flanks the *J* gene segments before recombination; the protein also has strong similarities to the class of proteins called integrases, which splice new DNA into the chromosomes of bacteria. It has not yet been demonstrated that RBP-2 plays a role in *V(D)J* recombination, however; RAG-1 clearly does.

According to Frederick W. Alt, an immunologist at Columbia University, it is not yet clear how to fit RAG-1, RBP-2 and *scid* into a coherent picture of recombinase activity. RAG-1 protein could be the principal recombinase, with RBP-2 and *scid* proteins acting as necessary accessory factors. Alternatively, RAG-1 may only be a master switch that turns on the recombinase proteins of RBP-2 and other genes. —*John Rennie*

Smelling Better

Smell-blindness "cure" may point to olfactory mechanism

Biological investigations often find their justification in the medical cures they may generate. Now a "cure" has been effected whose sole significance lies in the clues it provides to a continuing biological mystery: the mechanism by which smells are perceived.

The cure is for a deficiency in the ability to smell androstenone, a volatile steroid found in human perspiration and boar's saliva, whose odor has been variously likened to that of urine or of sweat. This olfactory blind spot, or specific anosmia, affects nearly half of the adult population. Almost everybody displays one of the more than 100 known anosmias, but aside from those for cooking gas, skunk scent and other auguries of danger, anosmias cannot be deemed handicaps. Indeed, in crowded urban settings they may confer adaptive advantages.

Writing in the *Proceedings of the National Academy of Sciences*, Charles J. Wysocki, Kathleen M. Dorries and Gary K. Beauchamp of the Monell Chemical Senses Center in Philadelphia report that they were able to overcome the anosmia for androstenone in 10 of 20 people by subjecting them to a six-week regimen of daily sniffing sessions.

It is surprising that an anosmia can be reversed at all. The syndrome's genetic basis, which has been demonstrated by studies of identical and fraternal twins, would be expected to make it stable over time, Wysocki says. He thought of doing the experiment only after finding that months of intermittent exposure to androstenone had conquered his own anosmia for it.

Wysocki thinks the key to the puzzle may lie in the two traits that distinguish olfactory nerve cells from other neurons: their ability to recognize a wide variety of foreign bodies and their ability to reproduce themselves. Both traits recall the functioning of the immune system, a resemblance noted in 1974 by Lewis Thomas when he was at the Memorial Sloan-Kettering Cancer Center in New York City. Thomas even suggested that there might be a genetic link between the two systems. Since then, studies have shown that mice can distinguish, simply by scent, other mice whose only genetic difference lies in the major histocompatibility complex, which plays an important role in the immune response. No such link between smell and the immune system has been demonstrated in human beings, but workers are beginning to look for one in tests of the olfactory responses of AIDS patients.

How might the anosmia have been reversed? Wysocki and his colleagues hypothesize that a particular odorous molecule, such as androstenone, stimulates certain olfactory receptors to multiply, much as an antigen stimulates particular lymphocytes to multiply by clonal expansion. "In one scenario, anosmic subjects have very few neurons with molecular receptors for androstenone, and sensitization causes those neurons to proliferate," Wysocki says. "In another scenario, the number of neurons stays constant, but the number of molecular receptors on each neuron increases."

In either case such proliferation during sensitization might have had the effect of enhancing a vestigial sensitivity to androstenone until the subject could smell the steroid. Wysocki says the hypothesis will now have to be tested—for example, by measuring changes in electric potentials in neurons of animals exposed to an odor or by examining the ability of isolated neuronal membranes to bind odorous molecules selectively.

There is more to the study of the nose than meets the eye. Olfactory receptors are thought to resemble synaptic receptors (by means of which one nerve receives electrochemical signals from another); understanding the one might help workers learn about the other. Olfactory sensitization may also illuminate the process by which new olfactory cells grow their axons to the brain and form new synapses there. "If we could learn what causes the turnover in olfactory neurons, we might learn how to stimulate other nerves to regenerate," Wysocki says. —*Philip E. Ross*

TECHNOLOGY

Down to the Wire

Coating a fiber with superconductors

Flexible wires made of high-temperature superconductors could lead to such magical applications as superconducting motors, but making such wires is difficult. Now workers at the Georgia Institute of Technology in Atlanta have developed a technique that will enable them to coat fibers with superconducting materials and then to bundle those fibers into wires.

W. Jack Lackey and his colleagues have put a new spin on a common coating technique: chemical vapor deposition (CVD). In conventional CVD, a substance is heated until it becomes a vapor; the gas then flows into another chamber and condenses onto a substrate.

Because reagents for superconducting materials made of the elements yttrium, barium and copper have low vapor pressures, they travel slowly into the substrate chamber, resulting in deposition rates of only about 1 micrometer per hour. The CVD process is also tricky for superconductors because each material must be vaporized separately and then released into the chamber together. This requires workers to control simultaneously the pressure, temperature and flow rate of three gases.

After months of wrestling with CVD, Lackey had an insight: Why not abandon vapors in favor of powder? By grinding the three constituents finely

"Those who have long enjoyed such privileges as we enjoy forget in time that men have died to win them."

FRANKLIN D. ROOSEVELT

Nearly 50 years ago President Roosevelt proclaimed these words when he designated Dec. 15, 1941, as Bill of Rights Day.

This day was to be "...set apart as a day of mobilization for freedom and for human rights, a day of remembrance of the democratic and peaceful action by which these rights were gained, a day of reassessment of their present meaning and their living worth."

Today, as we approach the bicentennial of the adoption of the Bill of Rights, let us all, as President Roosevelt asked of us then, "...rededicate its principles and its practice."

Philip Morris Companies Inc.

KRAFT GENERAL FOODS • MILLER BREWING COMPANY • PHILIP MORRIS U.S.A.

Join Philip Morris in support of the National Archives' celebration of the 200th anniversary of the Bill of Rights. For a free copy of this historic document, call 1-800-552-2222, or write Bill of Rights • Philip Morris Companies Inc. • 2020 Pennsylvania Ave. N.W. • Suite 533 • Washington D.C. 20006

and then mixing them in specific proportions with argon gas, Lackey created an efficient transport system for conveying the materials into the substrate chamber. Heating the chamber then vaporizes the powder and so coats the substrate.

The investigators report they have deposited high-quality layers of superconducting materials at rates of some 200 micrometers per hour. Although to date they have primarily worked with flat substrates, Lackey has designed a CVD system with pulleys that he thinks will make it possible to coat lengths of fibers as they are pulled through the device.

Lackey says that early samples have carried currents as high as 40,000 amperes per square centimeter, enough for some applications. The workers at Georgia Tech now hope to create long wires that will carry high currents. If they are successful, American Magnetics, Inc., in Oak Ridge, Tenn., will employ the wires in a prototype magnet. —*Elizabeth Corcoran*

Heli-hopper

Human-powered helicopter gets off the ground

For the first time in recorded history, a helicopter powered only by human muscle has ... flown? No. Hopped? Better than that. Hovered? Let's be precise: on December 10, 1989, a 100-pound craft called the da Vinci III, built by students at the California Polytechnic State University, rose eight inches off the ground for 6.8 seconds. That was long enough for an observer from the National Aeronautic Association to certify the event as an aviation record.

Polytechnic students launched the da Vinci project—named after the first helicopter designer—almost a decade ago. The da Vinci III consists of a 100-foot-wide rotor driven by a propeller at each tip; a cyclist provides the power.

The machine is more stable and 20 pounds lighter than its immediate predecessor, which failed in several flight attempts two years ago [see "Science and the Citizen," July, 1988]. Another crucial improvement was "a much stronger cyclist," explains Neal T. Saiki, a senior who headed the da Vinci III team.

The students hope to build and test a new-and-improved helicopter by this fall. Their ultimate goal: to win the $25,000 Igor I. Sikorsky prize, which the American Helicopter Society has

offered for the first human-powered machine that can hover at a height of 10 feet for a full minute. —*J.H.*

MEDICINE

Fallout

New radiation-risk estimates prompt calls for tighter controls

A comprehensive reassessment of the dangers of low-level ionizing radiation roughly trebles estimates of the risk of cancer per unit dose and finds no evidence for a safe dose—one below which the risk of carcinogenesis is insignificant.

The study, known as Biological Effects of Exposure to Low Levels of Ionizing Radiation (BEIR) V, comes from an impeccable source: the National Research Council. Based primarily on updated information from survivors of the Japanese atomic bomb blasts and from people exposed to radiation for medical purposes, it replaces a previous assessment the research council made in 1980. The new study—a fifth in a series—concerns only X-ray and gamma-ray radiation; it is already being cited as an argument for lower limits on the exposure of nuclear-industry workers.

The 1980 assessment was highly controversial: at the time, two of its authors issued statements disputing its estimates of cancer risk. The reassessment, developed with different computer models, supports the view that the 1980 study underestimated risks. The new study concludes, for example, that if 100,000 people receive a single exposure of 10 rems, the radiation would cause about 790 extra cancer deaths (in addition to the 18,330 otherwise expected). The risk of cancer—at least for most solid tumors—appears to be proportional to the dose of radiation. The study also increases the estimated risk of mental retardation for children exposed between eight and 15 weeks after conception. The risk of inducing a heritable disease in subsequent generations, however, appears to be lower than was previously thought.

Arthur C. Upton of New York University Medical Center, chairman of the committee that made the new assessment, says it "is not a revolutionary change" and should not greatly concern most members of the public. (The average annual exposure for Americans is less than half a rem, most of it attributable to radon break-

down products; moreover, the new estimates of cancer risk should be at least halved when exposure occurs over a long period.) Upton does think, though, that workers who accumulate relatively high doses perhaps "ought to be watched more carefully." He also stresses that the findings are provisional, because the youngest victims of the atomic bomb blasts have only now reached middle age; risk estimates may be revised up or down until the last victims die.

Energy Secretary Admiral James D. Watkins has instructed his department to assess the BEIR V study to determine whether changes in the department's operations are necessary. The National Council on Radiation Protection and Measurements and the Nuclear Regulatory Commission are also reviewing the new findings. And activist groups, such as the Nuclear Information and Resource Service in Washington, D.C., intend to use the new estimates in their campaigns. The main target is a proposal by the Nuclear Regulatory Commission to deregulate control of very-low-level radioactive waste. —*Tim Beardsley*

Safe Passage?

Study fuels debate over safety of birth centers

The idea has its appeal: having a baby in a homey, caring birth center, away from an impersonal hospital. But are free-standing birth centers as safe as hospitals?

The authors of a large nationwide study published last December in the *New England Journal of Medicine* say "yes," at least for women who have a low risk of a complicated delivery, as is the case for women normally admitted to such centers. (Women who, say, have diabetes or hypertension or are expecting twins are normally referred elsewhere.)

The answer to the safety question hinges largely on whether the centers can cope with unexpected, potentially life-threatening problems, such as prolonged fetal distress or profuse postdelivery bleeding by the mother. The centers, where births are generally attended by midwives rather than obstetricians, are not equipped for such emergency procedures as cesarean deliveries and resuscitation of certain newborns, and so the rare individual who gets in trouble may have to be rushed to a hospital for special care.

To assess the safety of birth centers, Judith P. Rooks of the School of Public Health at Columbia University, Eunice K. M. Ernst of the National Association of Childbearing Centers in Perkiomenville, Pa., and their colleagues followed nearly 12,000 women admitted in labor to roughly half of the known out-of-hospital birth centers in the nation (by some indications, probably the better-run half). As might be expected in any population of women in labor, serious complications that normally require immediate hospital care arose in a number of the deliveries, almost 8 percent of them. Only about half of the affected women and newborns were moved to hospitals, however; the others were not transferred, either because delivery was imminent or because the emergency was resolved by the time the baby was born.

In spite of the need for emergency transfer in some instances and the failure to accomplish the move in many of them, all of the mothers in the study survived, and the number of babies who died before or within a few weeks after delivery was low: 1.3 in 1,000. That rate is similar to rates reported in five studies of women who gave birth in hospitals and were determined retrospectively (by examining their records) to have had a low risk of complications. This similarity in death rates accounts for the investigators' assertion that free-standing birth centers are as safe as hospitals for low-risk populations.

Some physicians dissent from that verdict. In an editorial accompanying the December report, Ellice Lieberman and Kenneth J. Ryan of Brigham and Women's Hospital in Boston say they worry that the birth-center study did not include a prospectively followed control group of women (carefully matched by such characteristics as age and health status) who signed into hospitals or hospital-based birth centers. The two acknowledge that fulfilling their prescription might be difficult, but they contend that without a directly comparable control group, "it is impossible to make reliable inferences about the relative safety of hospitals and birth centers." The low infant death rate in the birth-center study, they say, might derive from having studied a particularly low-risk population.

"It is impossible to do the perfect study that would convince everyone," Rooks says. "This is a very large, well-designed study; the data are strong."

She also points out that no definitive data support the widely held assumption that hospitals are safer. Indeed, she notes, many hospitals cannot perform immediate cesarean deliveries and lack neonatal intensive-care facilities, and many interventions in hospitals can themselves complicate deliveries. "There will be babies that die because the mother goes to a birth center; there will be babies that die because the mother goes to a hospital," she says.

Rooks notes, too, that cesarean deliveries, which are not risk-free for either mother or baby, are much more commonly performed on women who undergo labor in hospitals. In the birth-center study 4.4 percent of the women had cesarean deliveries, half the rate cited for low-risk populations in the retrospective hospital studies that reported on the procedure.

Warren H. Pearse of the American College of Obstetricians and Gynecologists agrees that obstetric practices in certain hospitals need improvement. Yet he is disturbed by the need to transport women or newborns from birth centers to other facilities in emergencies. The number of preventable deaths—or serious health problems—that actually occur because of transport-related delays is undoubtedly extremely small. Nevertheless, because the dangers are real and the possible consequences catastrophic, many physicians, including Pearse, Lieberman and Ryan, prefer hospital-based birth centers to free-standing ones. —*Ricki Rusting*

OVERVIEW

Punctuated Equilibrium

Darwin survives as the debate evolves

In 1972 two paleontologists, Niles Eldredge of the American Museum of Natural History and Stephen Jay Gould of Harvard University, startled—and in some cases dismayed—biologists by suggesting that the view of evolution most of them held was an "insufficient picture." Eldredge and Gould maintained that life was no stately unfolding of gradually changing forms that slowly divided to create new species. Rather, they believed, species formed relatively quickly as a result of rapid bursts of evolutionary change. Eldredge and Gould thought their colleagues were wrong to blame the rarity of intermediate fossil sequences on gaps in the fossil record. Instead they proposed that such sequences were rare because evolution did not happen that way. Eldredge and Gould considered it more likely that

new species evolve within a few thousand years, the mere blink of an eye in geologic time. Once a species has evolved, they argued, it enters a period of stasis, remaining unchanged for possibly millions of years. The two scientists called their theory punctuated equilibrium.

Eldredge and Gould did more than offer a new picture of evolution's progress. They also suggested that Darwin's theory of natural selection acting on individuals was inadequate, because it predicted gradual evolution of species in response to environmental change rather than a "punctuated" pattern. In order to reconcile punctuated equilibrium with the long-term trends seen in fossil lineages—such as the progression from smaller to larger species of mammals—Eldredge and Gould therefore proposed a variant of Darwin's theory. They suggested that natural selection not only acts on the individual level but must also be acting at the species level, an idea since termed species selection. According to this idea, lineages thrive if they either resist extinction or produce frequent daughter species. Gould even wrote that the "synthetic theory"—Darwin's theory of natural selection combined with genetics—was "effectively dead" as an explanation for long-term evolutionary trends.

Not surprisingly, these ideas attracted attention. Popular accounts fostered the impression that Darwin's theory of evolution was wrong. Creationists gleefully played up the argument in their effort to reshape biology teaching, despite Eldredge and Gould's protests that their theory was not meant to question the basic tenets of evolution.

How has punctuated equilibrium fared since it was first proposed almost 20 years ago? Many biologists say Eldredge and Gould attacked a straw theory: evolution has never been viewed as exclusively slow and gradual, they argue, and natural selection is quite able to account for the formation of species over thousands of years, as required by punctuated equilibrium. Ernst Mayr of Harvard University makes the point by quoting Darwin, who wrote in later editions of *On the Origin of Species* that "the periods during which species have been undergoing modification... have probably been short in comparison with the periods during which these same species remained without undergoing any change."

Mayr, who is often regarded as the dean of evolutionary biology, proposed in 1954 that "genetic revolu-

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	**Old Address**
Street	Street
City	City
State and ZIP	State and ZIP

tions" in small, isolated populations may give rise to new species, an idea that Eldredge and Gould have acknowledged as inspirational. Mayr says that Eldredge and Gould have sometimes made "outrageous claims," but he believes they were correct to draw attention to stasis in fossil lineages. Mayr theorizes that such periods—some of which might last for tens of millions of years—are caused by genetic and developmental constraints that limit change.

Jeffrey S. Levinton of the State University of New York at Stony Brook is one of the strongest critics of punctuated equilibrium. He maintains there are "literally scores" of good examples of gradual change. He says Eldredge and Gould remind him of a Scandinavian aphorism: "People come crashing through open doors." How, for example, can the theory be tested when the formation of species cannot be clearly recognized in the fossil record? Levinton concludes that "the totality of the evidence makes it a theory not worth following up."

Antoni Hoffman of the Institute of Paleobiology in Warsaw, who has written a book aimed at refuting punctuated equilibrium, accuses proponents of the theory of creating a moving target. According to Hoffman's analysis, some versions of the theory claim merely that evolution varies in speed—which is undisputed. Later versions claim that gradual change is nonexistent or negligible; that, Hoffman maintains, is "blatantly false."

In 1987 Peter R. Sheldon, then at Trinity College in Dublin, seemed to score a direct hit against punctuated equilibrium when he reported finding that eight types of trilobite had each evolved gradually over a three-million-year interval during the Ordovician period, more than 440 million years ago. The finding prompted John Maynard Smith of the University of Sussex to comment that "we can forget about new paradigms and the death of neodarwinism." But the arguments persist, and each side continues to collect evidence. For example, Adrian M. Lister of the University of Cambridge reported in *Nature* recently that red deer on the island of Jersey underwent a sixfold reduction in their body weight in less than 6,000 years, suggesting that rapid evolution can indeed occur under some conditions.

Gould thinks Sheldon's interpretation of the trilobite record is weak, yet he maintains that paleontologists can discern the formation of new species in the fossil record through comparisons with living species. He admits that the evidence is not yet in and suggests that the technical difficulty of finding adequate data, in the form of well-preserved fossils from undisturbed sedimentary beds, has made for slow progress. While conceding that rapid change is not universal, Gould declares that he and Eldredge will be proved right if rapid species formation and stasis do turn out to dominate in the fossil record.

Where does that leave species selection? Most evolutionists seem to agree with Hoffman that it cannot explain features of individuals, although it might in principle explain some of the long-term trends in evolution. Yet there are no proven examples, and the idea may be, as Hoffman puts it, "an explanation in search of phenomena to explain." Montgomery W. Slatkin of the University of California at Berkeley argues that even if punctuated equilibrium does turn out to be common in evolution, it may nonetheless be driven by natural selection acting on individuals.

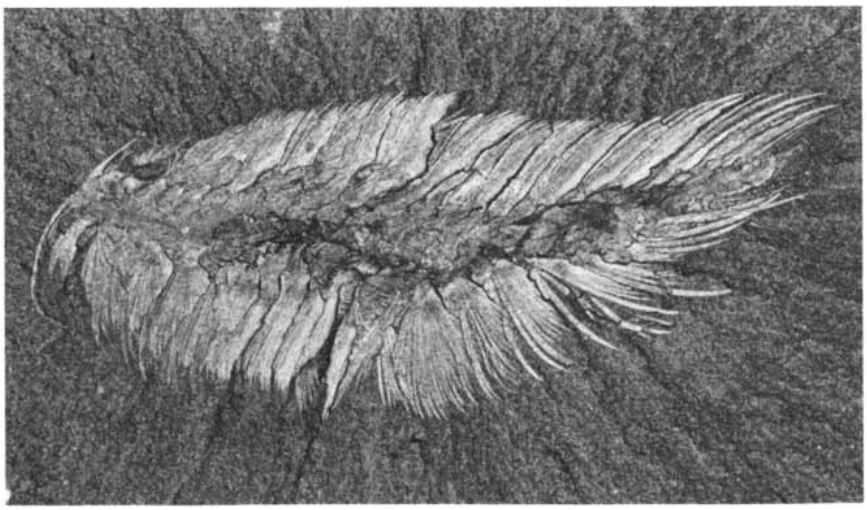
Species selection has even been criticized by a scientist whom Gould lionizes in his latest book, *Wonderful Life*. Gould recounts the story of the Burgess Shale in British Columbia, which contains the remains of bizarre soft-bodied animals that appeared on the earth some 530 million years ago, soon after the first emergence of multicellular creatures at the start of the Cambrian period. Work in recent decades has shown that not only are the Burgess fossils surprisingly complex, but few of them fit into known taxonomic classes.

One of the heroes of Gould's tale is Simon Conway Morris of the University of Cambridge, who painstakingly reconstructed some of the specimens. In a recent article in *Science*, Conway Morris writes that he sees no need for special evolutionary mechanisms (such as species selection) to explain the stunning diversity of the Burgess Shale. He adds that "there is no reason to think that any species did not arise by natural selection."

Gould responds that he finds it hard to imagine natural selection plays no part in the formation of species, but he emphasizes the role of chance in creating the genetic rearrangements that may generate them. Indeed, the importance of chance is the main lesson Gould· draws from the Burgess Shale fossils. He maintains it would have been impossible for any biologist to predict which of the myriad Burgess Shale animals would give rise to later groups: he sees the survival of species as a lottery controlled by historical contingency.

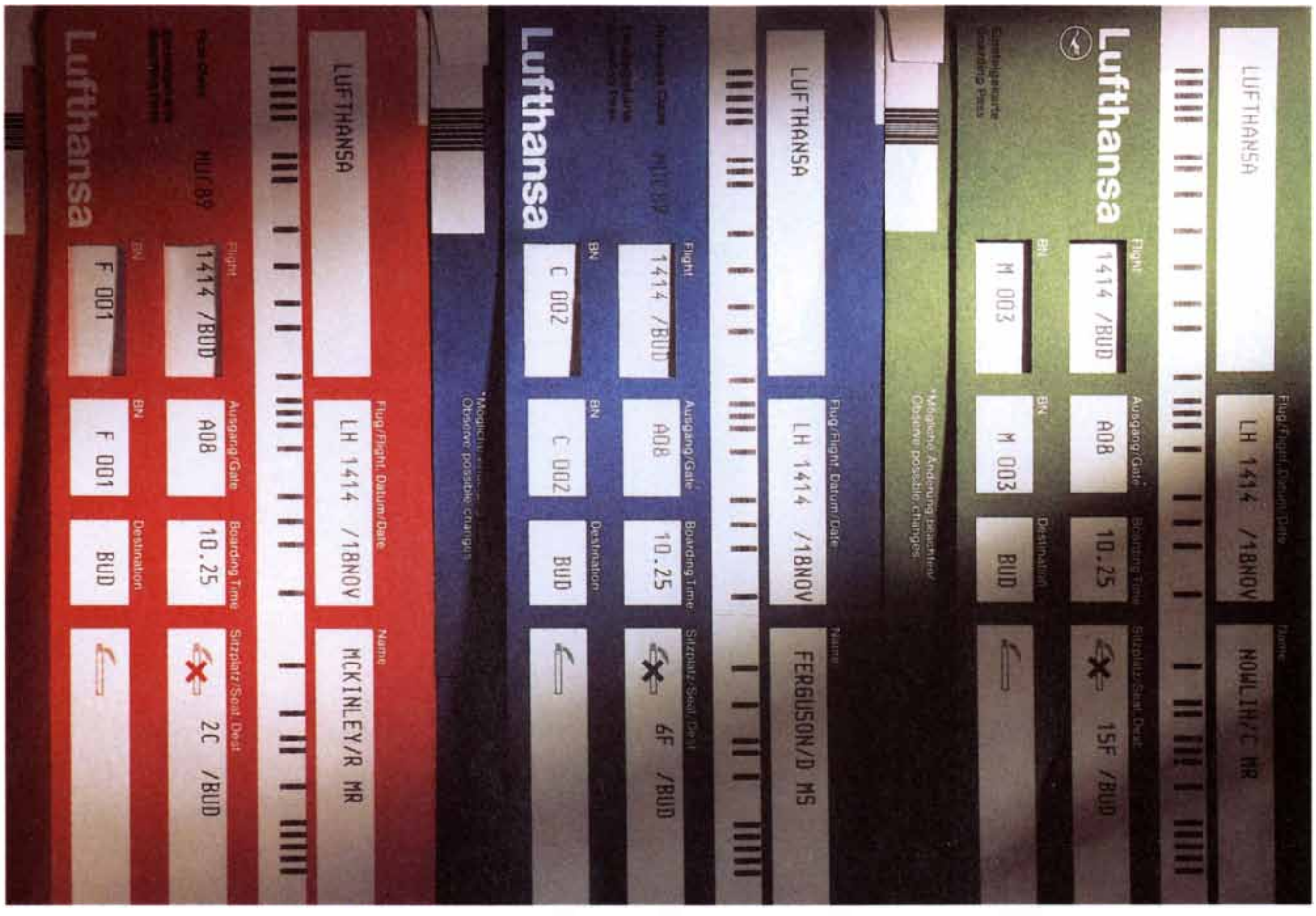
On that point, Gould and his critics agree for once. The implications of a starring role for Lady Luck are profound. Were evolution's drama to be "replayed," the earth's fauna and flora would be radically different, and the human species—that "improbable and fragile entity"—would presumably not exist. —*Tim Beardsley*

Burgess Shale fossils are unlike any others, confounding efforts to classify them



***CANADIA SPINOSA* is one of the strange animals whose remains were found in the Burgess Shale. Stephen Jay Gould thinks the few early types that survived and gave rise to later groups were simply lucky.**

We offer 3 separate classes on every Lufthansa plane, which separates us from just about every other airline.



When you fly to Europe, nearly all airlines offer separate First, Business and Economy Class cabins. But once you get there and take a connecting flight, the standards often change. On most airlines, Business Class mysteriously merges with Economy, or sometimes First Class simply disappears.

Lufthansa believes that service is best when it's also consistent. So every time you fly Lufthansa, you will have 3 distinct classes to choose from, with all the amenities you've grown to expect. Whether your flight lasts half an hour, or half a day, on Lufthansa the quality of service will always remain the same.

People expect the world of us.

Lufthansa

Lufthansa is a participant in the mileage programs of United, Delta, USAir and Continental/Eastern. See your Travel Agent for details.

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.*

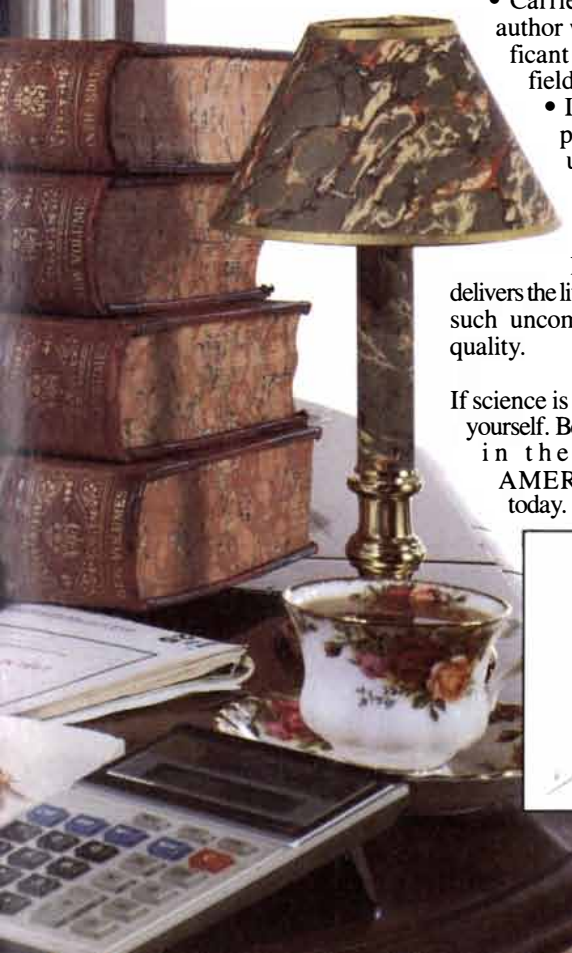
Written by leading scientists in accessible, compelling prose.

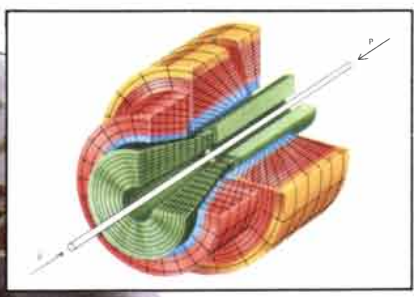
SCIENTIFIC AMERICAN LIBRARY books are created in a unique collaboration between our editors and some of the world's most renowned scientists (including many Nobel laureates). The result is a series of books that are readable *and* authoritative. Each volume...

- 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.

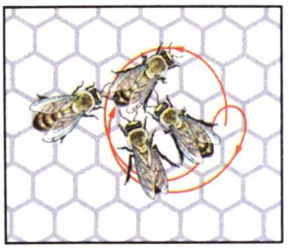
If science is your passion, reward yourself. Begin your membership in the SCIENTIFIC AMERICAN LIBRARY today.





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

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.

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.*

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.

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.

SCIENTIFIC AMERICAN LIBRARY

AN EXCITING NEW GENRE IN THE LITERATURE OF SCIENCE

SCIENTIFIC AMERICAN March 1990 Volume 262 Number 3

Observatories on the Moon

Hostile to life, the moon could be a haven for astronomy. Observatories on its surface could yield extraordinarily detailed views of the heavens and open new windows through which to study the universe

by Jack O. Burns, Nebojsa Duric, G. Jeffrey Taylor and Stewart W. Johnson

The harsh, lifeless surface of the moon may well be the best place in the inner solar system for human beings to study the universe around them. The near absence of an atmosphere, the seismic stability of its surface, the low levels of interference from light and radio waves (especially on the lunar far side) and the abundance of raw materials make the moon an ideal site for constructing advanced astronomical observatories.

Observatories on the moon will exceed the resolving power of current ground-based optical instruments by an extraordinary amount, perhaps by a factor greater than 100,000. Such observatories will also open an entirely new window on the universe by permitting detection of very low radio frequencies; they may even help pioneer new branches of astronomy through the study of gravitational waves and the elusive neutral particles known as neutrinos.

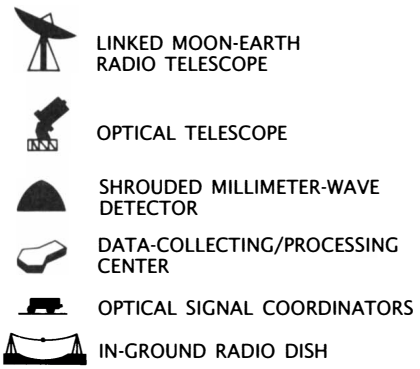
In the U.S. there is a renewed interest in the moon as a science platform, as a resource base or as a stepping-stone to Mars. In 1983 a National Academy of Sciences report, *Astronomy and Astrophysics for the 1980's,* advocated international planning for lunar-based astronomy. On July 20, 1989, President George Bush voiced his support for a permanently manned lunar base in his speech commemorating the 20th anniversary of the first Apollo landing. We and others in the scientific community have therefore begun serious efforts to plan for a permanent lunar base and for astronomical observatories that might be built on the moon in the 21st century.

Until then, improved astronomical observations will come from satellites in earth orbit. Four major scientific satellites—the *Great Observatories*—are expected to be launched by the close of the century. These satellites will examine a large fraction of the electromagnetic spectrum, which includes light as well as radio waves, X rays and gamma rays. Virtually everything that astronomers know about the cosmos has been learned by studying electromagnetic emissions from celestial objects. Each section of the electromagnetic spectrum yields information about the specific physical processes that produce radiation of that particular energy, or frequency.

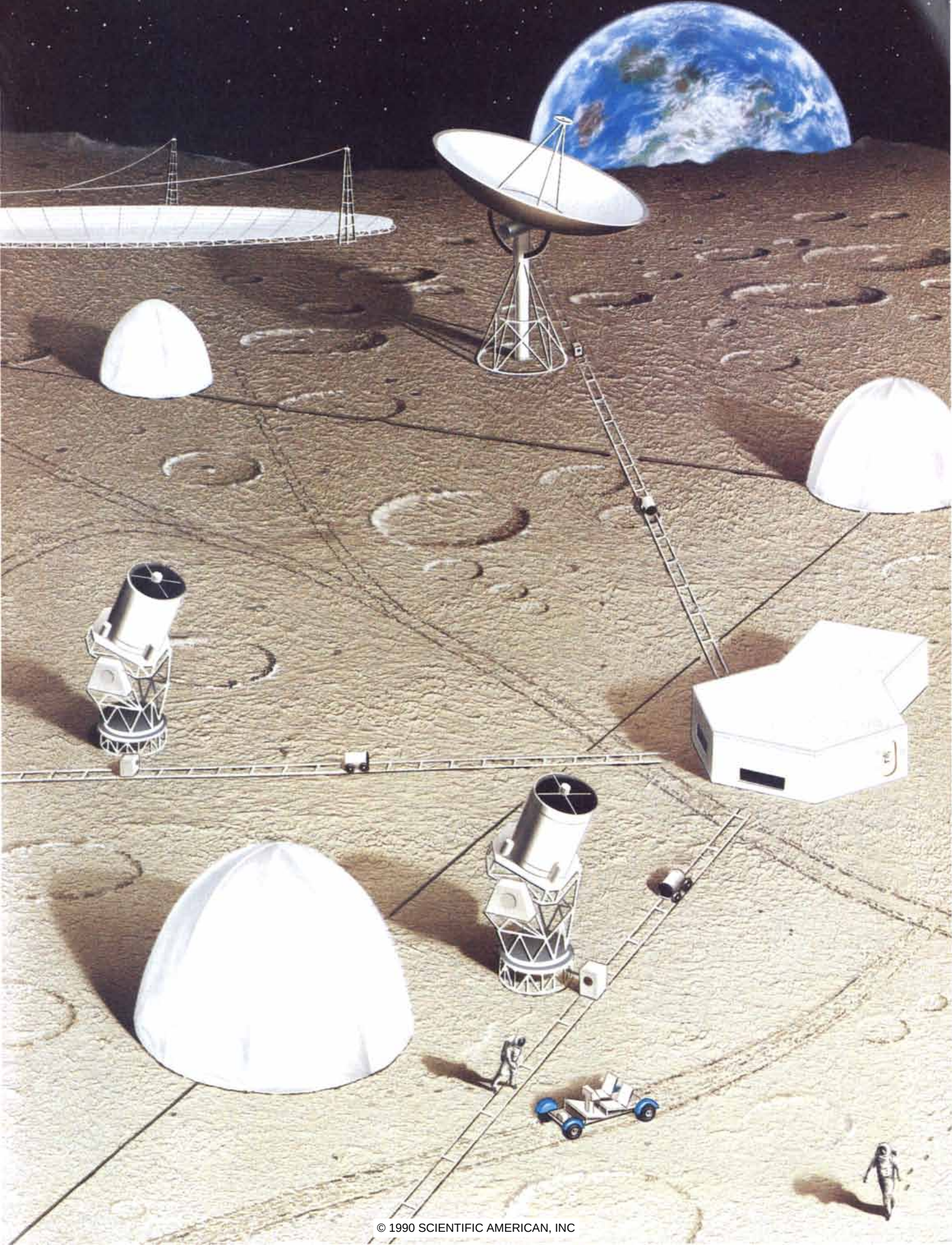
The *Hubble Space Telescope* will probe the universe at visible wavelengths with nearly 10 times the resolution of the largest ground-based telescopes when it is sent into space this year. The *Gamma Ray Observatory* is also slated for a 1990 launch; it will view radiation emitted by energetic processes that occur around dense, compact objects, such as neutron stars and black holes, with vastly improved sensitivity and resolution as compared with previous instruments.

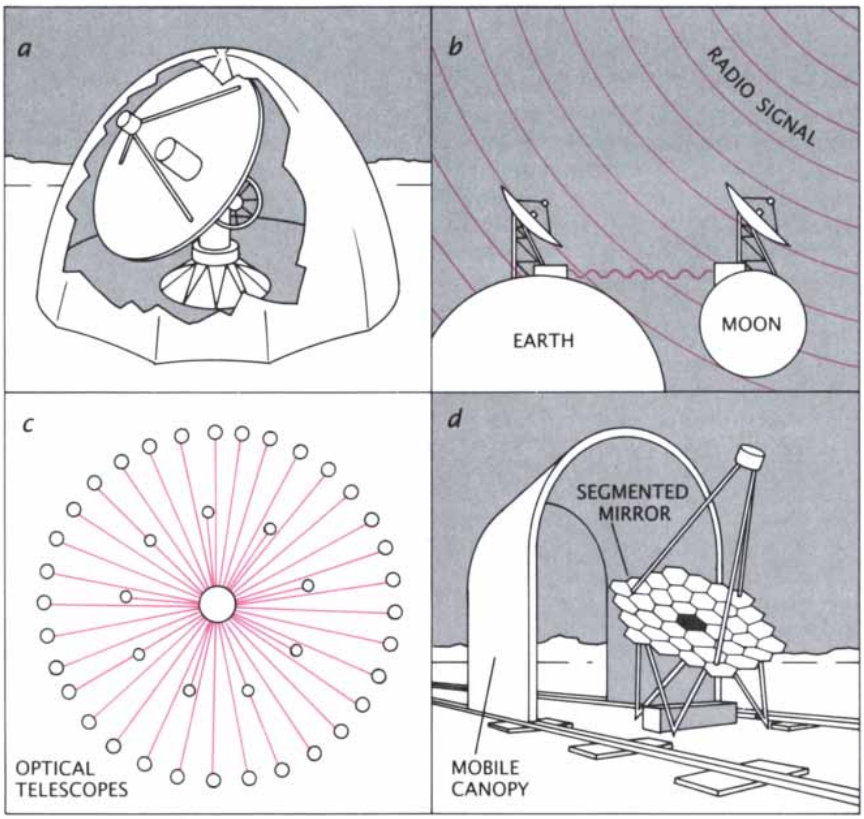
The *Advanced X-ray Astrophysics Fa-*

JACK O. BURNS, NEBOJSA DURIC, G. JEFFREY TAYLOR and STEWART W. JOHNSON share an interest in astronomy and exploring new frontiers. Burns is professor of astronomy and head of the astronomy department at New Mexico State University and has been designing and evaluating plans for lunar-based observatories for the past six years; he has a Ph.D. in astronomy from Indiana University. Duric is director of the Institute for Astrophysics and assistant professor of astronomy at the University of New Mexico. He received his Ph.D. in astronomy from the University of Toronto. Taylor is senior research scientist and assistant director of the Institute of Meteorics at the University of New Mexico and holds a Ph.D. in geology from Rice University. Johnson is adjunct professor of civil engineering at the University of New Mexico and is also principal engineer for advanced basing systems at BDM International. He earned a Ph.D. in civil engineering at the University of Illinois at Urbana-Champaign.



LUNAR OBSERVATORIES will greatly improve on existing facilities. Optical telescopes, unfettered by a distorting atmosphere, will be linked to a central facility to provide superior resolution. An array of detectors will map cosmic processes that emit millimeter-wavelength radio waves. Linking lunar radio telescopes with ones on the earth will permit high-resolution observations. Huge radio dishes built into craters will probe faint signals from deep-space structures.





TELESCOPES will be designed specifically for the lunar environment. Shrouds will protect delicate surfaces, such as the collecting dish of this millimeter-wave telescope, from meteoroid impacts (*a*). A radio telescope on the moon linked to one on the earth will have the resolving power of a single 400,000-kilometer-wide antenna (*b*). The authors envision linking 42 optical telescopes to form a 10-kilometer-wide array; signals received by the array would be collected and processed at a central facility (*c*). A proposed 16-meter telescope has a segmented mirror and simple mount to ease construction; a mobile canopy protects the optics (*d*).

cility, planned for a 1996 launch, will image hot gas in galactic clusters, in the environs of active galaxies and quasars and in the atmospheres of stars. Finally, the *Space Infrared Telescope Facility* will map cool interstellar gas clouds that are associated with the formation of stars and possibly planets; it is tentatively scheduled for a 1998 launch.

In spite of their considerable promise, we think the *Great Observatories* will likely be the last major astronomical instruments placed in low earth orbit. This is because at the low altitudes (500 to 600 kilometers above the earth) where most astronomical satellites orbit, telescopes encounter serious problems. First, the environment is dirty and the cloud of orbiting debris from space launches grows thicker each year. Much of the debris consists of high-velocity, micrometer-size objects that can seriously damage a telescope's sensitive optics and instrumentation.

Second, there is still a substantial amount of dust and gas at this altitude. The dust scatters light and creates an infrared background that masks emissions from the faintest celestial infrared sources. The high-velocity orbital motion of satellites excites atoms in the rarefied upper atmosphere, creating emission lines at optical wavelengths that can obscure or confuse astronomical observations.

Third, atmospheric drag causes satellites in low earth orbit to spiral slowly back toward the earth. The problem is exacerbated during times of high solar activity, when the increased intensity of solar ultraviolet radiation causes the earth's upper atmosphere to swell. Intense levels of solar activity have recently forced the National Aeronautics and Space Administration to reevaluate and raise the planned orbit for the *Hubble Space Telescope*.

Fourth, the earth itself is a major source of obscuring radiation. Light reflected off the earth can scatter into telescope optics and degrade the quality of the observations. Furthermore, the earth's magnetic field generates low-frequency radio noise that far exceeds the intensity of extraterrestrial sources. As a consequence, observatories in low orbit will have difficulty detecting and imaging low-frequency, kilometer-wavelength radio waves, the sole section of the electromagnetic spectrum that astronomers have not yet observed.

Finally, satellites orbiting close to the earth experience rapid thermal and gravitational changes, which limit the size, and hence the resolution and sensitivity, of telescopes that can be placed in such orbits. These changes deform mirrors and radio dishes, preventing them from producing optimal images and wasting valuable observing time while the instrument stabilizes. Eliminating thermal strains necessitates the construction of elaborate masks or sunshades.

Substantial improvements in the next generation of space observatories will require a better location from which observations can be made. One solution is to launch satellites into higher, geosynchronous orbits (where one orbit requires 24 hours), about 37,000 kilometers above the earth. Another is to place instruments on the surface of the moon (which orbits 384,000 kilometers from the earth).

Transportation costs increase with growing distance from the earth, but the *rate* of increase declines with distance: traveling from the earth to geosynchronous orbit requires 2.6 times the amount of fuel needed to travel from the earth to low earth orbit, but only an additional 50 percent is needed to reach the moon. More significantly, the environment for astronomy greatly improves with distance; the moon is an optimal location for astronomical observations.

One major advantage of the moon is that it provides an enormous natural platform on which large structures can be built. Large-aperture telescopes and huge, electronically linked arrays of smaller telescopes on the moon will lead to major improvements in the detail of astronomical observations. Signals from a pair of linked telescopes are said to "interfere," and so such linked devices are known as interferometers. Through the careful combination of the signals, it is possible to achieve the same resolving power from a pair of small telescopes as from a single telescope whose diameter equals the distance between the pair. In earth orbit, interferometers would require giant platforms or complex and costly station-keeping of all the telescopic ele-

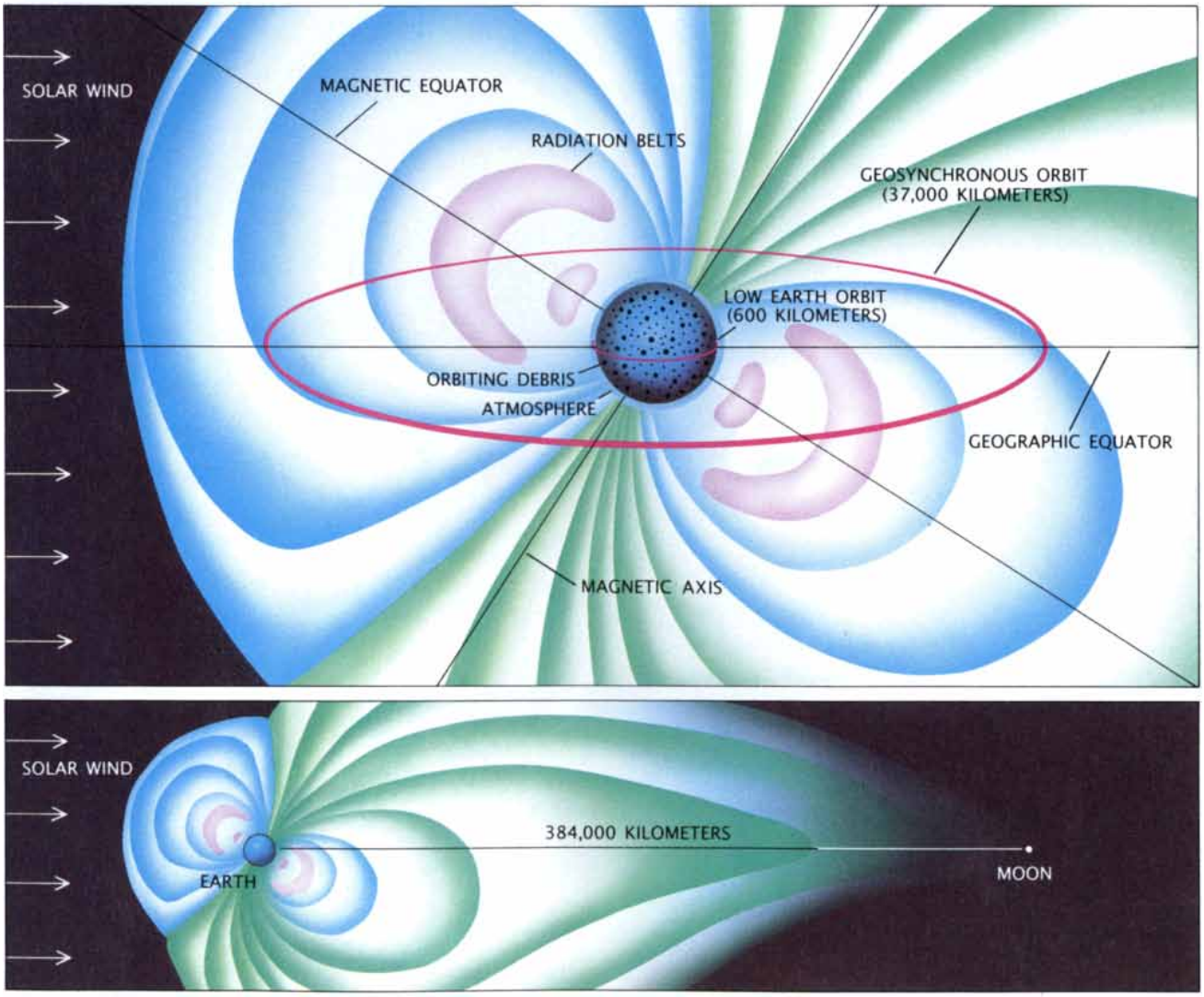
ments; on the moon, such instruments simply could be placed on the surface.

Fabricating structures in the reduced-gravity environment of the moon will be easier than in the zero-gravity environment of earth orbit, as the Apollo and space-shuttle missions have revealed. Construction of observatories on the moon can be adapted from techniques developed on the earth, with the advantage that the moon's weaker gravitational pull makes it possible to build larger devices than are practical on the earth. Ferhat Akgul and Walter H. Gerstle of the University of New Mexico and one of us (Johnson) have drawn up plans for a lunar radio telescope made of advanced graphite-epoxy materials. They foresee no practical obstacles to building fully steerable radio-collecting dishes 500 meters across or 16-meter-wide mirrors for observing visible light and infrared radiation (which has somewhat longer wavelengths than visible light) on the moon.

The moon is a remarkably stable platform: the energy of a typical seismic event on the moon is 100 million times less than an average one on the earth, and moonquakes produce ground motions of only about one billionth of a meter. Such stability is a boon for optical interferometers, which require that the distance between telescopic elements in a linked system be known to within a fraction of a wavelength of light, or roughly one ten-millionth of a meter.

Another feature of the moon is that it offers a clear view of virtually all astronomically interesting radiations. The moon's atmosphere is essentially nonexistent: the mass of the entire lunar atmosphere is equal to that of the air inside a typical basketball stadium on the earth! As the moon orbits, one hemisphere always faces away from the earth. On the far side, the mass of the moon serves as a giant wall, blocking the natural and man-made sources of interference that make low-frequency radio observations impossible on the earth and difficult in earth orbit. Indeed, the far side of the moon is the only location within the inner solar system that is both large enough and sufficiently shielded to make it possible to study these frequencies.

There are other reasons why the lunar environment is ideal for astronomy. James D. Burke of the Jet Propul-



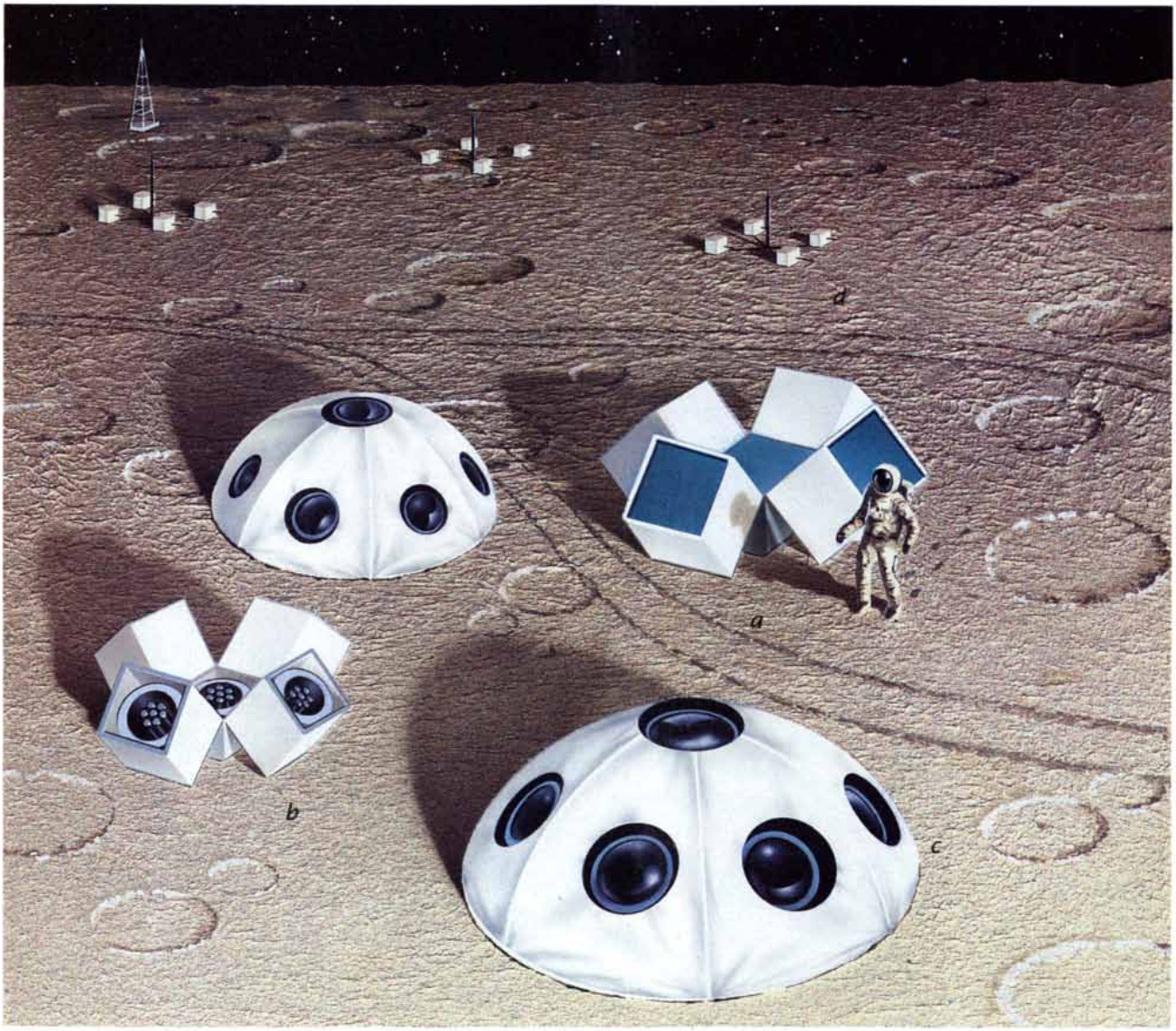
DIRT AND NOISE plague earth-orbiting astronomical satellites. Human-generated debris surrounding the earth threatens satellites with potentially devastating collisions. The thin upper atmosphere creates friction that generates unwanted light and causes satellites' orbits to decay. Even at higher altitudes, light from the earth degrades astronomical observations. The earth's radiation belts create radio noise and interfere with instrument operations. The prospects for astronomy are much brighter on the moon: it lies far from the troublesome environment surrounding the earth, and it is large enough to act as a shield, protecting observatories on its far side from natural and human-generated radiation emitted by the earth.

sion Laboratory in Pasadena, Calif., notes that some craters near the moon's poles appear to be constantly shadowed; the temperature in these locations is believed to be about 70 kelvins. Most astronomical detectors have to be cooled to or below this temperature to reduce the thermal noise that is always present in instrument electronics. Daniel F. Lester of the University of Texas at Austin suggests that the lunar poles may be ideal locations for infrared telescopes; the frigid environment would passively cool not only the electronics and infrared-sensitive detectors but also the telescope structures, which at higher temperatures are themselves major sources of infrared radiation.

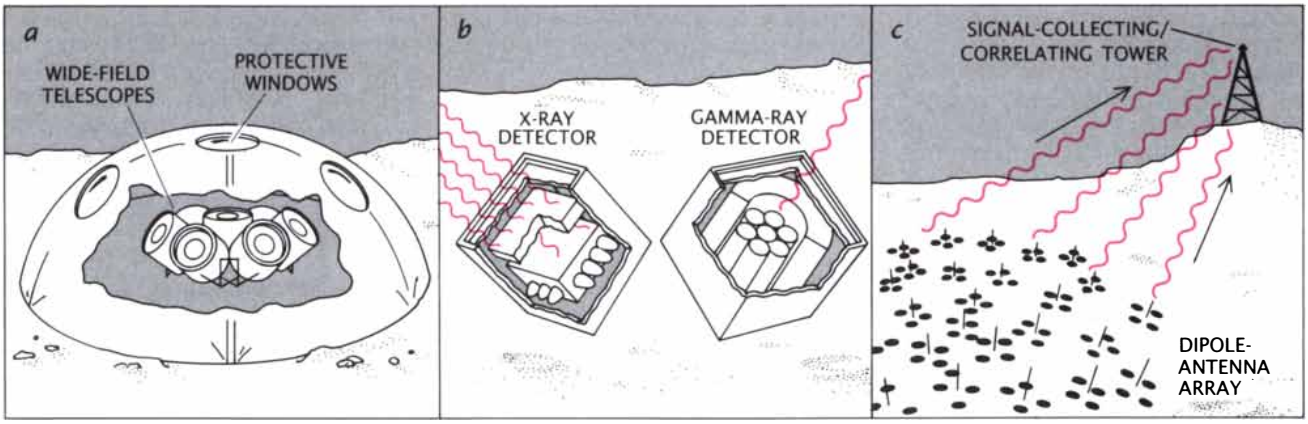
The moon has an abundance of raw materials that could be processed to yield aluminum, ceramics and superstrength glasses for the construction of telescopes. James D. Blacic of the Los Alamos National Laboratory notes that the extremely dry lunar environment will make it possible to produce glasses with tensile strengths comparable to that of steel and with low coefficients of thermal expansion. Such glasses could be used not only for optical mirrors but for the telescope structures as well. Mining, refining and processing on the moon could also serve commercial purposes and become an integral part of a permanently manned lunar base. By the middle of the 21st century, few components of astronomical telescopes may need to be brought from the earth.

There are, of course, difficulties associated with placing a base on the moon. The earth has a magnetic field that deflects charged cosmic-ray and solar-wind particles away from its surface and from near-earth space. There is no comparable field on the moon, and so precautions must be taken to protect human beings and sensitive electronics from hazardous radiation exposure.

The moon is also constantly bombarded by small meteoroids. On the earth, these particles disintegrate in the upper atmosphere; on the airless moon, they rain down on the surface



NEW VIEWS OF THE UNIVERSE will result from lunar astronomy. X-ray variability monitors (*a*) and gamma-ray burst detectors (*b*) will scan the skies for mysterious, rapidly blinking sources of high-energy radiation. Wide-field optical and ultraviolet telescope arrays, seen here shrouded (*c*), may at last reveal the true nature of these sources. Dipole-antenna clusters (*d*) linked together will form the Very Low Frequency Array (VLFA), which will detect low-energy radio waves emitted by solar flares, supernova remnants and active galaxies. Such radio waves cannot be observed from the earth's surface.



ADVANCED TECHNOLOGIES will help exploit the potential of lunar-based astronomy. Full-sky monitoring arrays will be highly automated (*a*). Gamma- and X-ray monitors (*b*) will incorporate advanced detectors made of germanium and special masks to improve resolution. Self-guided robots will help place the VLFA's roughly 200 antennas on the moon's far side; each antenna will be linked to a central data-collecting facility via optical fibers or sophisticated radio transmitters (*c*).

at velocities ranging from 10 to several tens of kilometers per second. Analysis of data from the Apollo missions suggests that tiny craters from one micrometer to 10 micrometers wide will be common on exposed surfaces on the moon. Sensitive surfaces—optical mirrors, for example—will need to be protected with domes or tubes.

The moon's slow rotation (once every 27.3 days) and lack of an atmosphere create drastic and rapid temperature variations—from 100 kelvins at night to 385 kelvins during the day. Provisions must be made to minimize the thermal strains that would afflict lunar telescopes. Analyses by Akgul, Gerstle and Johnson suggest that a judicious choice of composite materials (such as graphite-epoxy) that resist expansion and contraction may prevent thermal changes from being a serious problem.

A vigorous base on the moon—involving mining, manufacturing and regular launch activities—could provide important resources but could also pollute the lunar atmosphere. Richard R. Vondrak of Lockheed has noted that, under extreme circumstances, mining activities could generate a long-lived lunar atmosphere opaque to ultraviolet radiation that would interfere with lunar astronomical activities.

Fortunately, analyses by Ilias Fernini of the University of New Mexico and us indicate that lunar pollution would be largely limited to the vicinities of mining or launch facilities. Atmospheric atoms quickly would be adsorbed by the lunar soil or swept away by the solar wind (the stream of particles from the sun). As a result, observatories more than 10 to 100 kilometers from mining operations should be largely unaffected by atmospheric pollutants.

The first lunar observatory will probably be modest in scope. Yet even a simple one-meter-diameter optical telescope on the moon could perform valuable studies; in the absence of a distorting atmosphere, such a telescope would have a resolution of about .1 arc second, better than the largest earth-based instruments.

Telescopes on the moon will have access to ultraviolet and infrared radiations (which have wavelengths somewhat shorter and longer than visible light, respectively) and so could study a wide range of astronomical phenomena. A lunar-based telescope designed to give a wide field of view would be ideal for producing a map of the entire sky that would show fainter objects and finer details than the Palomar Sky Survey, the current standard reference. Such a telescope also could be used to monitor solar activity and warn astronauts of impending solar cosmic-ray bursts.

Michael Zeilik of the University of New Mexico proposes dedicating a one-meter optical telescope to monitoring changes in the brightness of variable stars and quasars. Freed from the earth's bright and turbulent atmosphere, lunar-based brightness measurements would require less than 1 percent of the exposure time necessary on the earth, and they could be conducted continuously, without the short day-night cycles and the weather interruptions that hinder such observations on the earth.

Other fairly simple instruments well suited to the lunar environment include gamma-ray burst detectors and X-ray variability monitors. These devices could search the sky for the brief, enigmatic bursts of gamma-ray emission (lasting from .01 second to 80 seconds) that rise hundreds or thousands of times above the quiescent background. Little is known about gamma-ray bursters, primarily because astronomers have been unable to locate and identify optical counterparts to the bursts. Typical gamma-ray detectors can locate objects only to within a few degrees. A gamma-ray monitoring station on the moon, coupled with others scattered through the solar system, could pinpoint these sources with arc-second precision by comparing the arrival times of the bursts at the various detectors. X-ray variability monitors could study the flickering emissions thought to be associated with hot disks of matter surrounding neutron stars and black holes.

One particularly compelling reason for building a lunar observatory is that it would enable astronomers to study low-frequency radio waves (those with a wavelength of about one kilometer), which are observable only from the moon's far side. During the past two decades, as new technologies have made it possible to observe celestial X rays and infrared radiation, new classes of sources and surprising physical phenomena have been discovered. Radio frequencies below about 30 megahertz are the last unexplored section of the electromagnetic spectrum. These frequencies are inaccessible from the earth's surface because the upper atmosphere reflects the incoming radiation.

This tantalizing window on the cosmos could be opened by constructing the lunar Very Low Frequency Array

(VLFA), a concept proposed by James N. Douglas and Harlan J. Smith of the University of Texas at Austin in 1985 and developed by us. The VLFA would consist of about 200 dipole antennas, each resembling a television reception antenna about one meter in length. The antennas would be distributed in a circular pattern over a 20-kilometer-wide area and would be sensitive to frequencies between about 50 kilohertz and 30 megahertz. Groups of antennas would be electronically coordinated to "point" the array at different sources without physically moving the elements. Signals from each group would be relayed to a central location for correlation and calibration; the data could then be analyzed by computers on either the moon or the earth to reveal the surface brightness (and hence the structure) of the source producing the radiation.

Building the VLFA will pose several technological challenges. The array would be sited on the lunar far side, quite distant from any likely manned base, and so probably would be deployed remotely. Robotic vehicles will need to be designed; they must be able to traverse a 20-kilometer-wide area of varied terrain and place antennas at optimal locations. Because of the communications time delay either to the near side of the moon or to the earth (about three seconds to and from the earth), the vehicles will need to incorporate self-contained artificial-intelligence programs. Transmitting the data from each dipole to a central processing station will necessitate developing a sophisticated multiplexed radio or laser transmitter-receiver.

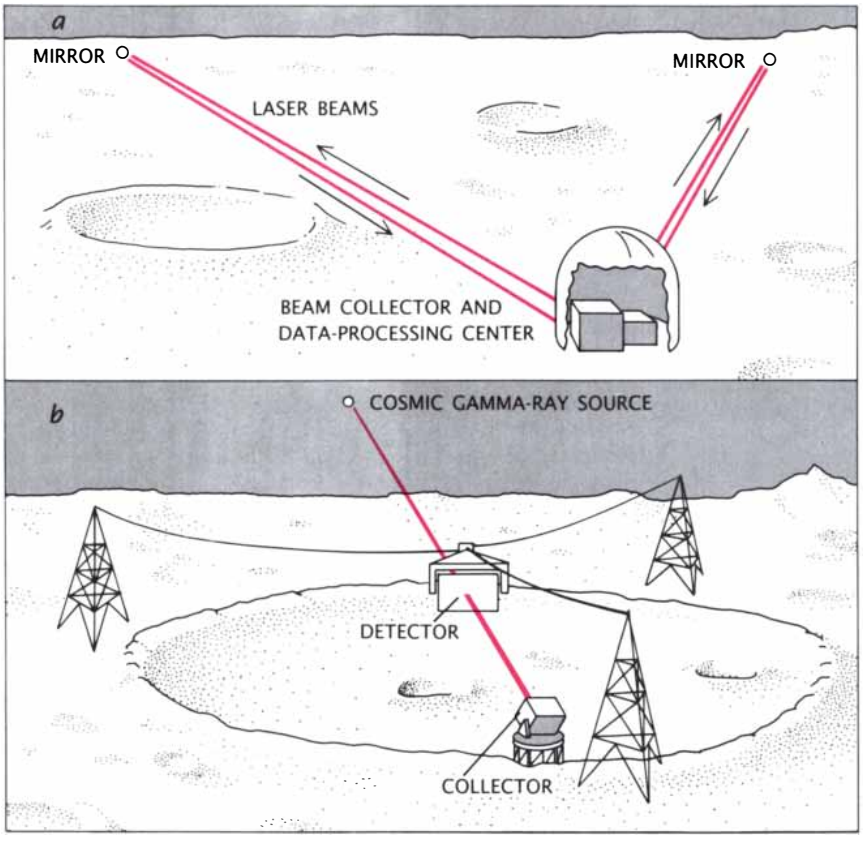
Once these challenges are overcome, VLFA can be expected to yield a scientific bonanza. Astronomers will be able to study the processes that accelerate high-energy particles in solar flares, making it possible to produce an early-warning system for energetic solar outbursts; this would complement the work of a one-meter optical solar monitor. Researchers will be able to study the flux of high-energy electrons in planetary magnetic fields, in supernova remnants, in pulsars and in extragalactic radio sources that are the sites of powerful but still poorly understood astrophysical processes. The VLFA will be able to probe the structure of the interplanetary and interstellar media. Finally, low-frequency observations will allow astronomers to glimpse the subtle low-energy processes that accompany the more visible behaviors of active galaxies and quasars.

Pierre Bely of the Space Telescope Science Institute in Baltimore and Garth D. Illingworth of the Lick Observatory recently have outlined a concept for a 16-meter optical-infrared telescope on the moon. Such a telescope would be able to detect objects 40 times fainter than those accessible to the space telescope. The 16-meter-wide light-collecting mirror would be made up of several hexagonal segments, which would ease construction and assembly. The telescope would be supported by a simple, lightweight mount consisting of three pairs of legs; computer-controlled actuators in the legs would compensate for shifts in the telescope's foundation. J. Roger P. Angel of the University of Arizona has suggested that the telescope may be able to detect earthlike planets around other stars by searching for the characteristic emission signal of atmospheric ozone.

A far more ambitious project for optical astronomy would be to place a large optical interferometer on the moon. The eagerly anticipated *Hubble Space Telescope* will provide only a tenfold improvement in resolution as compared with ground-based telescopes. The telescope we propose, called the Lunar Optical-Ultraviolet-Infrared Synthesis Array (LOUISA) will improve on the resolution of the largest ground-based telescope by a factor of 100,000. Such an instrument could in principle resolve a dime on the earth from its vantage on the moon!

Bernard F. Burke of the Massachusetts Institute of Technology first proposed building an instrument of this kind; we envision LOUISA as a series of electronically linked optical telescopes placed in a circular array 10 kilometers wide. An optical interferometer of this kind cannot be built on the earth because of atmospheric turbulence and seismic motions of the earth's crust. In low earth orbit, the gravitational gradients over a 10-kilometer baseline would necessitate constant, highly sophisticated and expensive repositioning of each of the telescope elements.

In February, 1989, some 50 of the leading researchers in the field of optical interferometry gathered at a workshop at the University of New



ASTRONOMICAL FRONTIERS will be rolled back by lunar observatories. Gravity waves are thought to exist but have never been clearly detected. A laser distance-measuring facility (*a*) on the moon's highly stable surface will search for distortions that might arise from these elusive waves. Long-focal-length gamma-ray telescopes (*b*), possible on the airless moon, will view energetic processes in supernovas and distant galaxies with far greater sensitivity than existing instruments.

Mexico and settled on a preliminary design for LOUISA. The array would consist of two concentric rings of telescopes with 1.5-meter-diameter mirrors. The outer ring would contain 33 telescopes and would be 10 kilometers in diameter; the inner ring would contain nine and be half a kilometer in diameter. Light collected by each telescope would be transmitted to a central station for processing and storage. LOUISA would observe a broad wavelength range from the ultraviolet (.1 micrometer) to the near infrared (one micrometer).

LOUISA is the most technically challenging of the telescopes being considered for construction on the lunar surface, but it potentially has the largest scientific payoff. With a resolution of about one hundred-thousandth of an arc second, LOUISA will open entirely new classes of problems for study. LOUISA will be capable of detecting possible earthlike planets orbiting nearby stars (and perhaps determining their atmospheric composition); this is the first step in ascertaining the feasibility of life elsewhere in our galaxy. In our own solar system, LOUISA could provide images of planets and asteroids with a level of detail exceeding even the views provided by the Voyager spacecraft. The surface features of stars could be viewed directly, thereby revealing the connection between activity on the sun and on other stars and making it possible to study subtle motions of stellar surfaces; these motions yield information about the stars' internal structures and provide clues to their evolution. Astronomers will be able to explore the dynamics of galactic nuclei, view disks of matter spiraling into black holes and neutron stars and probe the structures of disrupted galaxies. In the area of cosmology, LOUISA will measure the proper motion of quasars and reveal any unevenness in the expansion of the universe.

Other lunar telescopes that would offer vast improvements over current instruments have also been proposed. A moon-earth radio interferometer would extend the technique of interferometry to a baseline stretching from the earth to the moon, essentially creating a 384,000-kilometer-wide radio telescope. Such a device would have a resolution of about one hundred-thousandth of an arc second, at a frequency of 10 gigahertz.

Frank D. Drake of the University of California at Santa Cruz has proposed constructing a giant radio-dish antenna, perhaps 1,500 meters wide, in a lunar crater. An antenna this large could conduct high-sensitivity observations of radio emissions from neutral hydrogen in space; such observations may soon become impossible on the earth because of the increasing levels of human-generated interference. Paul Gorenstein of the Harvard-Smithsonian Center for Astrophysics has proposed constructing a clustered array of X-ray telescopes on the moon that would improve significantly on the resolution and sensitivity of the *Advanced X-ray Astrophysics Facility.*

The construction of an observatory on the moon also may help expand astronomy into the largely unexplored realms of gravity waves and neutrinos. Gravity waves are predicted by Einstein's theory of general relativity; the detection of gravity waves would provide important tests of his theory as well as of models of galaxy formation.

Gravity waves are believed to be generated by the collapse of the stellar core that occurs during a supernova explosion and possibly by exotic high-energy processes relating to so-called cosmic strings that may have formed in the early universe. The predicted intensity of these waves is so low that sensing them will likely require highly stable detectors more than 1,000 kilometers long. The moon's seismic steadiness and its large surface area make it an attractive location for a gravity-wave detector.

Neutrinos are extremely abundant and may in fact be the dominant form of matter in the universe. They may also hold the key to understanding the thermonuclear reactions in the sun, which do not seem to behave as predicted. Neutrinos are difficult to detect because they interact extremely weakly with matter. On the earth, observing neutrinos from astronomical sources is further complicated by the constant showers of high-energy particles (including neutrinos) that form when cosmic rays crash into the atmosphere. Michael Cherry and Kenneth Lande of the University of Pennsylvania have noted that the neutrino background on the moon is less than one thousandth that on the earth, making the moon a superior location for a neutrino detector.

Two recent reports, *Pioneering the Space Frontier* and *Leadership and America's Future in Space* (often referred to as the Sally Ride report), advocate the establishment of a permanent lunar base as one of several desirable goals for the U.S. space program. The reports recognize the great scientific and resource potential of the moon. Learning to live and work on an alien world relatively close to the earth is also an important prerequisite to embarking on more ambitious and challenging manned missions, such as those to Mars.

Sizable lunar construction projects such as VLFA, LOUISA and large radio dishes will undoubtedly stimulate major advances in facilities engineering. Moreover, a lunar observatory will promote integrated, collaborative efforts linking astronomy, physics and electro-optics, as well as various branches of engineering.

As the level of East-West tension eases, the U.S. will presumably spend a declining fraction of its gross national product on defense; yet the U.S. must maintain its technological readiness. Lunar astronomical facilities could help the U.S. create a clearly stated, coordinated long-range plan for technology development and application. A lunar base would also provide a prime opportunity for international cooperation.

Political and economic considerations offer solid reasons for the U.S. to choose a return to the moon as a high-priority goal for the 21st century. Even if a lunar base is established primarily for nonscientific reasons, its existence will make construction of lunar observatories relatively simple and inexpensive. By analogy, the Apollo missions may have been spurred by political motivations, but they nevertheless returned a wealth of scientific data. It is time to take advantage of the staggering astronomical potential that could result from a permanent return to the moon in the 21st century.

For ages humans have looked up at the moon and dreamed of reaching it. Soon we may find ourselves gazing up from the moon itself in order to explore the universe as never before.

FURTHER READING

LUNAR BASES AND SPACE ACTIVITIES OF THE 21ST CENTURY. Edited by Wendell W. Mendell. Lunar and Planetary Institute, 1985.

ENGINEERING, CONSTRUCTION, AND OPERATIONS IN SPACE. Edited by Stewart W. Johnson and John P. Wetzel. American Society of Civil Engineers, 1988.

FUTURE ASTRONOMICAL OBSERVATORIES ON THE MOON. Jack O. Burns and Wendell W. Mendell. NASA Conference Publication 2489, March, 1988.

EXTRATERRESTRIAL FACILITIES ENGINEERING. Stewart W. Johnson in *Encyclopedia of Physical Science and Technology, 1989 Yearbook*. Edited by Robert A. Meyers. Academic Press, 1989.

Interleukin-2

The first hormone of the immune system to be recognized, it helps the body to mount a defense against microorganisms by triggering the multiplication of only those cells that attack an invader

by Kendall A. Smith

Ever since Edward Jenner introduced vaccination in 1797 as a means of preventing smallpox, biologists have been fascinated and mystified by the immune system. During the past two centuries, investigators have slowly learned that the activity of the immune system depends on a symphony of highly specialized types of cells in the blood and tissues, each cell type performing a unique function. Yet as the roster of immune cells grew, it remained unclear how this unconnected multitude of cells was able to orchestrate its activities into a selective defense against diseases. Consequently, until recently, immunity was a mysterious phenomenon, and immunology seemed hopelessly complex and esoteric even to biologists in other disciplines.

Within the past 10 years, however, immunology has been transformed by the demonstration that the immune system is regulated by hormones in much the same way as are most other organ systems. Today immunology is no longer a science set apart by a specialized vocabulary and array of mechanisms. Instead the discovery and characterization of the interleukins, the hormones of the immune system, have made it clear that the immune response operates according to the same principles governing the classic hormones and their receptors.

Interleukin-2 (IL-2) was the first of the new immunologic hormones to be discovered and characterized. Although eight interleukins are now recognized by immunologists, the functions of some of these are undetermined or are not connected directly to immunity; IL-2, in contrast, is pivotal for the generation of an effective immune response. Accordingly, an understanding of IL-2 and its receptors opens the way for the development of therapeutic approaches to a wide range of conditions, including cancer, autoimmune disorders, chronic infectious diseases, acquired immunodeficiency syndrome (AIDS) and organ-transplant rejection.

An understanding of the role of IL-2 has done much to explain the specific properties of the immune system. Just as the nervous system senses and responds to changes in light, sound and other environmental stimuli, so the immune system senses and responds to the invasion of molecules in the form of foreign microorganisms, such as bacteria, viruses, fungi and parasites. To accomplish its tasks, the immune system combines three novel characteristics. First, immune reactivity is both highly specific and extremely diverse, properties that enable the body to recognize and respond to any possible microbe or foreign molecule (antigen). The immune system also has an exquisite ability to discriminate between the self and the nonself and only rarely turns on the body's own normal tissues. Finally, the immune system has a memory: after the initial exposure to an antigen, mysterious changes in the immune system make it capable of responding much more quickly and forcefully to reexposures.

Niels K. Jerne, who received the Nobel prize for medicine in 1984, laid the groundwork for the current understanding of the immune system in 1955 while working at the California Institute of Technology. He proposed that immune reactivity is based on the Darwinian principle of natural selection. It was already known that antibody molecules in the blood reacted specifically with antigens; Jerne suggested that everyone carries an initially small number of antibodies against every possible antigen. When an antigen enters the body, the antibodies capable of binding to it are positively selected (to use Darwin's terminology) and increase in number.

Four years later another Nobel laureate, Sir Macfarlane Burnet of the Walter and Eliza Hall Institute of Medical Research in Melbourne, gave Jerne's natural-selection theory a cellular basis in a small book entitled *The Clonal Selection Theory of Acquired Immunity*. In it, Burnet proposed that each antibody molecule is the product of a single cell. According to Burnet, the antigen reacts directly with the antibody-producing cell to stimulate antibody production. (This idea was first suggested in 1905 by Paul Ehrlich, the pioneering immunologist and Nobelist.) Implicit in Burnet's hypothesis was the idea that after selective activation by an antigen, the antibody-producing cell proliferated to form a clone, a set of cells with one common ancestor.

At the time of Burnet's proposal, the cells responsible for reacting with antigens were unknown, although it was

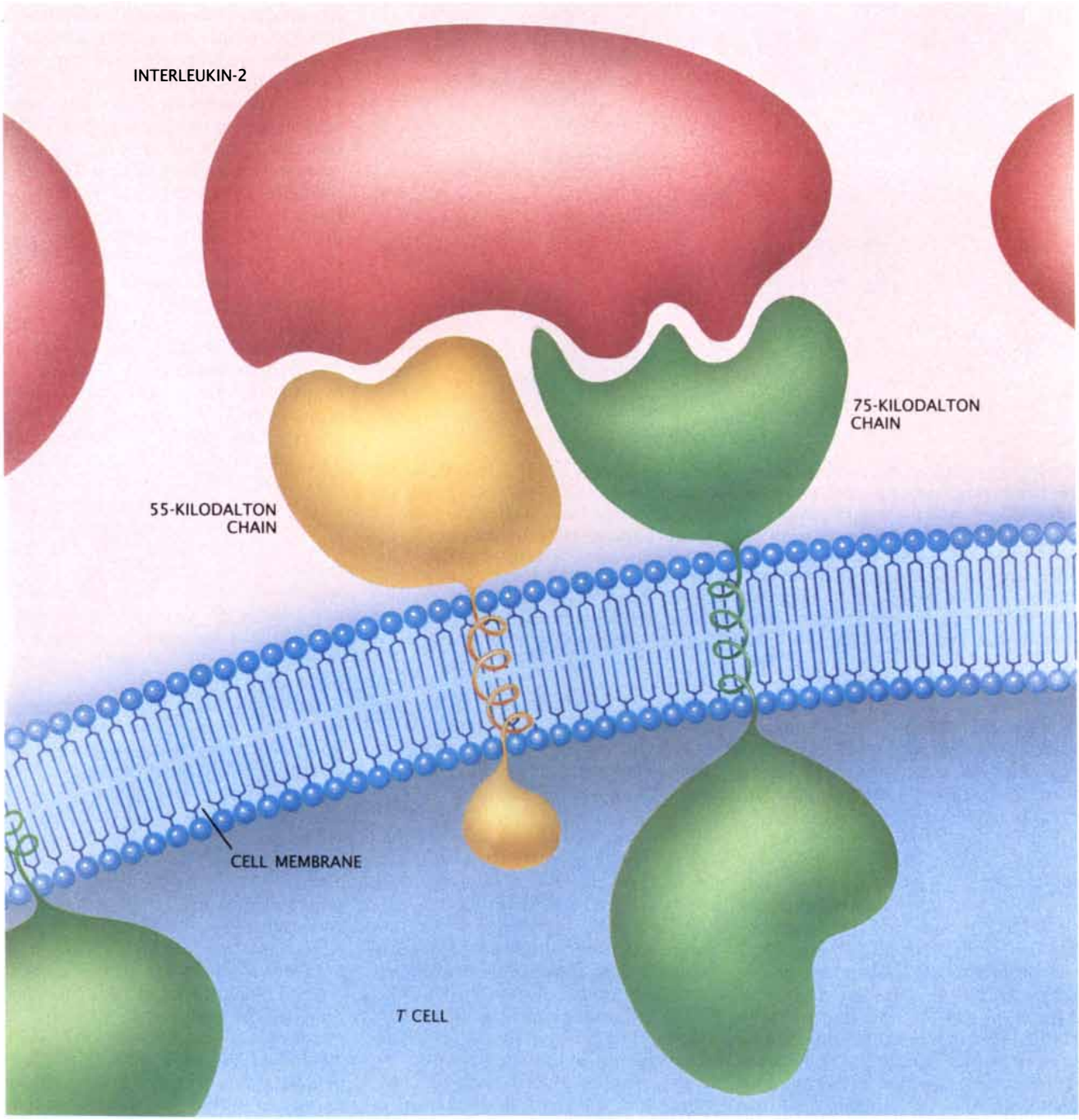
KENDALL A. SMITH is professor of medicine at Dartmouth Medical School. He received his medical degree summa cum laude from the Ohio State University College of Medicine in 1968. After completing his clinical medical training at Yale University, he devoted several years to postdoctoral research at the National Cancer Institute, Dartmouth and the Institute for Cancer Research and Immunogenetics in Villejuif, France. At Dartmouth, Smith studies the structural biology of interleukin-2 and its receptor, the molecular events that signal cell proliferation, the biochemical mechanisms responsible for immunologic memory and the potential therapeutic utility of the interleukins.

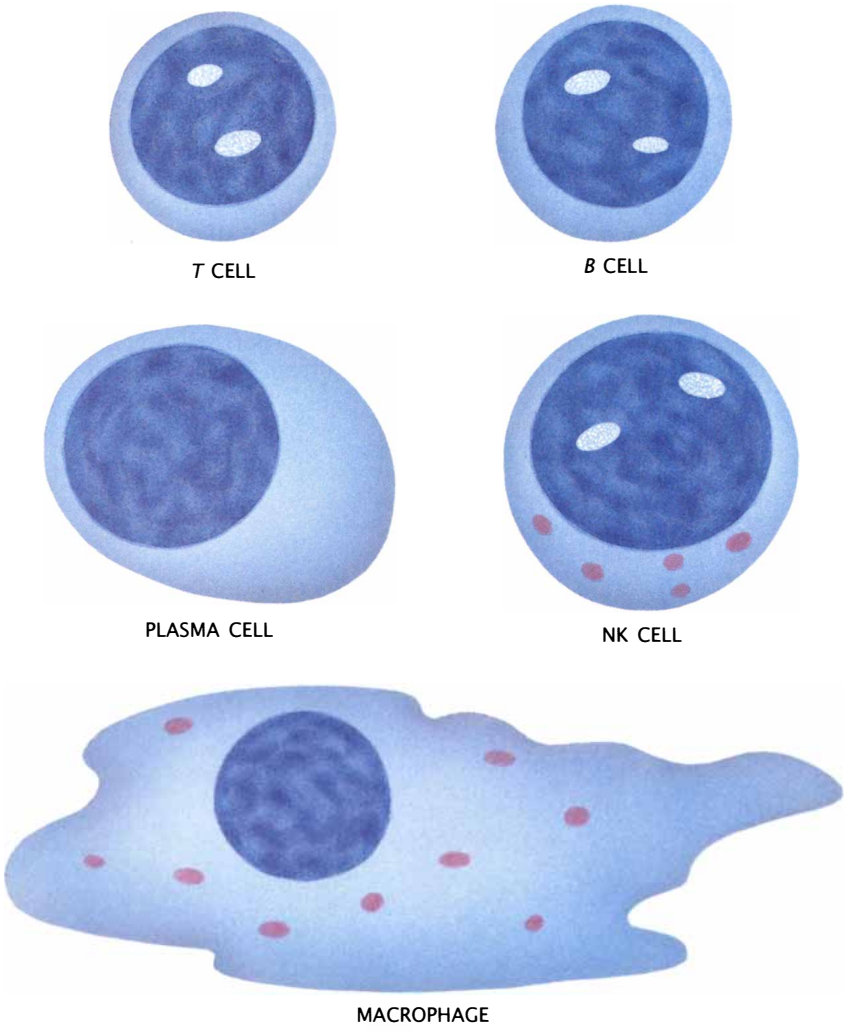
INTERLEUKIN-2 (IL-2) binds with a receptor on the outer membrane of antigen-stimulated *T* cells in the immune system. The receptor consists of a large (75-kilodalton) protein chain and a smaller (55-kilodalton) chain, each of which binds separately to the IL-2 molecule. The interaction between the IL-2 and the large chain signals the *T* cells to proliferate, thereby enhancing the immune system's specific assault on any microorganisms.

fairly certain that the white blood cells called plasma cells produced large amounts of antibodies. Lymphocytes, the predominant cells in the lymph nodes, were the most obvious candidates to react initially with antigens and thereby serve as the precursors for plasma cells, but lymphocytes were generally believed to be incapable of proliferating. Yet only one year later, in 1960, Peter C. Nowell of the University of Pennsylvania discovered that lymphocytes could proliferate, given the proper chemical stimulus.

In the decade after Burnet's hypothesis and Nowell's discovery, a sound cellular understanding of the immune system was established. Lymphocytes were found to be of two major types, *B* cells and *T* cells. *B* cells, which are derived from bone marrow, express antibody molecules on their surface, as Burnet had envisioned. Experiments by Gustav Nossal of the Hall Institute and by Jerne and Albert A. Nordin of the University of Pittsburgh demonstrated that when stimulated by a specific antigen, each *B* cell becomes a plasma cell that secretes antibodies with a single specificity.

T cells, which mature in the thymus, do not produce antibodies, but they do bear specific antigen receptors on their surface that strikingly resemble antibody molecules and that selectively bind antigen. Like *B* cells, *T* cells react to antigen stimulation by secreting molecules that mediate their immune function. On the basis of the molecules they secrete, *T* cells have been subcategorized as helper *T* cells and cytotoxic *T* cells. Today it is clear





PRINCIPAL CELLS of the immune system perform a variety of tasks. *T* cells, which are of two types, react to specific antigens (foreign molecules). Cytotoxic *T* cells kill infected cells and invading microorganisms by secreting toxic molecules. Helper *T* cells secrete interleukins, or growth factors, that enhance the immune response. *B* cells resemble *T* cells but do not attack microorganisms directly. Instead, when stimulated by an antigen, they differentiate and become plasma cells that secrete antibodies to combat the invaders. Natural killer, or NK, cells are not antigen-specific; they attack any foreign microbe. Large macrophages ingest microorganisms and present the antigens to *T* cells to trigger the specific immune response.

that helper *T* cells fulfull their role by secreting interleukins. Cytotoxic *T* cells, in contrast, make direct contact with infected cells and, by secreting toxic molecules, kill the cells and the microbes they contain.

T cells, in particular, turned out to serve as a good laboratory model of the immune response. When an antigen is injected into the body, the only detectable immune reaction is against that antigen specifically. In culture, *T* cell proliferation is also antigen-specific: only those cells that react with a given antigen survive and multiply. For this reason, the short-term culture of *T* cells was adopted in 1965 as a test-tube reproduction of the immune response. Also, the behavior of the cultured *T* cells helped to explain the phenomenon of immune memory: exposure to an antigen selectively increases the number of cells capable of responding to it in the future.

Immunologists had always assumed that antigen was the sole stimulus for *B* and *T* cell division. This assumption, which became dogma, had to be overturned before immunology could proceed to a more detailed and accurate description of how the immune system works.

The beginnings of the current understanding of the mechanisms responsible for stimulating lymphocyte growth can be traced back to 1965, when two teams of investigators at the Royal Victoria Hospital in Montreal, Shinpei Kasakura and Louis Lowenstein and, independently, J. Gordon and Lloyd D. MacLean, published papers simultaneously in *Nature*. Both teams reported that culture media "conditioned" by proliferating lymphocytes contained an unidentified substance that enhanced lymphocyte growth when antigen was present.

Many reports about these mysterious growth-enhancing substances appeared during the next decade. Nevertheless, most immunologists ignored these reports and continued to maintain that antigens were the sole agents responsible for lymphocyte proliferation. At most, the factor in the lymphocyte-conditioned medium was thought simply to amplify proliferation that had already been triggered by an antigen. With this idea in mind, various investigators had published several papers by the early 1970's that described ways of growing lymphocytes in culture for prolonged periods through the repetitive addition of antigen. It was possible to stimulate the proliferation of cultured lymphocytes for as long as four months by these techniques, and the lymphocytes retained their antigen specificity.

Immunologists doubted that a factor released by cells could specifically stimulate lymphocyte division because such a factor, it seemed, should amplify the proliferation of all lymphocytes regardless of whether they had encountered their specific antigen. In 1976, however, Doris A. Morgan, working with Francis W. Ruscetti in Robert C. Gallo's laboratory at the National Cancer Institute (NCI) in Bethesda, Md., reported that normal human *T* cells could be cultured without antigen for up to nine months, provided that a lymphocyte-conditioned medium was added to the cultures at regular intervals.

Actually, Morgan's observation was serendipitous. A self-proclaimed novice in lymphocyte culture methods, Morgan was trained in hematology and in hematopoiesis (the process by which blood cells develop). While trying to establish long-term cultures of leukemia cells, Morgan employed a lymphocyte-conditioned medium as a stimulant because lymphocytes were known to release factors that promoted the growth of the early blood-forming cells. Much to her initial anguish, the cells from leukemia patients that grew in the lymphocyte-conditioned medium appeared to be normal *T* cells instead of leukemia cells. For immunologists, however, her findings were important in that they suggested that

some kind of factor in the conditioned medium other than antigen was responsible for *T* cell growth.

Although Morgan's report appeared in the prominent journal *Science,* its significance was lost on most of the immunology community. Morgan and her co-workers were well known in hematology and virology circles but were unfamiliar to most immunologists. The title of the paper stressed that the cultivated *T* cells had come from bone marrow, which meant that they might be immature or otherwise unrepresentative of most *T* cells. More important, the cultured cells were not shown to perform any antigen-specific functions, and immunologists were traditionally indifferent to phenomena that lacked antigen specificity.

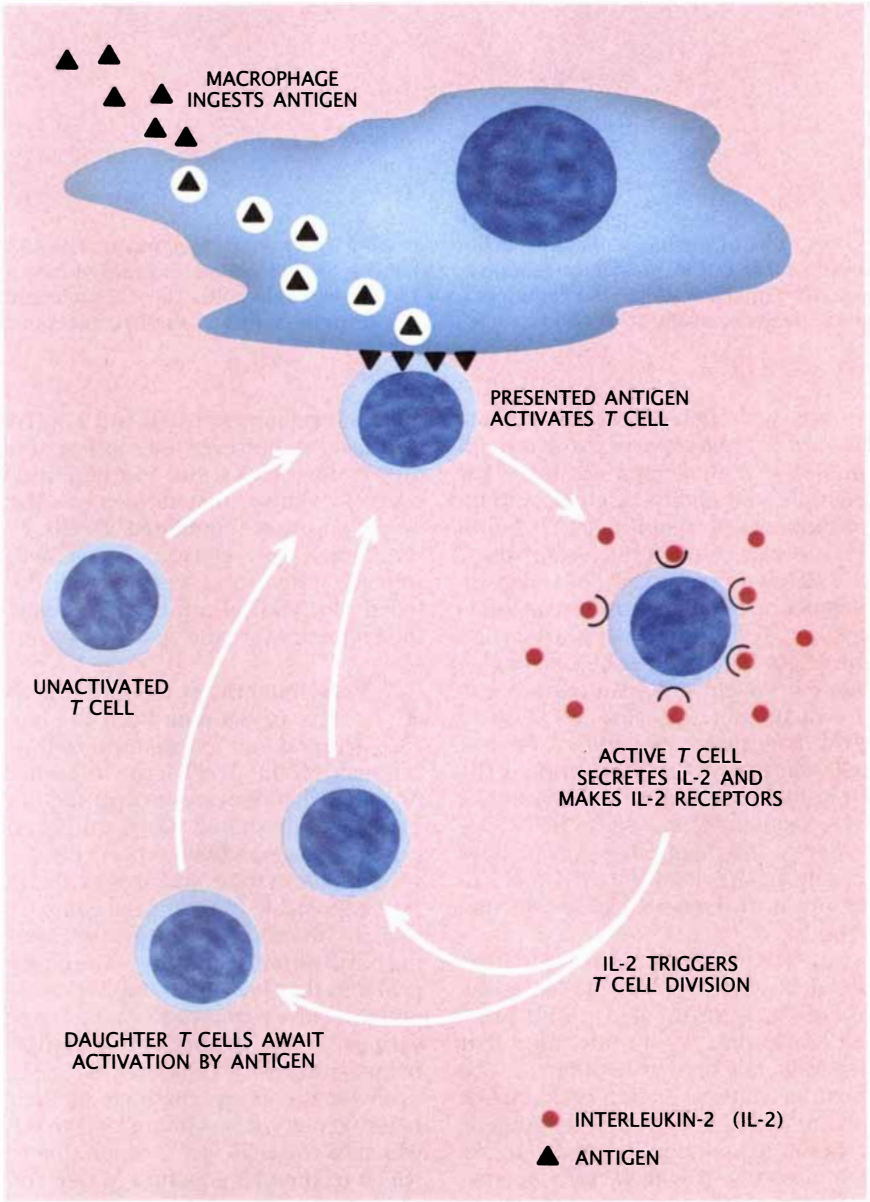
Like Morgan, I was trained in hematology and hematopoiesis rather than in immunology. As a postdoctoral fellow, I had worked with George Mathe of the Institute of Cancer and Genetic Research in Villejuif, France, who was one of the first investigators to try to treat leukemia with immunotherapy. I had become intrigued by the possibility of stimulating the growth of cytotoxic lymphocytes to kill leukemia cells. In 1974, therefore, as assistant professor of medicine at Dartmouth Medical School, I initiated a research program directed toward understanding the fundamental determinants of cellular proliferation. By 1976 my colleagues and I had demonstrated that cytotoxic *T* cells could kill mouse leukemia cells in culture, and yet we were frustrated by our inability to maintain *T* cells in culture for more than a few days. We therefore decided to try new methods to achieve long-term, antigen-specific *T* cell growth.

Unencumbered by immunology's dogmas, we combined the methods that others had already tested. First we vaccinated mice repeatedly with irradiated tumor cells to expand the population of tumor-reactive *T* cells in the animals. *T* cells from the mice were then mixed with tumor cells in tissue culture; in this short-term culture, only those *T* cells reacting with tumor antigen would survive and proliferate. After one to two weeks the surviving *T* cells were transferred to a lymphocyte-conditioned medium similar to that described by Morgan. Steven Gillis, who was a first-year graduate student at the time, was the first in the laboratory to succeed at creating long-term cultures of *T* cells.

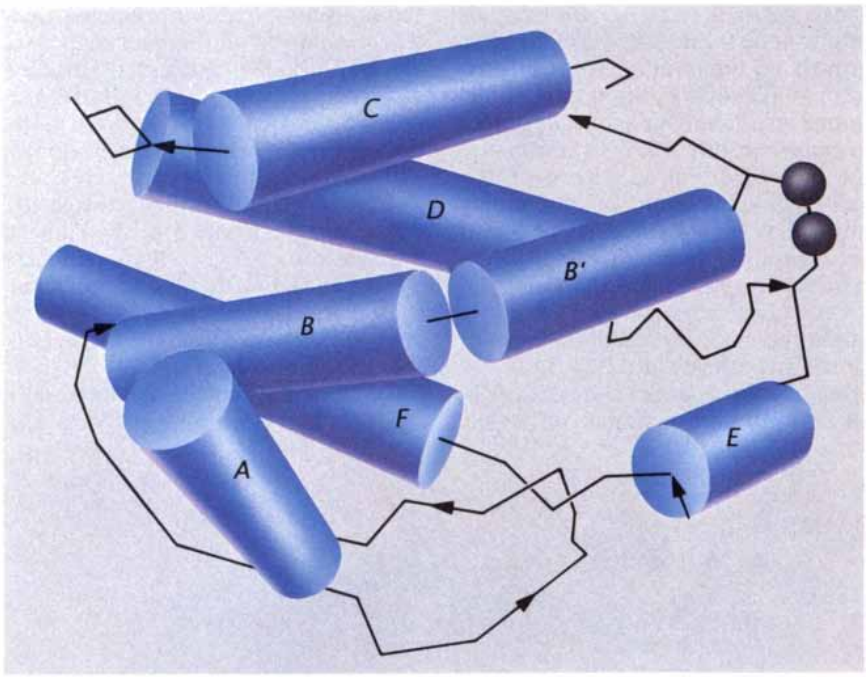
In short, despite the prevailing opinion that this approach would fail because we had removed the specific tumor antigen, we saw the long-term growth of tumor-antigen-specific cytotoxic *T* lymphocytes in the conditioned medium. Our paper appeared in *Nature* in July, 1977. In contrast to the lack of attention accorded Morgan's paper, the immunologic community was very interested in our report because it emphasized the culture of functional, antigen-specific *T* cells.

Our success suggested that we might be able to validate Burnet's clonal hypothesis directly by developing clones of antigen-specific cells, all derived from a single cell. Again, the prevailing opinion predicted that this would be an almost impossible task because single cells usually grow poorly (if at all) in culture. Yet the addition of the lymphocyte-conditioned medium led to *T* cell clones with surprising efficiency. Each clone showed antigen-specific cytotoxicity that could be ascribed to the clone's descent from a single cell. We reported the derivation of the first monoclonal cytotoxic *T* cell lines in 1979, 20 years after Burnet had formulated his clonal-selection theory.

Fundamental problems in immunology that had been impossible to ap-



PROLIFERATION of *T* cells is controlled by IL-2 after an antigen, ingested and presented by a macrophage, activates individual *T* cells. The antigen stimulates the *T* cells to secrete IL-2 and to make IL-2 receptors. Subsequently, the binding of IL-2 with its receptors signals the *T* cells to divide, thereby producing pairs of daughter cells that can also be activated by the antigen. In this way, a clone of identical antigen-specific *T* cells grows until the immune system eliminates the antigen from the body.



STRUCTURE of the IL-2 molecule has been revealed by X-ray crytallography. The 133 amino acids in the protein backbone of the molecule are coiled and folded into a roughly spherical shape; the cylinders (*A–F*) represent tight coils. The IL-2 molecule must be structurally intact to react with both protein chains in the IL-2 receptor.

proach with heterogeneous populations of lymphocytes were solved with antigen-specific cloned cell lines. For example, the ability to grow unlimited numbers of identical *T* cells led to a molecular characterization of the *T* cell antigen receptor. It was also instrumental in demonstrating unequivocally that a protein structure called the major histocompatibility complex plays an essential role in the recognition of antigen by *T* cells. The ability to grow clones of helper and cytotoxic *T* cells was important for identifying the molecular mechanisms underlying the cells' activities. In brief, the ability to derive and manipulate *T* cell clones in culture provided detailed molecular proof of Burnet's clonal-selection hypothesis.

Our first experiments provided the initial insight into how the selection of a clone by an antigen-specific process can initiate cell proliferation that depends only on a growth factor. The most logical explanation is that when a lymphocyte is activated by antigen, it develops a unique capacity to respond to the growth factor; the vast majority of lymphocytes (which do not respond to the same antigen) remain inactive.

My colleagues and I conducted a series of experiments to test this theory, and we proudly submitted the results to one of the most prestigious immunology journals in 1978. To our chagrin, however, the notion of a growth factor was still too heretical: everyone "knew" that antigen was the only thing that stimulated *T* cells to proliferate. The reviewers, it seemed, were skeptical and insisted on detailed biochemical information about the factor that we did not then have.

Consequently, we devoted our efforts to defining both the biological and biochemical characteristics of the *T* cell growth factor. Although lymphocyte growth factors had been described since 1965, no one had developed a quantitative assay for their activities, and it was therefore impossible to detect and compare relative amounts of the factors during purification procedures. The major problem that had hindered efforts by previous investigators to characterize various factors had been the heterogeneous cultures of cells that had been used for the assays. Because of their heterogeneity, it was impossible to tell which of the cells were responding to which of the many factors in the lymphocyte-conditioned medium.

Because we had already established monoclonal lines of growth-factor–dependent cells, however, the problem of cell heterogeneity in the assay was solved. Also, as a postdoctoral fellow, I had developed a quantitative assay for erythropoietin, the red-blood-cell growth factor; it was relatively simple to adapt the method to measuring *T* cell growth factor.

Armed with a rapid and quantitative assay, we performed a series of experiments published in several papers between 1978 and 1983 that collectively described for the first time the biological and biochemical characteristics of the *T* cell growth factor now known as interleukin-2. These studies showed that the immune system, after it senses an antigen, transfers the control over the immune response from an antigen-regulated mechanism to a hormonelike regulatory system.

The general mechanism works as follows: When an antigen is introduced into the body, it is ingested by macrophages (a class of scavenger cells) and *B* cells. These cells digest the antigen and present short segments of the antigen molecules on their cell surface. Most *T* cells in the body do not recognize the presented antigens, and they continue to move quietly through the bloodstream and lymphatic system. A few *T* cells, however, have antigen receptors that bind to the presented antigens, stimulating the cells. Thereafter, these antigen-stimulated *T* cells become autonomous growth-factor factories, both secreting IL-2 and responding to it by proliferating. The end result of this process is an expansion of only those *T* cell clones that react to the antigen stimulus.

Although we were able to demonstrate that only antigen-activated *T* cells responded to IL-2, we did not know how IL-2 interacted with the cells to produce its effects. I suspected that the mechanism would probably involve cell-surface receptors for IL-2, just as receptors for insulin mediate the activity of that hormone. Our experiments soon showed that activated *T* cells could absorb IL-2 activity, as would be expected if they bore IL-2 receptors. These findings encouraged us to produce and purify radioactively labeled IL-2, with which we could observe the process directly. The results from our very first experiments were unambiguous: IL-2 binds to cells because of a high-affinity receptor on the cell surface.

The application of this hormone-and-receptor concept to the immune system has had an extraordinary impact on models of immune regulation. Cellular immunology had previously maintained that macrophages, *B* cells and *T* cells signaled one another exclusively through intimate contact. After the discovery of IL-2 and other soluble factors and the advent of the idea

that the expression of IL-2 receptors determines which cells participate in the immune response, the interactions of the immune-system cells lost their mystical aura. It was possible to understand them according to principles borrowed from endocrinology—principles that describe the interactions of hormones and their receptors.

It was evident by 1982 that understanding the molecular mechanisms signaling *T* cells to proliferate was going to require a knowledge of the structures of IL-2 and its receptor. Structural knowledge would also be necessary to design agents that could therapeutically block or mimic the hormone-receptor interaction. An important first step toward these goals was the isolation of the gene for IL-2, which was first accomplished in 1983 by Tadatsugu Taniguchi and his colleagues at Tokyo University.

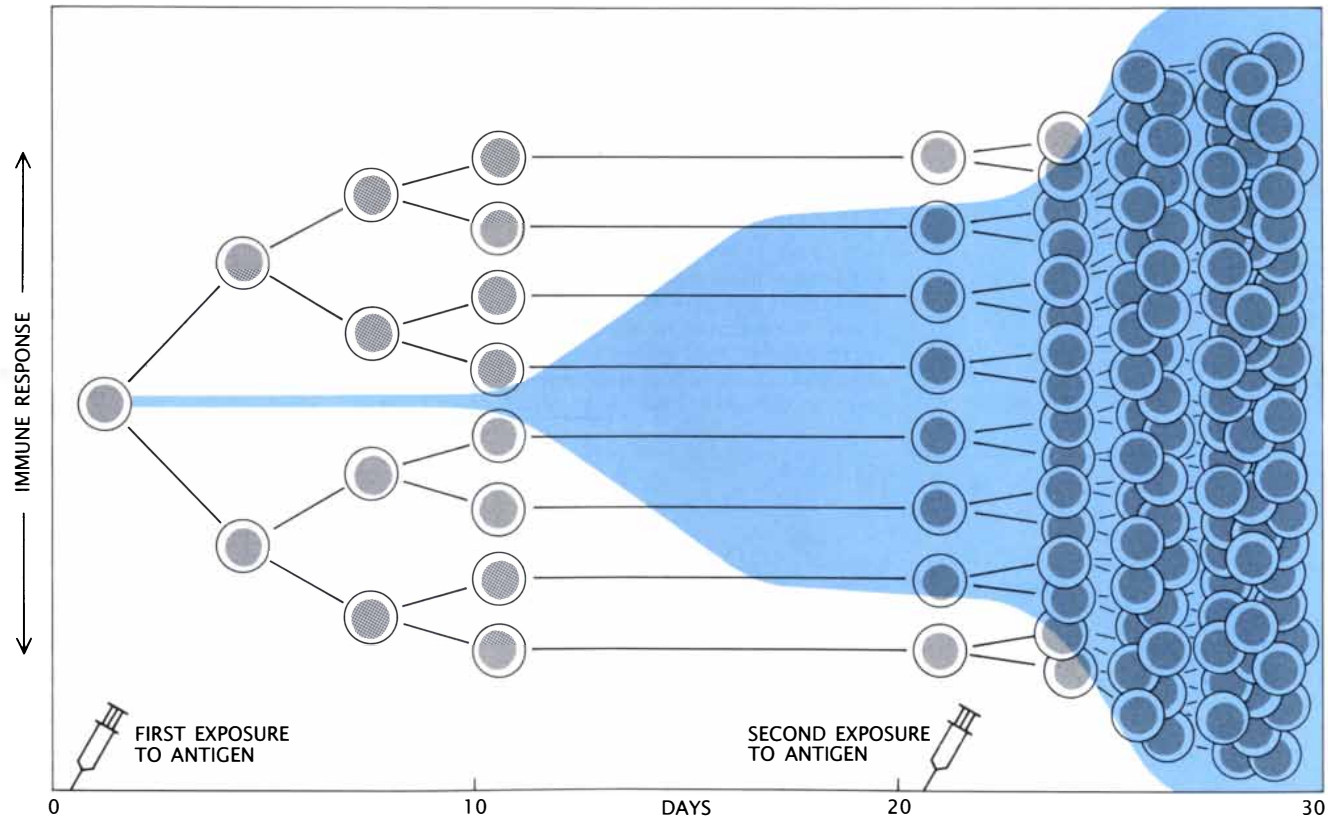
Once a gene has been isolated, gene-cloning technology makes it feasible to produce the gene's protein product in virtually unlimited quantities. In the past six years, IL-2 has been made available by biotechnology firms and distributed throughout the world. The availability of large quantities of pure IL-2 permitted David B. McKay and his group at the University of Colorado at Boulder to grow crystals of IL-2 and to deduce the three-dimensional structure of the molecule by X-ray crystallography in 1987.

In 1984 Warren J. Leonard and others, working in Thomas A. Waldmann's laboratory at NCI, and Toshio Nikaido, working with Tasuku Honjo, Takashi Uchiyama and others at Kyoto University, simultaneously reported the isolation of a gene for a putative IL-2-receptor protein. This protein chain reacted with a monoclonal antibody against the IL-2 receptor developed by Uchiyama; however, the chain's small size and low affinity for IL-2 meant that it could not be the complete IL-2 receptor. Then, in 1986, Keisuke Teshigawara of Kyoto University, working in our laboratory at Dartmouth, and Mitsuru Tsudo, also of Kyoto University but working in Waldmann's laboratory, independently made the startling discovery of a second, larger IL-2-receptor chain.

It turned out that the IL-2 receptor is made up of two chains: one that is 55 kilodaltons in size and reacts with Uchiyama's antibody and an additional chain that is 75 kilodaltons in size. Michael Sharon of the National Institute of Child Health and Human Development, working with Leonard, also had data consistent with this interpretation. By the fall of 1986 most workers in the field began to race toward the isolation and characterization of the new 75-kilodalton chain.

Tsudo eventually won this race after taking a new position in Tokyo at the Metropolitan Institute for Medical Sciences. By mid-1988 he had derived monoclonal antibodies that reacted with the 75-kilodalton chain. Subsequently, he collaborated with Masanori Hatakeyama, a young hematologist working in Taniguchi's laboratory at Osaka University. Using Tsudo's antibodies, Hatakeyama was able to identify and isolate the gene encoding the 75-kilodalton chain.

While these studies were proceeding, we devoted our efforts at Dartmouth to defining the role of the IL-2-receptor system in the immune response. In a series of experiments,



CELLULAR BASIS of the immune response explains the phenomenon of immune memory, in which the body responds to an antigen faster and more powerfully after reexposures than it did after the first exposure. In this diagram, the strength of the immune response is proportional to the pink band's thickness. After the first exposure to an antigen, the immune response does not become detectable for 10 to 14 days because it takes time for the clone of antigen-specific *T* cells to grow to an effective size. After the second exposure, however, the lag time is shorter because the number of *T* cells in the blood that can react with the antigen is already large. Exponential growth by the clone heightens the immune response.

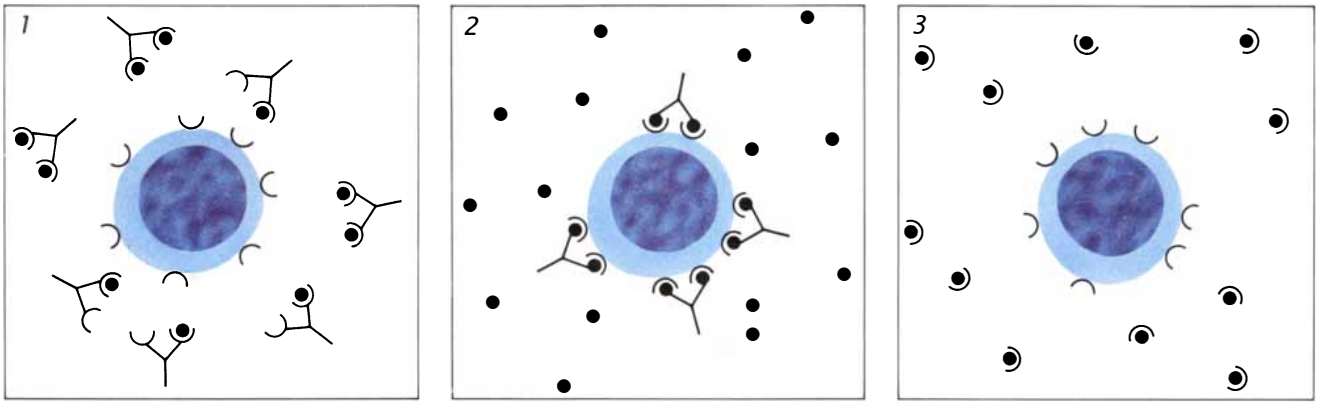
Doreen A. Cantrell determined how the IL-2-receptor mechanism functions during a model *T* cell immune response. She discovered that there are only three parameters important for regulating *T* cell proliferation after antigen activation: the concentration of IL-2, the density of IL-2 receptors on the cell surface and the duration of the IL-2-receptor interaction. Apparently, before a cell will irrevocably commit to dividing, it must experience a number of IL-2-receptor interactions over several hours.

Huey-Mei Wang, a recent graduate student in my laboratory, extended these findings by showing that the IL-2 receptor functions as an "on-off" switch. IL-2 binds rapidly to the receptor by interacting with the binding site on the 55-kilodalton chain. It is then tightly held by an interaction with the binding site on the 75-kilodalton chain. (It is because IL-2 must interact with both receptor chains that the molecule must be intact to function.) This interaction between the IL-2 molecule and the 75-kilodalton chain turns on the intracellular mechanisms that signal the *T* cell to become active. When the molecule dissociates from the receptor, which happens only very slowly because of the strength of the interaction with the large chain, the signaling stops.

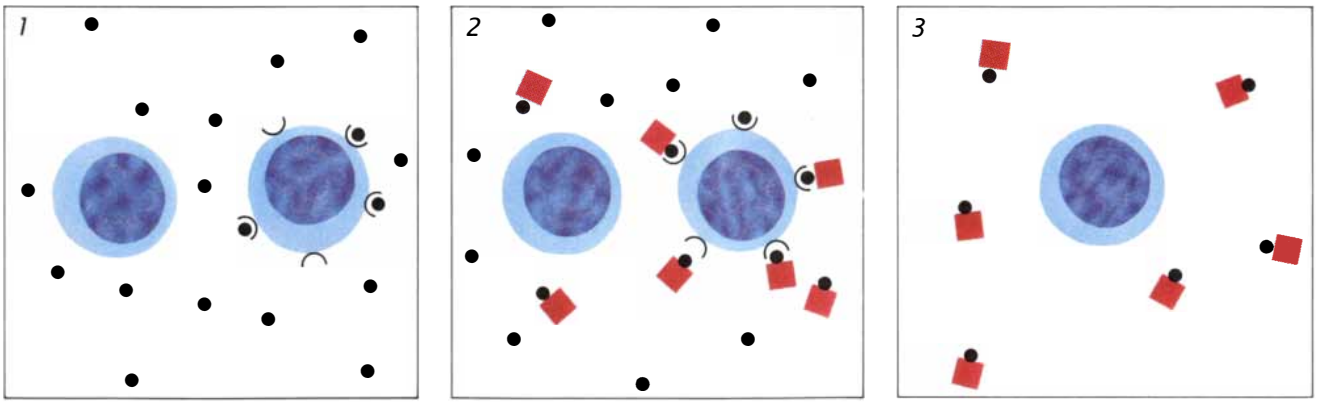
As the immune response proceeds, antigen gradually disappears from the body, and the antigen-reactive *T* cells stop receiving signals from their antigen receptors. Consequently, the number of IL-2 receptors on the cells gradually declines and the expanded clone of cells stops proliferating. These lingering *T* cells make up the memory population of the immune system.

T cells are not the only cells stimulated by IL-2. As early as 1981 Christopher S. Henney and his colleagues at the University of Washington reported that cells with natural killer (NK) activity are stimulated by IL-2. NK cells make up about 10 percent of the total circulating lymphocyte population and are thought to participate in the immune surveillance against cancer cells and in the body's initial responses to viruses; they may serve as a first line of defense by virtue of their immediate responsiveness to IL-2. Unlike *T* cells, NK cells have no antigen receptors and appear always to be in a state of activation. Studies have revealed that NK cells express the 75-kilodalton chain of the IL-2 receptor continuously.

B cells, like *T* cells, remain inactive until they make contact with antigen. They then undergo clonal expansion and become plasma cells, releasing



SELECTIVE SUPPRESSION of the immune response would be desirable for treating organ-transplant patients and people with autoimmune diseases. One approach might be to interfere with the binding between IL-2 molecules and their receptors on *T* cells, for example, by injecting antibodies against IL-2 into the body (*1*). Antibodies against the IL-2 receptor should have the same effect, because the IL-2 molecules would be unable to find unoccupied receptors (*2*). Another possibility would involve injected soluble IL-2 receptors, which could compete with the cellular receptors (*3*).



DELETING a specific clone of *T* cells is another route to suppressing the immune system selectively. One way to do so would be to administer IL-2 conjugated to a bacterial toxin. Only *T* cells that have been activated by antigen make IL-2 and IL-2 receptors (*1*). Hence, the IL-2-toxin conjugate would bind with and kill only the active *T* cells (*2*). The inactive *T* cells would not be harmed by the conjugate molecules (*3*).

large amounts of antibody. IL-2's role in these cellular changes is still somewhat controversial, but most investigators agree that it promotes the proliferation of antigen-activated *B* cell clones in the same way that it stimulates *T* cell growth. In addition, recent experiments by Marian E. Koshland of the University of California at Berkeley and Kenji Nakanishi from the Hyogo College of Medicine have demonstrated that IL-2 participates in the differentiation of *B* cells by helping them to start secreting antibodies.

All of these findings indicated that the IL-2-receptor mechanism plays a crucial role in regulating the first events of an immune response. It was therefore easy to imagine ways of exploiting this hormone-receptor interaction either to enhance or to suppress the immune response for therapeutic purposes. Actually, the two most effective types of immunosuppressive drugs prescribed today, cyclosporine and the glucocorticoids, turned out to work by inhibiting the production of IL-2.

My colleague Thomas L. Ciardelli of Dartmouth hopes to capitalize on the IL-2-receptor system by creating modified forms of the IL-2 molecule that could serve either as IL-2 antagonists (by tying up the receptor without activating it) or as "superleukins" with greater immunostimulatory effects. A similarly imaginative approach to immunosuppression has been developed independently by John R. Murphy of Boston University and Ira H. Pastan of NCI. With genetic-engineering techniques, these investigators have coupled bacterial toxins to IL-2. These toxic hybrid molecules bind to and kill antigen-activated *T* cells and *B* cells, thereby deleting an antigen-specific clone from the body.

Monoclonal antibodies that react with IL-2 or its receptor can also suppress antigen-activated *T* cell responses. Antibodies that block the binding of IL-2 to its receptor completely inhibit the antigen-specific proliferation of *T* cells in culture; the antibodies have also proved to be effective antigen-specific immunosuppressants in heart-transplantation experiments. The same function might be served by a different preparation, consisting of the IL-2-binding portion of the receptor molecules, which could compete with cell-surface receptors for IL-2.

IL-2 itself has several obvious applications as an immune stimulant. Because IL-2 stimulates the clonal expansion of *T* cells and *B* cells after the introduction of antigen, one obvious use for IL-2 will be as an immunopotentiator: a substance that boosts the effectiveness of vaccinations. The safety and efficacy of IL-2 has already been tested in human beings by Stefan C. Meuer of the University of Heidelberg during immunization trials with a hepatitis B vaccine. As additional vaccines are developed by genetic engineering, the availability of an effective immunostimulator such as IL-2 may well determine their eventual success or failure.

IL-2 has already been applied as an experimental immunotherapy for cancer. Pioneered by Steven A. Rosenberg and his colleagues at NCI, this strategy culminates an almost century-long search for an effective way of harnessing the immune system to destroy cancer cells. So far, patients with at least three types of cancers that are usually extremely resistant to conventional therapies—malignant melanoma, kidney cancer and colon cancer—have responded to the administration of IL-2 in combination with IL-2-stimulated NK cells. The proportion of patients that have had a significant response to therapy is still low, about 20 percent. Yet many of the patients who have responded have had remarkably stable, disease-free remissions without any further therapy.

One of the most promising areas for immunoenhancement therapy lies in the treatment of infectious diseases. Since the introduction of antibiotics in the late 1940's and 1950's, most of the developed world has been relatively free of the most common bacterial diseases. Equally effective antimicrobial agents against fungi, parasites and viruses have not yet been developed, however. Moreover, infectious diseases continue to be major health problems in developing countries.

Many chronic infectious diseases that do not respond to conventional treatments, such as tuberculosis, leprosy and leishmaniasis, involve microbial infestations of macrophages. These infected macrophages steadily release inflammatory substances that can eventually destroy the surrounding tissues. From results obtained by Gilla Kaplan, Zanvil A. Cohn and their colleagues at the Rockefeller University, it now seems that the activation and expansion of *T* cells and NK cells by IL-2 can tip the scales in the favor of the immune system: IL-2-stimulated cells can kill the infected cells and the microbes they contain. The Third World could be the first place to benefit from such immunotherapy.

AIDS shares many features of these other chronic infectious diseases. The human immunodeficiency virus (HIV), however, infects helper *T* cells as well as macrophages, with the result that patients eventually lose the very cells that are critical for mounting an effective immune response. Consequently, people with AIDS become susceptible to any and all the common microbes in the environment. Investigators are now beginning clinical trials of IL-2 for the treatment of HIV infection, administering it before the overt immunodeficiency develops. If HIV-infected cells can be killed by IL-2-activated *T* cells and NK cells before the virus becomes widespread, patients may be spared the devastating loss of their immune cells.

The converse goal of immunosuppression is important primarily in developed countries, where organ transplants and autoimmune diseases are more common. The currently available drugs are broadly immunosuppressive and must be taken over long periods; consequently, organ-transplant recipients and patients with autoimmune diseases run serious risks of potentially fatal infections during treatment. An approach is needed that, for example, suppresses only those immune cells reactive to a grafted organ or tissue while sparing the rest of the immune cells. Any approach that capitalizes on the blocking of the IL-2-receptor interaction will provide that much-needed specificity.

In the 30 years since Burnet first enunciated the clonal-selection theory, a detailed cellular and molecular understanding has been crafted that explains how the immune system is regulated. The appreciation of the fundamental role of IL-2 in mediating the expansion of a clone of immune cells after it has been selected by an antigen has been instrumental to this understanding. New means of stimulating or suppressing the immune system for the treatment of diseases can now be created based on a rational understanding of immunity.

FURTHER READING

CELLULAR IMMUNOLOGY. Sir Macfarlane Burnet. Melbourne University Press, 1969.

INTERLEUKIN 2. Edited by Kendall A. Smith. Academic Press, 1988.

INTERLEUKIN-2: INCEPTION, IMPACT, AND IMPLICATIONS. Kendall A. Smith in *Science,* Vol. 240, No. 4856, pages 1169–1176; May 27, 1988.

THE INTERLEUKIN 2 RECEPTOR. Kendall A. Smith in *Annual Review of Cell Biology.* Edited by G. E. Palade. Vol. 5, pages 397–425; 1989.

New Radioactivities

An atomic nucleus can spontaneously restructure itself, occasionally ejecting rare clusters of protons and neutrons. Observations of these new radioactivities have illuminated theories of nuclear dynamics

by Walter Greiner and Aurel Sandulescu

Radioactivities emerge as ambassadors from a distant world, the nucleus of an atom. The nucleus sends forth a flash of energy, a couple of protons, a few neutrons or other particles. Each of these ambassadors carries a particular message. Most often they announce the decay of a nucleus from an energetic, unstable state to a less energetic, more stable one. Radioactivities also reveal subtle clues about nuclear structure. By piecing these clues together, investigators have created detailed models of the nucleus. The models not only account for most nuclear phenomena but also predict many new types of radioactivities. The triumph of the past decade has been the discovery of several of these new nuclear fragments.

The discoveries settled a 40-year-old quandary in nuclear physics. Until the 1980's it appeared as though the nuclear fragments from radioactive processes came in roughly three sizes: four, 100 or 200 nucleons—a term that refers to both protons and neutrons. In the four-nucleon range is the alpha particle, or helium nucleus. If an alpha particle emerges from an atom, it leaves behind a nucleus composed of approximately 200 nucleons. In the 100-nucleon range are the fragments from fission, a process in which a heavy nucleus splits roughly in half. The restricted range of sizes raised an intriguing question: Why did a nucleus not emit a fragment composed of other quantities of nucleons—why not 14 or 24?

Today it is known that a nucleus can indeed eject a fragment of this size or any other. These new radioactivities form when a large number of nucleons within the nucleus spontaneously rearrange themselves in certain configurations. Because these large-scale rearrangements occur at random, the emission of a new radioactivity is in general a much rarer event than, say, the emission of an alpha particle. By the end of the 1980's, physicists had succeeded in observing many of these new nuclear ambassadors.

Their achievement marked a new era in nuclear physics and was the culmination of a series of experiments that began nearly a century ago when A. Henri Becquerel first discovered radioactivity. Soon after, Ernest Rutherford distinguished two forms, the alpha ray and the beta ray. Later the French physicist Paul U. Villard identified a third, which Rutherford named the gamma ray.

For years these investigators and others debated the composition and origin of these rays; today few doubts remain. A gamma ray is electromagnetic radiation whose wavelength is about one million times shorter than that of light. A gamma ray carries away some of the energy from a nuclear reaction just as light conveys energy from a chemical reaction. A beta ray consists either of an electron and an elusive particle called an antineutrino or of their antimatter counterparts. A beta ray released from the nucleus signals the transformation of a proton to a neutron. An alpha ray consists of two protons and two neutrons. This combination, which became known as an alpha particle, has the same composition as the nucleus of helium 4. (The isotope number, here "4," represents the number of protons and neutrons in the nucleus.)

From the standpoint of nuclear structure, the emission of an alpha particle is perhaps the most interesting of these types of radiation. Alpha emission represents the decay

WALTER GREINER and AUREL SANDULESCU share an avid interest in the theory of nuclear fission. Since 1965 Greiner has been professor of physics and has directed the Institute for Theoretical Physics at the Johann Wolfgang Goethe University in Frankfurt. Greiner was educated at the universities in Frankfurt, Darmstadt and Freiburg, where he received his doctoral degree. Sandulescu has conducted research at the Institute for Atomic Physics in Bucharest since 1956. He got his doctoral degree from the University of Bucharest. From 1983 to 1986 Sandulescu was vice director of the Joint Institute for Nuclear Research in Dubna, U.S.S.R.

URANIUM NUCLEUS, which consists of 92 protons and 140 neutrons, should decay into lead (82 protons and 126 neutrons) and neon (10 protons and 14 neutrons). A uranium nucleus, like all other nuclei,

of a parent nucleus into two daughter nuclei: one a helium 4 nucleus, the other a more massive nucleus. At the time that alpha decay was discovered, however, investigators were ill equipped to explain the phenomenon. They lacked a basic knowledge of the shape of the nucleus and a fundamental understanding of the interactions within it.

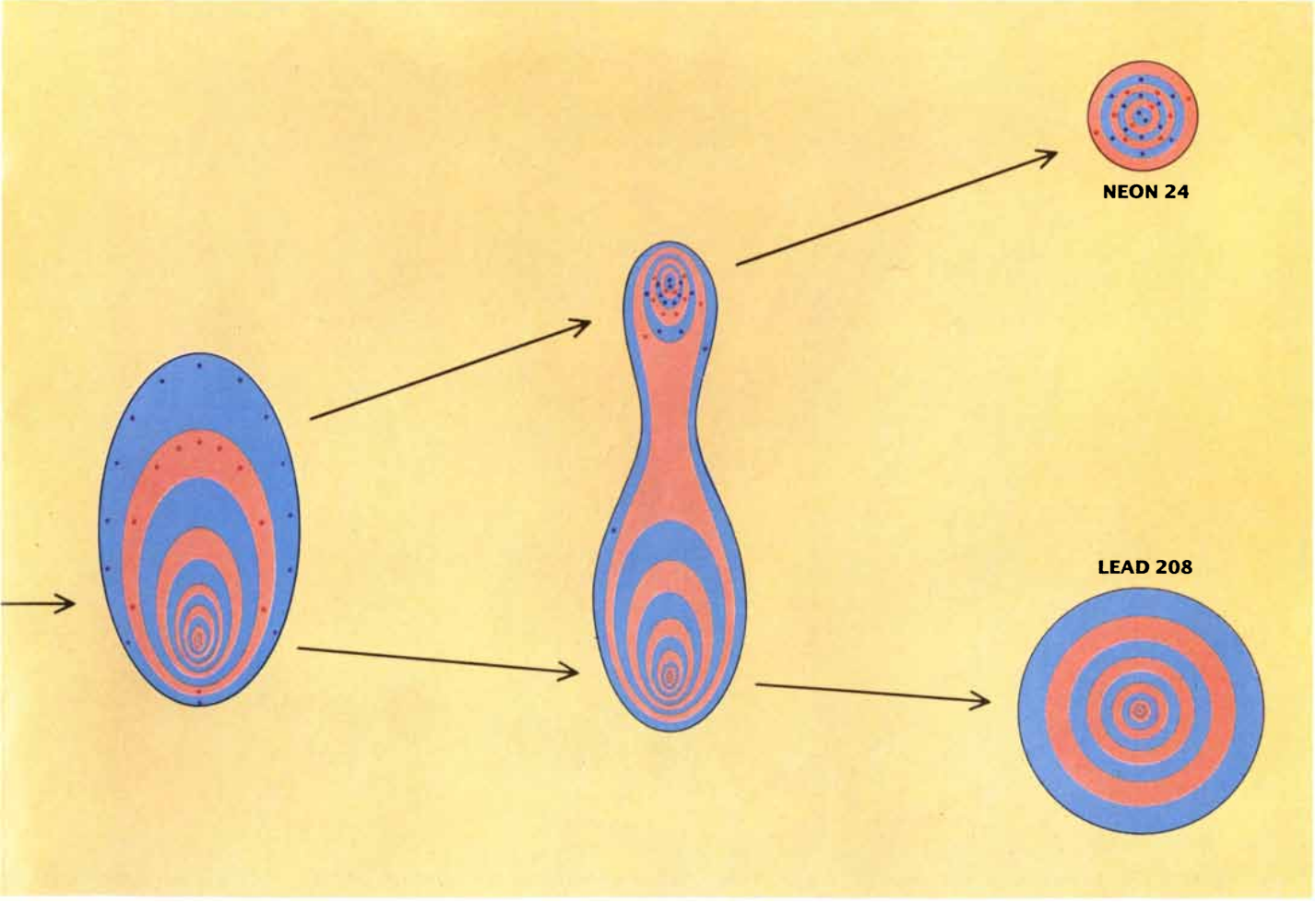
It is now known that the atomic nucleus is a more or less spherical object whose diameter is about a few fermis—a unit of measure equal to one quadrillionth of a meter, or simply 10^{-15} meter. Electrons orbit the nucleus at a distance of about 100,000 fermis. (For comparison, the radius of the moon's orbit is only about 30 times greater than the diameter of the earth.) Packed in this fermi-size nucleus is nearly all of the mass of an atom and all of its positive electric charge. The mass of the nucleus comes mainly from nucleons. Protons carry the positive charge.

The structure of the nucleus arises from two types of interactions: strong and electromagnetic. As a result of the strong interaction, or nuclear force, protons bind to neutrons and to each other. The nuclear force binds nucleons very tightly but acts over a very short range. To separate two neutrons that are one fermi apart, for instance, requires an energy of about one million electron volts. On the other hand, only about 10 electron volts is needed to dissociate two nucleons that are 10 fermis apart.

As a result of the electromagnetic interaction, or Coulomb force, protons repel other protons. Although the Coulomb force is weaker than the nuclear force, it acts over a much longer range. If two protons are one fermi apart, the Coulomb force is about 100 times weaker than the nuclear force. Yet at a distance of 10 fermis, the Coulomb force is about 10 times stronger than the nuclear force.

These relations have important implications for alpha decay. Before an alpha particle emerges from the nucleus, it encounters both the nuclear attraction from other nucleons and the Coulomb repulsion from other protons. An alpha particle can escape the nucleus if it gains enough energy to travel to the point where the Coulomb repulsion exceeds the nuclear attraction.

The energy for alpha decay can come from an external source, but for certain kinds of nuclei enough energy is bound up within the nucleus to allow an alpha particle to be emitted spontaneously. The criterion for spontaneous emission is that the mass of the parent nucleus be greater than the sum of the masses of the daughter nucleus and the alpha particle. Be-



consists of shells occupied by a certain number of protons (*light-red bands*) or neutrons (*light-blue bands*). The outermost neutron shell of the uranium nucleus can hold up to 58 neutrons but here has only 14 (*blue dots*). The outermost proton shell contains 10 protons (*red dots*) out of a possible 44. Because protons (or neutrons) do not fill the shells of the uranium nucleus, the nuclear structure is unstable. Hence, the uranium nucleus can spontaneously deform and break apart into two stable nuclei. Workers recently discovered this process, which is known as the superasymmetric fission of uranium.

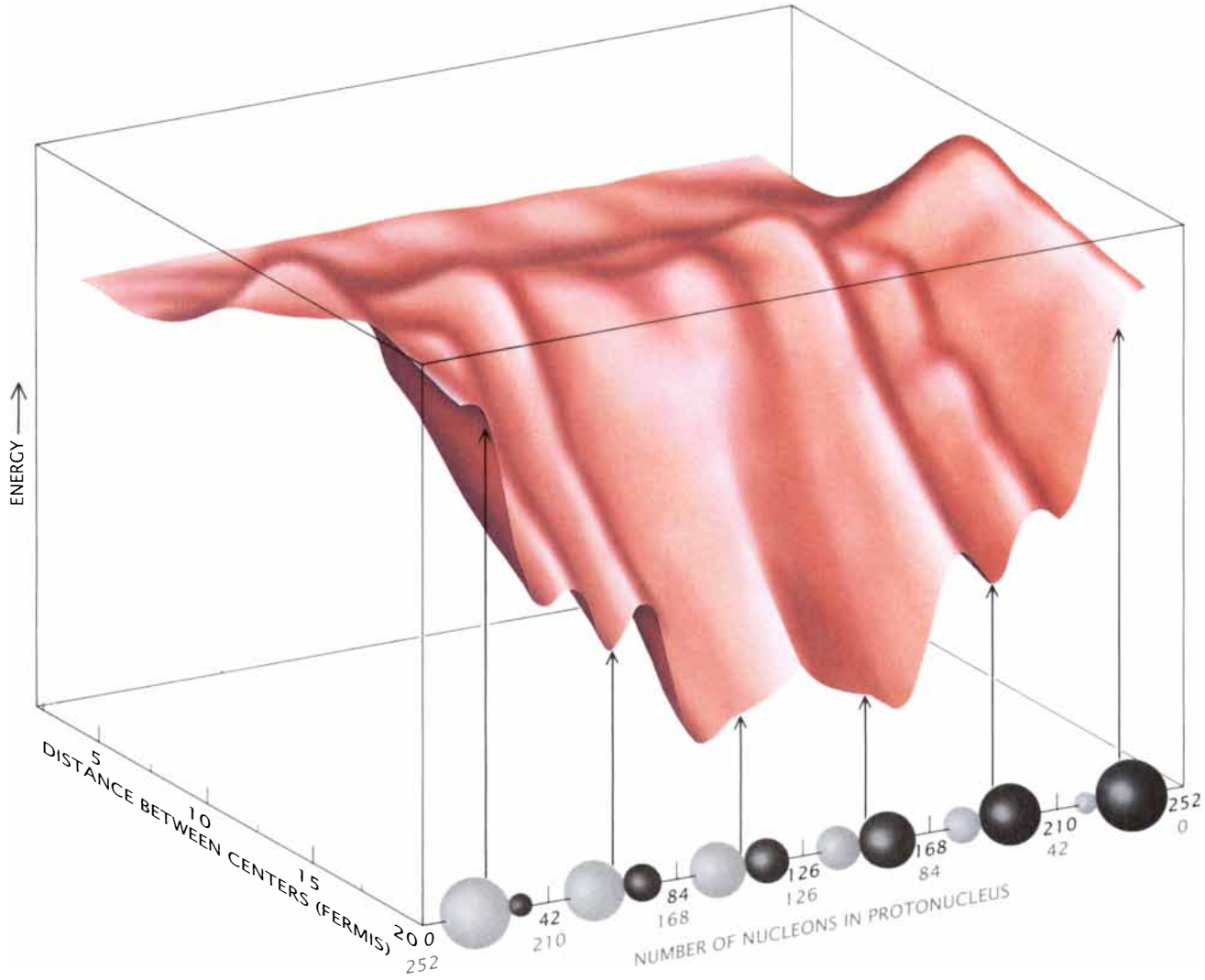
cause mass is equivalent to energy, it can also be said that if the energy of a nucleus is greater than the total energy of both the daughter nucleus and the alpha particle, then the nucleus can decay, emitting an alpha particle.

Why does the parent nucleus have more energy than the alpha particle and the daughter nucleus combined? The energy of the parent nucleus includes not only the energy associated with the mass of all the protons and neutrons but also a binding energy—the energy required to hold the nucleons together. Much of the binding energy arises from the energy needed to overcome the Coulomb repulsion of the protons. The binding energy is on average eight million electron volts (MeV) for every nucleon in the nucleus. The binding energy per nucleon can deviate greatly from the average. It is about 7 MeV for helium 4 and about 9 MeV for iron 56.

A high-energy nucleus can spontaneously transform to a lower-energy nucleus and in the process eject an alpha particle at great speed. Until 1928 physicists could not predict how much energy should be released in the decay process. In that year the Russian-American physicist George Gamow explained the process of alpha decay in terms of the newly conceived theory of quantum mechanics.

To understand Gamow's contribution, one might consider how the alpha particle behaves in the realm of the nucleus. An alpha particle traveling within a nucleus is somewhat like a ball rolling on a hilly surface. The height of each point on the surface is equivalent to the energy required to move a particle a given distance away from the center of the nucleus. The shape of the surface is actually somewhat like a deep depression on the top of a hill because energy is required to overcome the nuclear force and then energy is liberated as the Coulomb force pushes the alpha particle away. The energy required to climb out of the depression is known as the potential barrier.

The energy that the alpha particle absorbs from other nucleons propels the alpha particle (ball) within the deep depression. According to theories of classical physics, the only way for an alpha particle to escape the



ENERGY SURFACE of the nucleus of nobelium 252 suggests the many ways that the nucleus can spontaneously evolve into two protonuclei—precursors to the fragments emitted in fission. Energy is shown as a function of the distance between the protonuclei and of their size. As depicted in the lower portion of the graph, one protonucleus can have more nucleons (protons and neutrons) than the other protonucleus, yet the number of nucleons in the nobelium 252 nucleus must equal the total number of nucleons in the two protonuclei. Nobelium 252 is most likely to decay into protonuclei that are relatively low in energy. Therefore, the valleys in the energy surface indicate a likely pathway for fission. One valley shows the formation of gadolinium 158 and strontium 94; a second valley indicates osmium 192 and iron 60; a third leads to radon 214 and sulfur 38.

nucleus is if the alpha particle is propelled over the potential barrier (the ball is pushed out of the depression). In general, the alpha particle cannot gain enough energy, however, to surmount the potential barrier. According to quantum mechanics, there is a chance that the alpha particle will penetrate the potential barrier (the ball will tunnel through the hill). Gamow was able to calculate this chance as well as the kinetic energy of the emitted alpha particle.

After 1900 no new nuclear reactions were discovered until the experiments of the German physicists Otto Hahn and Fritz Strassman in 1939. By bombarding a sample of uranium with neutrons, they found that the uranium nucleus would fission: the nucleus broke apart into two fragments that had roughly the same mass and charge. One year later Konstantin A. Petrzhak and Georgii N. Flerov of the Joint Institute for Nuclear Research in Dubna, U.S.S.R., discovered that uranium could fission spontaneously. From 1939 to 1980 the large fragments from nuclear fission along with alpha, beta and gamma rays were the only radioactivities observed in the laboratory.

In the early 1940's Danish physicist Niels H. Bohr introduced a theory to explain nuclear fission for heavy elements such as uranium. Bohr exploited the fact that nuclear forces act over a short range. He assumed that a nucleon interacts only with its nearest neighbors, just as a molecule in a liquid affects only molecules nearby. A large nucleus, he thought, could be modeled after a drop of liquid.

A nuclear drop vibrates to some extent as it absorbs energy. Because of the vibrations, the drop can deform into two smaller nuclear drops connected by a long neck. As the distance between the two smaller drops increases, the potential barrier (the nuclear forces between the two drops) decreases. The smaller drops can then penetrate the potential barrier as long as the energy of the decay products (the small drops) is less than the energy of the deformed nucleus. Bohr calculated the chance that the two fragments will penetrate a nucleus, a quantity called the barrier penetrability. His predictions agreed quite well with the measurements of spontaneous fission in uranium.

Yet several questions lingered. Why did certain nuclei fission and not others? What determined the rate at which nuclei decay? What makes one nucleus more stable than another? Over the years three models of the nucleus tried to answer these questions: the shell model, the collective model and the two-center shell model.

In 1949 Maria Goeppert Mayer of the Argonne National Laboratory and Johannes H. D. Jensen of the University of Heidelberg independently introduced the spherical shell model of the nucleus. Like electrons in an atom, they proposed, the nucleons move according to the general principles of quantum mechanics: protons and neutrons occupy a set of discrete states, each corresponding to a particular amount of energy and a particular range of motion. The Pauli exclusion principle underlies this arrangement. The principle holds that a proton cannot occupy an energy state filled by another proton. The same is true of neutrons. As a result, each proton fills one energy state, starting with the state that has the least energy and filling as many states as there are protons. The neutrons fill another set of energy states. If the protons and neutrons are arranged in the lowest-energy configuration, the nucleus is said to be in its ground state.

The difference in energy between one state and the next higher state defines an energy gap. The size of these gaps is small in most cases, but in others it is large. The relatively large gaps separate the energy states into groups, or shells. In similar fashion, the energy states that electrons occupy in atoms are separated into shells.

The resemblances between nuclear shell structure and atomic shell structure are striking. If the electrons of an atom completely fill one or more shells, as is the case for helium and neon, the atom is stable: it is chemically inert. If the shells of a nucleus are completely filled, as are those of calcium and lead, the nucleus is stable and consequently spherical.

The first nuclear shell can be filled with as many as two protons and two neutrons; the second shell can be filled with up to six of each nucleon; the other shells are also filled with a certain number of protons and neutrons. The result is that one can usually predict the stability of a nucleus just by counting the number of protons and neutrons. Stable nuclei usually consist of a "magic number" of protons or neutrons; that is, they have 2, 8, 20, 28, 40, 50, 82, 126 or 184 protons or neutrons. Nuclei that have double magic numbers are particularly stable—for example, calcium 48 (20 protons and 28 neutrons) or lead 208 (82 protons and 126 neutrons).

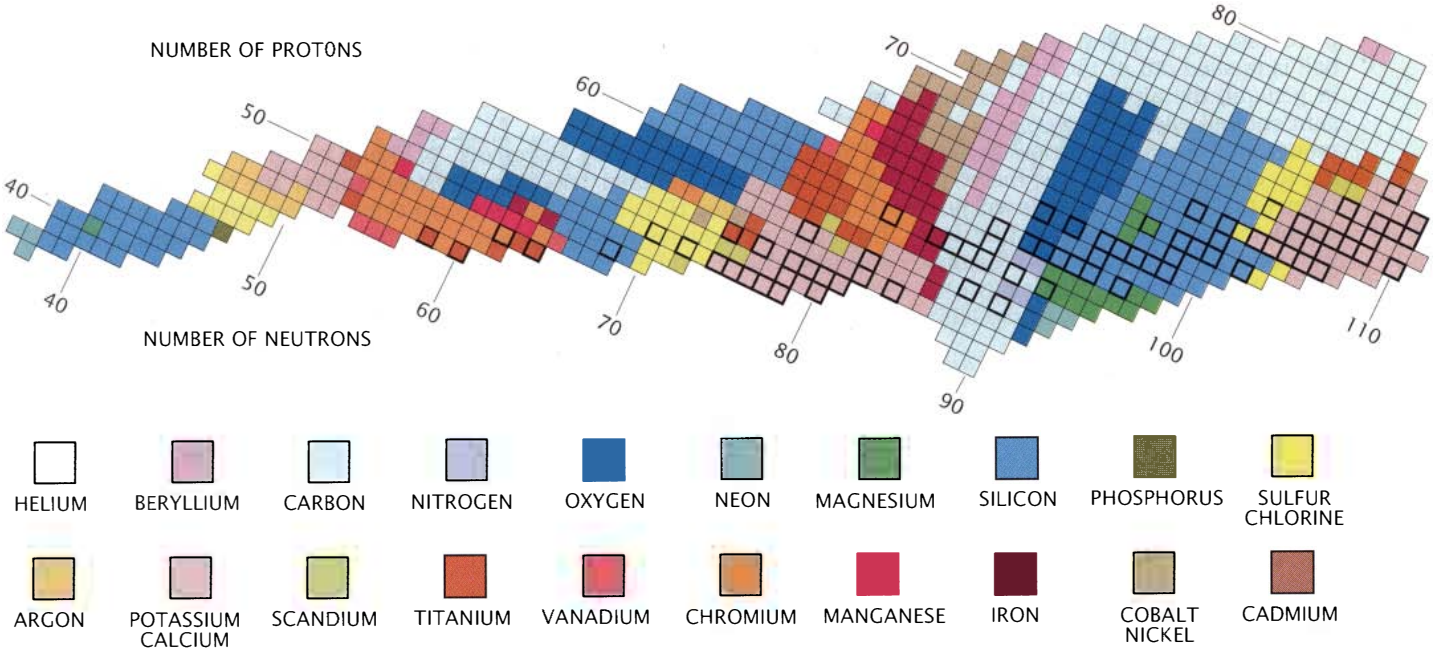
Although the shell model gave a fairly accurate description of nuclear stability, it did not yield the shape of the nucleus, nor could it reveal how that shape might change. In 1952 Aage N. Bohr (son of Niels) and Ben R. Mottelson of the Niels Bohr Institute in Copenhagen attempted to address this problem. They proposed the collective model, which combines certain aspects of both the shell model and the liquid-drop model. From the shell model Bohr and Mottelson retained the idea that the nucleons in the unfilled outer shells of a nucleus move under the influence of the filled inner shells. Yet, whereas the shell model assumes a somewhat rigid structure, the collective model holds that the outer part of the nucleus can deform when the outer nucleons move with respect to the nucleons of the inner nucleus. This collective motion, or deformation, derives from the liquid-drop model.

Most nuclei are prolate spheroids (cigar-shaped); some are oblate (disk-shaped). These deformations require that the nucleus gain or lose energy. A nucleus whose shape changes slowly as energy is added is called a hard nucleus; a nucleus whose shape deforms rapidly from additional energy is referred to as soft.

Although the collective model described how nuclear shape could change, it did not suggest how a nucleus could break apart or how two nuclei could fuse together. The explanation came in 1969, when one of us (Greiner) and his students at the University of Frankfurt proposed the two-center shell model. The model was based on the work of Sven G. Nilsson and Peter Möller of the University of Lund in Sweden. Exploiting a technique developed by Villen M. Strutinsky of the Nuclear Research Institute in Kiev and V. V. Pashkevich of the institute in Dubna, Joachim A. Maruhn of the University of Frankfurt and one of us (Greiner) later investigated the implications of the two-center shell model for fission processes.

Like the original collective model, the two-center shell model predicts the shapes of individual nuclei that are close to the ground state. In addition, the two-center shell model describes the numerous ways that a nucleus can spontaneously evolve into two protonuclei: precursors to the fragments emitted in fission.

For example, lead 208, which has 82 protons and 126 neutrons, can transform into a protonucleus of zirconium 104 (40 protons and 64 neutrons)



and a protonucleus of molybdenum 104 (42 protons and 62 neutrons). Lead 208 is a very stable (double magic) nucleus whose nuclear shells are completely filled. Even so, the energy of the nucleons in the outer shells exceeds the energy of the potential barrier. In other words, a nucleon in the outer shell can overcome the attraction of the inner nucleons and can move away from the center of the shells.

As the nucleon does so, it attracts some of the other nucleons in the outer shell, and they too start to move away from the original center. Eventually the new clusters can form two stable protonuclei: molybdenum 104 and zirconium 104. The protonuclei can remain in this state. The protonuclei can also overcome their mutual attraction and fission into independent nuclei. In other cases the protonuclei recombine to form lead 208.

The two-center shell model yields several methods for depicting the many ways in which a nucleus can spontaneously evolve into two protonuclei. For instance, one can consider two protonuclei that are a certain distance apart. Then one assumes that the first protonucleus has a certain number of nucleons and that the second protonucleus has all the remaining nucleons. The two-center shell model predicts how much energy is required to transform the original nucleus to this given state. One then asks what would happen if the first protonucleus gains another nucleon or what would happen if the protonuclei move further apart. In this way, one develops a three-dimensional map of the energy with respect to the mass of a protonucleus and the distance between the protonuclei.

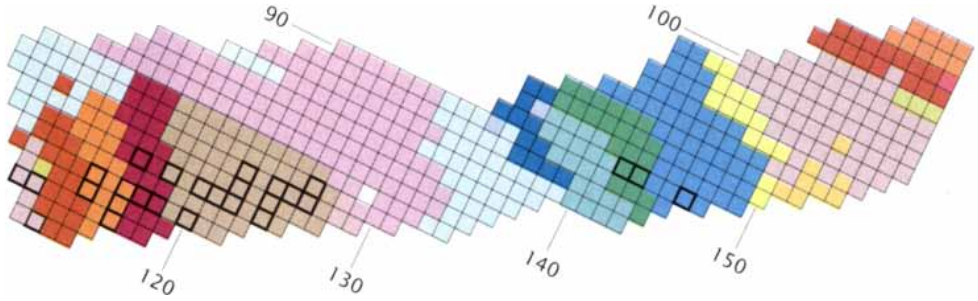
The result is a potential energy surface that has many mountains and valleys [*see illustration on page 60*]. The mountains correspond to high-energy, unstable states. The valleys correspond to relatively stable nuclei and protonuclei. Like the alpha particle, there is a small chance that a protonucleus will tunnel through a mountain to reach another valley. This chance (or barrier penetrability) depends on the height and width of the energy mountain. The maps also reveal how much kinetic energy is available to propel the fragments.

The two-center shell model has provided a precise description of many nuclear phenomena. It has predicted many new fission processes as well as the fusion of light nuclei into superheavy nuclei. Just as the valleys in the energy surface indicate that a particular nucleus is likely to decay into two lighter nuclei, the valleys can also show that two particular nuclei are likely to fuse to form a heavier nucleus. Thus, the two-center shell model led to the prediction and discovery of both new elements and new radioactivities [see "Creating Superheavy Elements," by Peter Armbruster and Gottfried Münzenberg; SCIENTIFIC AMERICAN, May, 1989].

In 1977 we predicted the existence of a nuclear process called superasymmetric fission and the emission of a new type of radioactivity. Unlike ordinary fission, superasymmetric fission produces two fragments that differ greatly in mass and charge. The emission of the smaller of these two fragments produces radiation known as cluster radioactivity. The cluster is usually several times larger than an alpha particle.

Superasymmetric fission can be observed in many nuclei. For example, a nobelium 252 nucleus can decay into radon 214 and sulfur 38, a cluster emission. This reaction represents a decay from an unstable state to a relatively stable state. Whereas nobelium 252 is very unstable, a radon 214 nucleus, which consists of 86 protons and 128 neutrons, is close to the very stable structure of lead 208, whose shells are completely filled with 82 protons and 126 neutrons.

Why does nobelium 252 decay into radon 214 instead of the even more stable lead 208? The answer becomes apparent when one employs the two-center shell model. One can calculate the potential energy surface for nobelium 252 as a function of the distance between the protonuclei and of the mass of one of the protonuclei. The surface reveals three valleys, two caused by the superasymmetric-fission process, the third by the ordinary fission process. One of the valleys corresponding to superasymmetric fission indicates the formation of a protonucleus having a mass number of 214. More specifically, two protonuclei formed from nobelium 252 and sepa-



SUPERASYMMETRIC FISSION should cause many species of nuclei to release clusters of nucleons. For each species the colored boxes indicate the cluster most likely to be emitted. For instance, a tin 112 nucleus, which is composed of 50 protons and 62 neutrons, most often ejects the cluster chromium, a nucleus that contains 24 protons. The number of neutrons in the cluster is not specified in the chart. But the white box refers specifically to the cluster helium 5, which has three neutrons, and the lavender boxes denote the cluster beryllium 8, which has four neutrons. Boxes bordered by heavy lines show stable nuclei found in nature. The chart is based on calculations by Dorin N. Poenaru of the Central Institute of Physics in Bucharest.

rated by 20 fermis are in a low-energy state when the mass of one protonucleus is 214 and the mass of the other is 38. Once the protonuclei become independent nuclei, they are again susceptible to other fission processes.

In 1978 we and Dorin N. Poenaru of the Central Institute of Physics in Bucharest started to predict the types of nuclei that should produce cluster radioactivity. For each candidate we tried to determine the barrier penetrability. We found that a radium 224 nucleus is one million times less likely to emit a cluster radioactivity of carbon 14 as it is likely to emit an alpha particle. Similarly, a sample of thorium 230 will emit the cluster neon 24 about 1,000 times less frequently than it will emit helium 4. In 1979 Poenaru, Marin Ivascu and one of us (Sandulescu) demonstrated that alpha decay is a type of superasymmetric fission. From this result they tried to predict the absolute half-life of several elements. The absolute half-life is the time required for half of a sample of a given element to undergo alpha decay or any other type of fission. Our calculations agreed well with measurements of absolute half-lives.

To guide experiments in the search for new radioactivities, we and our co-workers needed to answer several questions: What species of nuclei will emit cluster radioactivity at the highest rate? What types of cluster radioactivities can be emitted by a nucleus? What are the relative emission rates of one type of cluster radioactivity versus another? In order to solve these problems, one must search systematically through about 2,200 kinds of nuclei. For each possible parent nucleus, one has to trace all combinations, which amounts to one million possibilities. This is a difficult task to accomplish, even with modern computing technology.

In 1980 a team of Romanian and German theorists therefore began to create a computer program known as the analytical superasymmetric-fission model. Since its development, Poenaru, Ivascu, their colleagues and one of us (Greiner) have exploited the model extensively. We produced a comprehensive table of various kinds of cluster radioactivity for more than 800 kinds of nuclei. We found that any isotope that has more than 40 protons can decay through these new modes.

We expected, however, that available technology would allow us to observe emission rates only for nuclei that break apart into lead 208 or into its neighboring nuclei. These include more than 150 kinds of cluster radioactivities whose estimated half-lives are shorter than 10^{23} years and whose kinetic energies are about 2 MeV per nucleon.

Because the clusters are emitted against a large background of alpha particles, the most important factor is the branching ratio: the ratio of cluster emissions to alpha emissions for the same nucleus. We calculated the branching ratios for the emission of various light nuclei and found that the cluster emissions most easily detected are carbon 14, neon 24, neon 25 and magnesium 28. The branching ratios for the emissions of carbon 14 from radium 223 turned out to be one billion alpha emissions for every one carbon 14 emission. This was the largest predicted branching ratio.

After a decade of advances in nuclear theory, experiments began to demonstrate the predictive power of the two-center shell model. Then in 1984 Hans Jürgen Rose and G. A. Jones of the University of Oxford found the first convincing evidence for cluster radioactivity. To prepare a sample for study, they chemically separated actinium 227 from natural uranium 235. Actinium 227 naturally decays into radium 223, which was expected to emit the cluster radioactivity carbon 14.

The actinium source was positioned near a detector. As the source emitted nuclear fragments, the detector would register their energy. Although the detector could distinguish between carbon 14 fragments and alpha particles, it could still be inaccurate. If three alpha particles hit the device at about the same time, the triplet could mimic the energy signature of a carbon 14 fragment. In order to avoid the problem, the experiment was designed so that the record of a triplet would be a much rarer event than the detection of a carbon 14 fragment.

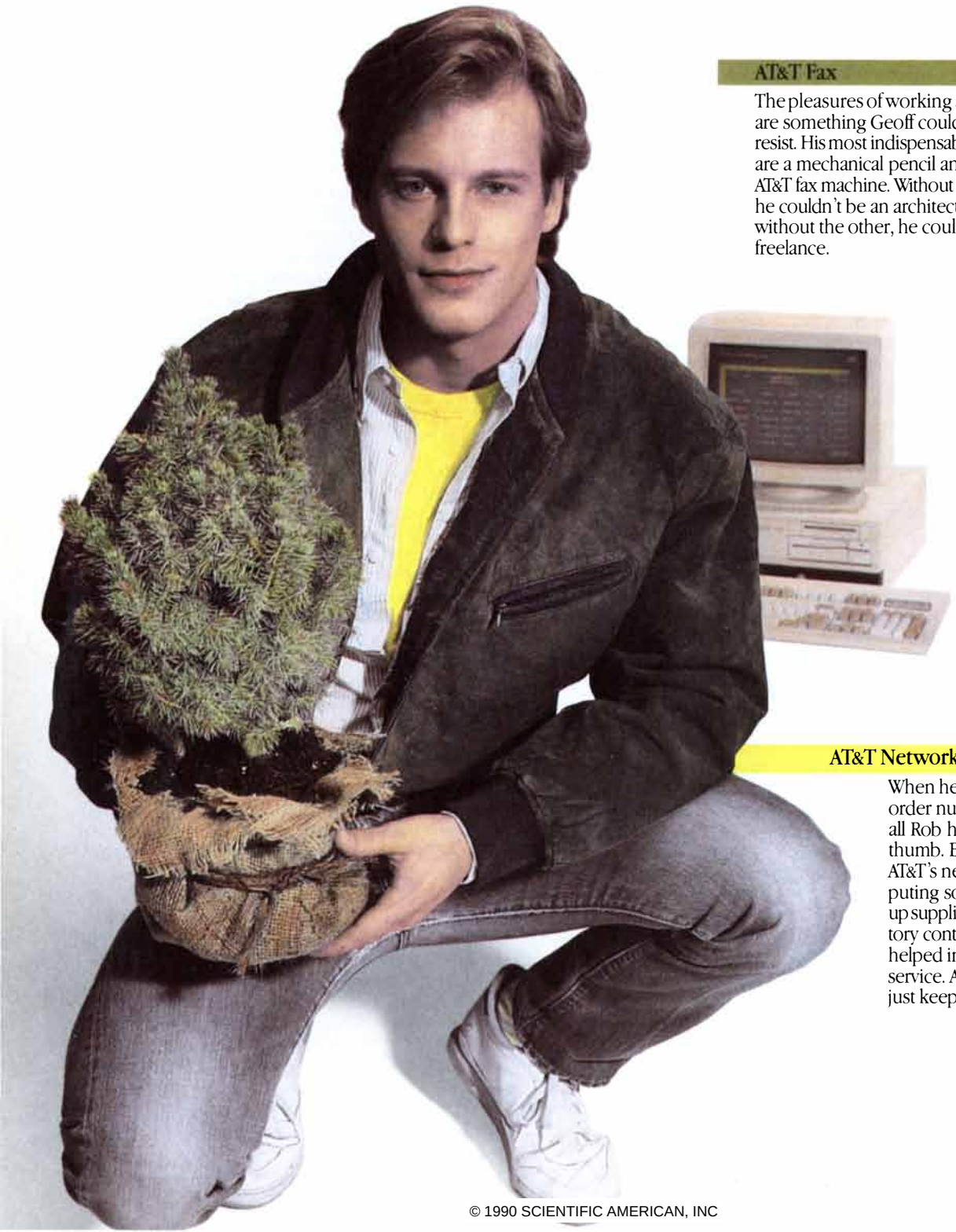
The search ran for 189 days. During that time the detector registered 65 billion alpha particles and 11 events involving carbon 14. Rose and Jones concluded that carbon 14 was ejected as radium 223 decayed into lead 209. They calculated that radium 223 releases one carbon 14 nucleus for every billion or so alpha particles. A few months later Alexei A. Ogloblin and his co-workers at the I. V. Kurchatov Institute of Atomic Energy in Moscow performed a similar experiment that confirmed the work of Rose and Jones.

To improve these measurements and detect other types of cluster radioactivity, investigators in France and the U.S. adopted a different strategy. They developed devices to filter the cluster radioactivities from the multitude of alpha particles. In France, E. Hourani and his colleagues at the University of Paris built the superconducting solenoid spectrometer SOLENO; in the U.S., Walter Kutschera, Walter F. Henning and their co-workers at the Argonne National Laboratory constructed an Enge split-pole magnetic spectrograph.

The SOLENO spectrometer generates a magnetic field by means of a super-

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

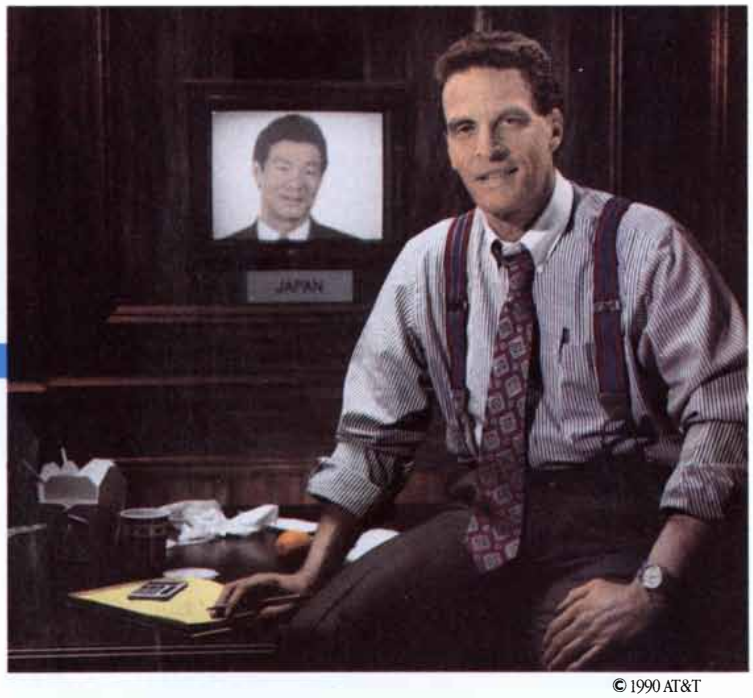
The pleasures of working at home are something Geoff could not resist. His most indispensable items are a mechanical pencil and his AT&T fax machine. Without the one, he couldn't be an architect, and without the other, he couldn't freelance.

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.

AT&T Healthcare Solutions

How does she find the time to always give her patients that little extra attention and encouragement? With the help of AT&T's physician information network, Grace has immediate access to everything from lab results to ultrasound images. Giving her more time to do what she always wanted to do. Take care of people.



AT&T International Videoconferencing

After the merger was announced, Michael wanted to know how the news was received on the Tokyo Stock Exchange. For him, it wasn't just a question of seeing the numbers. He wanted to see the reaction. And AT&T's international business video allowed him the luxury of a face-to-face meeting on the spot. After all, when it comes right down to it, every business is a people business.

The way things get done.

AT&T
The right choice.

conducting solenoid, or coil of wire. The magnetic field guides charged nuclei from the source through the spectrometer. Because each type of nucleus has a characteristic mass and charge, the field guides each type to a particular point within the spectrometer. Because alpha particles are light and most have a positive charge of two, alpha particles tend to strike a point in front of the detector. Carbon 14 nuclei, which are heavier and have a charge of six, however, converge on the detector itself.

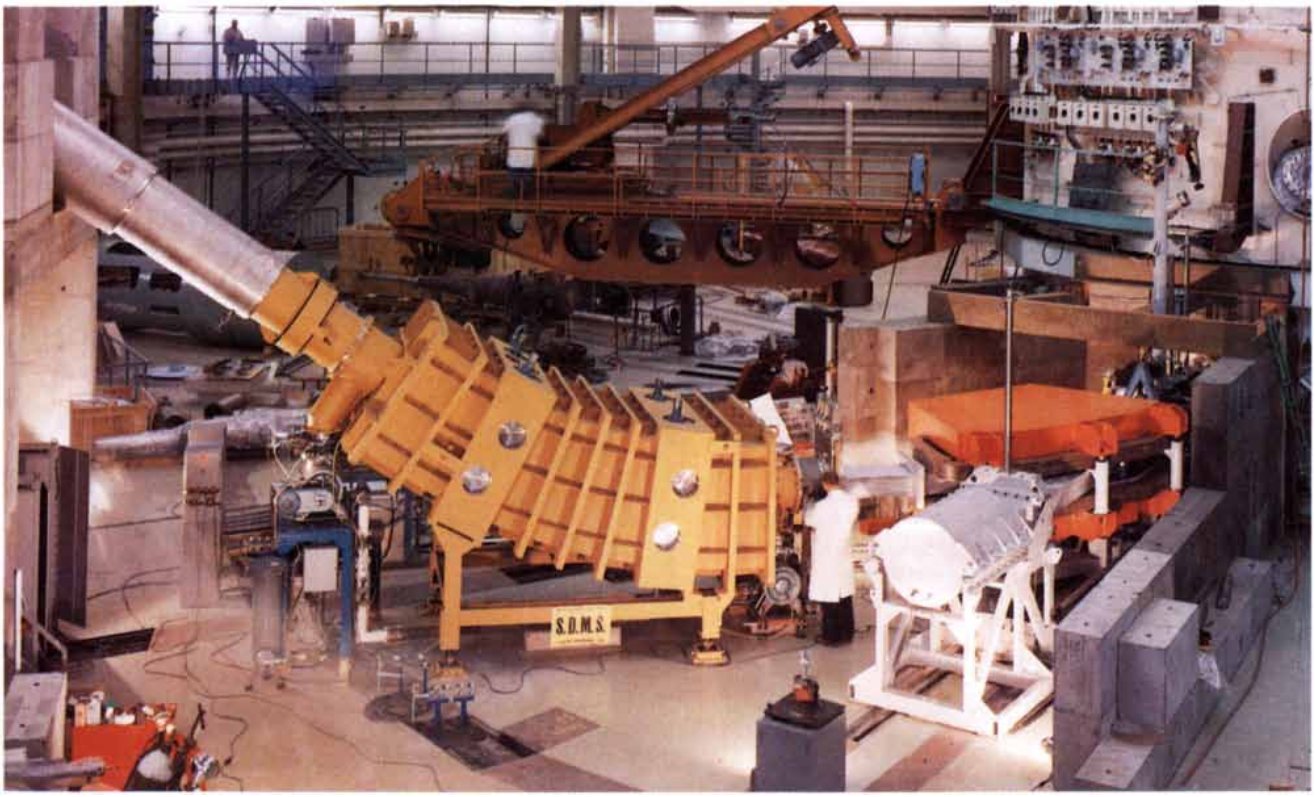
In 1985 investigators from the University of Paris used the SOLENO spectrometer for the first time to observe the decay of radium 223 nuclei. In just five days they succeeded in registering 11 events involving carbon 14 nuclei. They also proved beyond a reasonable doubt that the detected nuclei had a mass number of 14. Later the workers demonstrated that radium 222 nuclei and radium 226 nuclei also spontaneously emit carbon 14 nuclei. They do so about 10 times less frequently than does radium 223.

Shortly thereafter workers at the Argonne National Laboratory began experiments with the Enge spectrograph. The instrument operates in a manner similar to the SOLENO spectrometer. It has a different kind of detector, however, which is filled with gas that ionizes when hit by cluster radioactivity. During a six-day trial the spectrograph registered 24 carbon 14 nuclei emanating from radium 223.

The most sensitive devices developed for the detection of cluster radioactivity are track detectors, which record radioactivities on a special film. A large nucleus that strikes the film produces a defect. A small nucleus such as helium 4 passes through the film and only rarely leaves a mark. When the film is developed in a chemical bath, it will be etched away faster in the area of the defect than in unaffected areas. In fact, the etch rate of a defect area is proportional to the charge and the mass of the nucleus that created the defect. If one views the film under a microscope, one can therefore determine the point of impact of the nucleus as well as its charge and mass.

In 1984 P. Buford Price and his coworkers at the University of California at Berkeley built the first such track detector at CERN, the European laboratory for particle physics near Geneva. They made the detector out of a polycarbonate film, which is sensitive to nuclei that have more than two protons. As a source of nuclear material, they exploited beams of nuclei. The nuclear beams were generated at the Isotope Separator On-line (ISOLDE) facility at CERN. The ISOLDE facility produces many types of beams, each containing nuclei that have the same number of nucleons but different proportions of neutrons and protons. For the cluster radioactivity experiments, Price and his co-workers used beams containing nuclei that had either 222 or 223 or 224 nucleons. The beams were directed into the detector, a small chamber lined with polycarbonate film. When the nuclei in the beam hit a plate at the back of the chamber, they sometimes emitted cluster radioactivities that exposed the film.

A similar device led to the discovery of a second type of cluster radioactivity, namely, neon 24. The discovery was made at the beginning of 1985 by Price and his colleagues and independently by S. P. Tretyakova, her co-workers and one of us (Sandulescu) at the institute in Dubna. Both groups employed polyethylene terephthalate detectors, which record only those particles hav-



***LOHENGRIN* SPECTROMETER at the Laue-Langevin Institute in Grenoble, France, extracts nuclei of a particular mass and charge from among many others. The nuclei emerge from the reactor (*right*) and travel through a magnetic field generated by the apparatus painted orange. The magnetic field focuses particular nuclei to a point behind the concrete walls (*left*).**

ing more than six protons. The Dubna workers observed the emission of neon 24 nuclei from protactinium 231, uranium 233 and thorium 230. The Berkeley group, which found that uranium 232 emits neon 24 nuclei, measured one neon 24 emission for every 500 billion alpha particles.

In 1986 Price and his co-workers at Berkeley developed a phosphate glass track detector to study the decay of nuclei into neon, magnesium and silicon. The glass track detector does not record many extraneous signals because the lightest particle capable of producing a track is oxygen. The Berkeley group chose to investigate uranium 234 nuclei based on predictions by Poenaru and us that uranium 234 emits significant amounts of both neon and magnesium. Three detectors were exposed for about three months. The group observed 14 neon events and three magnesium events. In this way, they discovered a new radioactivity, magnesium 28, and the first nuclear species, uranium 234, known to emit three kinds of cluster radioactivities: helium, neon and magnesium.

These three forms and carbon have been the only cluster radioactivities detected so far. These clusters are emitted from various nuclei at particular rates; half-lives range from 10^{11} to 10^{26} years. The analytical superasymmetric-fission model has reliably predicted, within 1.5 orders of magnitude, the half-lives of all nuclei that emit these clusters.

Cluster radioactivity was only one triumph of the two-center shell model. Another was the prediction of cold fission. In this newly discovered process, a nucleus splits into two "unexcited" nuclei. More specifically, nucleons that form a low-energy state of one nucleus can rearrange themselves to form the low-energy states of two new nuclei. Unlike the ordinary (hot) process, the energy released in cold fission does not excite the emitted nuclei into high-energy states. The nuclear fragments from cold fission are therefore more spherical and less elongated than the nuclear fragments from ordinary fission. According to the two-center shell model, a nucleus that is a candidate for ordinary fission has a small chance of undergoing cold fission instead.

A number of workers have devised techniques for isolating cold-fission events from among the many ordinary (hot) fission events. In 1981 Claude Signarbieux and his colleagues at the research center in Saclay, France, performed the first such experiments. They bombarded a sample with neutrons to induce fission. Using two detectors on opposite sides of the sample, they determined the mass and kinetic energy of the fragments.

Soon after, Friedrich Gönnenwein of the University of Tübingen in West Germany and his co-workers developed another technique for detecting cold fission. They worked with two fission-fragment spectrometers, known as *Lohengrin* and *Così fan tutte,* at the Laue-Langevin Institute in Grenoble, France. The spectrometers allow measurements of fragment masses, nuclear charges and kinetic energies. Recently Gönnenwein and his colleagues constructed yet another type of device to measure cold-fission events. Their new device surpasses the sensitivity of other detectors by a factor of 10.

These experiments provided indirect evidence for the cold fission of uranium 233, uranium 235 and plutonium 239. The amount of energy that is released in fission is at most the Q value, the difference in energy (mass) between the parent nucleus and its daughter products. This energy is distributed between the motion of the whole nucleus and the excitation of the nucleons within the nucleus. In the experiments the kinetic energy measured for some fission events equaled the Q value. This implied that no energy was available for fragment excitation and therefore that cold fission had occurred.

Workers continue to search for evidence for cold fission. Furthermore, Martin Greiner (son of Walter) and Werner Scheid of the University of Giessen have predicted that cluster radioactivities can also be emitted in a hot—that is, excited—state.

Yet another triumph of the two-center shell model was the prediction of bimodal fission. Through the process of cold fission, some heavy nuclei divide into two symmetric hard nuclei in which the nucleons fill a number of the nuclear shells. Because these heavy nuclei can also fission in the more common manner (they can split into two elongated nuclei of the same size), they are said to exhibit bimodal fission. Many workers have studied bimodal fission, including Maruhn, Pashkevich, Möller, Karl Depta of the University of Frankfurt, James R. Nix of the Los Alamos National Laboratory and Adam Sobiczewski of the Institute for Heavy-Ion Research (GSI) in Darmstadt.

These workers contributed to the prediction that bimodal fission would be most evident in the decay of fermium 264 nuclei and its close nuclear relatives. The ordinary fission of the fermium 264 nucleus should produce two elongated fragments. The cold fission of fermium 264 should create two hard spherical nuclei of tin 132. The shell structure of tin is particularly stable because it contains a magic number of protons (50) and a magic number of neutrons (82). In certain situations one can think of fermium 264 as a kind of nuclear molecule containing two tin nuclei.

Over the past four years E. Kenneth Hulet of the Lawrence Livermore National Laboratory and his colleagues have constructed a detector sensitive to nuclei that have more than 100 protons and more than 156 neutrons. They have found evidence for bimodal fission in nobelium 258, kurchatovium (rutherfordium) 260 and three isotopes of mendelevium (258, 259 and 260).

Experiments continue in the areas of new radioactivities, cold fusion and bimodal, superasymmetric and cold fission. We expect that the sensitivity of detectors will be greatly improved by the application of technologies already available. We also hope that the productive interplay between theory and experiment will continue to reveal new radioactivities and increase the depth of understanding of them. We can only anticipate the new messages these ambassadors might bring.

FURTHER READING

HIGH MASS-ASYMMETRY DISTRIBUTIONS OF FISSIONING NUCLEI. A. Sandulescu, H. J. Lustig, J. Hahn and W. Greiner in *Journal of Physics G: Nuclear Physics,* Vol. 4, No. 11, pages L-279–L-285; 1978.

NEW TYPE OF DECAY OF HEAVY NUCLEI INTERMEDIATE BETWEEN FISSION AND α DECAY. A. Sandulescu, D. N. Poenaru and W. Greiner in *Soviet Journal of Particles and Nuclei,* Vol. 11, No. 6, pages 528–541; November/December, 1980.

CLUSTER RADIOACTIVITIES. Walter Greiner, Marin Ivascu, Dorin Poenaru and Aurel Sandulescu in *Treatise on Heavy Ion Science,* Vol. 8. Edited by D. A. Bromley. Plenum Press, 1989.

COMPLEX RADIOACTIVITY. B. Price in *Nuclear Physics A,* No. 502, pages 41c–58c; 1989.

TOPICAL REPORT: A NEW RADIOACTIVITY. A. Sandulescu in *Journal of Physics G: Nuclear and Particle Physics,* Vol. 15, pages 529–554; 1989.

Earthquakes in Stable Continental Crust

Earthquakes can strike even in stable crust, well away from the familiar earthquake zones at the edges of tectonic plates. What accounts for these enigmatic events?

by Arch C. Johnston and Lisa R. Kanter

Charleston, South Carolina, is a long way from earthquake country. The rim of the Pacific, the Mediterranean and central Asia experience most of the world's earthquakes, both spectacular and small. Yet in 1886 a quake several times larger than the event that struck San Francisco in October of last year severely damaged much of Charleston. Near the coast, the water-saturated soil liquified and erupted in geysers of sand and mud. Scores of people were killed, and the city's recovery from the devastation of the Civil War, 20 years before, was slowed.

By most standards Charleston occupies a geologically quiet zone. Plate boundaries, where the 100-kilometer-thick plates of rock that make up the earth's outer shell collide, rift apart or slide past each other, are the scene of mountain building, volcanism, active faults and most earthquakes. The San Andreas fault, for example, the boundary between the plate that carries most of North America and the plate underlying the Pacific Ocean, cuts directly through California. Yet Charleston lies well away from a plate boundary; the eastern edge of the North American plate is thousands of kilometers offshore, in the middle of the Atlantic Ocean. The area has not experienced the tectonic (crust-deforming) activity of a plate margin since the opening of the Atlantic some 180 million years ago.

The earthquake that collapsed buildings in Newcastle, in eastern Australia, in December of last year took place in crust with a similar history. As these earthquakes and other, larger events—in the Mississippi Valley in 1811 and 1812 and in the province of Kutch, India, in 1819—demonstrate, major earthquakes can and do occur in plate interiors. The rarity of such earthquakes and the fact that the largest ones took place before good instruments were developed have made their study a seismological backwater, but their mere occurrence is both a fascinating geophysical puzzle and a source of concern in planning critical facilities such as nuclear power plants. Questions of seismic hazard, in fact, helped to motivate our own systematic study of these earthquakes, which was sponsored by the Electric Power Research Institute in Palo Alto, Calif., and done with Kevin J. Coppersmith and his co-workers at Geomatrix Consultants in San Francisco and Ann G. Metzger of Memphis State University.

We sought answers to two major questions: Just how much seismic activity does take place within the stable parts of continents? And are there specific geologic features that make some areas of stable crust particularly susceptible to earthquakes? We began by studying North America alone, but it soon became clear that the fairly short record of these rare events on a single continent would not provide enough data for reliable analysis. Hence, we decided to substitute space for time—to survey earthquake frequency and distribution in stable continental areas worldwide.

What counts as stable continental crust? The challenge was to identify the very quietest parts of continents, well away from the tectonic activity of plate boundaries. Clearly, the bedrock shields, some of them more than three billion years old, that form the ancient hearts of continents and the "platforms" of sediment-covered bedrock that surround the shields qualified as stable crust. At the other end of the spectrum, the plate boundaries themselves had to be excluded. After that, though, the choices of what to define as stable crust became more complex.

Tectonic activity driven by plate interactions can extend well away from plate boundaries. Where plates converge and an oceanic plate is subducted beneath a continental plate, volcanic activity and mountain building may take place in a band hundreds of kilometers wide on the overriding continent, as it does in the Andes. If both plates bear continents or continental fragments, the collision can produce a region of folded, uplifted and faulted terrain several thousand kilometers wide, typified by the Himalayas, the Tibetan plateau and central Asia as far north as Lake Baikal.

The converse process—the rifting of a continent to form two new plates separated by a spreading center such as the Mid-Atlantic Ridge, where new ocean floor is generated—also results in a broad zone of deformed crust. As the rifting begins, the continental crust becomes stretched and thinned, riddled with faults and veined with

ARCH C. JOHNSTON and LISA R. KANTER are, respectively, director of the Center for Earthquake Research and Information (CERI) at Memphis State University and assistant research professor there. Johnston, who is also professor of geophysics at Memphis State, joined CERI in 1979 after getting his Ph.D. from the University of Colorado at Boulder. He has often served as an adviser to both private and public organizations, including congressional subcommittees, on seismic hazards in the eastern U.S. Kanter got her Ph.D. at Stanford University in 1983 and moved to CERI in 1985. Johnston writes that the present study was triggered when nobody, in an entire roomful of eminent seismologists at a workshop, could identify the largest earthquake recorded in stable continental crust.

volcanic material. Once a full-fledged oceanic spreading center develops between the rifted edges of the continent, the continental crust stops stretching. Its edges cool and subside, creating what are known as passive margins.

As plate boundaries migrate and reorient themselves over geologic time, regions deformed by compression or stretching become part of the plates' stable interiors. They are no longer the scene of intense tectonic activity, even though they do still experience stresses—in general, compressive ones—transmitted from the distant edges of the plates. Along with the ancient shields and platforms, then, we counted as stable continental crust mountain belts older than about 100 million years (the Appalachians and Urals but not the Alps and Himalayas, for example) and passive margins older than about 25 million years (the margins of the Atlantic but not of the Red Sea). We also included ancient failed rifts, where a full-fledged spreading center did not develop and the stretched crust subsided into a broad valley or a sediment-filled trough. By these criteria nearly two thirds of all continental crust qualifies as stable.

To determine the earthquake activity within these least seismic parts of the continents, we could not rely on recent instrumental records alone; even global records are too sparse. We had to cast as wide a net as possible and include earthquakes in historical accounts. In North America such accounts begin in the 16th and 17th centuries, but in Europe they cover 1,000 years and in China more than 2,000.

Simply counting the events in the historical record was not enough; we needed a measure of their size. Current practice recognizes the moment-magnitude scale, devised by Hiroo Kanamori of the California Institute of Technology and Thomas C. Hanks of the U.S. Geological Survey, as the most reliable way to gauge earthquake size. Earlier scales, such as the Richter scale, rank earthquakes according to the amplitude of specific seismic waves, which can vary for an earthquake of a given energy depending on the wave frequency at which the amplitude is measured. The moment-

FAULT IN ANCIENT CRUST was the site of a 1988 series of moment-magnitude 6 earthquakes near Tennant Creek, in Australia's Northern Territory. The events are among the few recorded stable-continent earthquakes in which the surface ruptured. J. Roger Bowman of the Australian National University made the image.

magnitude scale, in contrast, is based directly on the physical process at the heart of an earthquake: the slip of rock along a fault.

The scale is based on seismic moment: the size of the force couple (two opposite forces) that caused the fault to rupture. Seismic moment is equal to the surface area of the rupture multiplied by the average displacement of the rock along the rupture and the rock's rigidity. Moment magnitude (designated **M**) is linearly proportional to the logarithm of the moment, so that an event of **M** = 7 is 32 times larger than one of **M** = 6 and 1,000 times the size of an **M** = 5 event.

Seismic moment can be calculated even if the ruptured fault is inaccessible—hidden underwater or buried at a depth of many kilometers. Thanks to techniques developed by Keiiti Aki, who is now at the University of Southern California, and many later workers, an earthquake's seismic moment, and hence its moment magnitude, can be determined from low-frequency components of seismic waves, which can be recorded at a distance by seismographs. Yet nearly half of the stable-continent earthquakes in our list, and the majority of the largest ones, took place before the invention of the seismograph in the late 19th century. We had to develop a way of estimating seismic moment from the evidence in historical accounts: "intensity areas," or descriptions of an earthquake's effects on people, buildings and the landscape and their areal extent.

Fortunately, we had both intensity areas and seismic moments directly computed from instrumental data for more than 50 stable-continent earthquakes. By applying statistical regression techniques to the data, we were able to develop a correlation between intensity areas and seismic moment for earthquakes in stable crust; we could then assign moment magnitudes to other events that lacked instrumental data. The largest stable-continent earthquakes on record turned out to be the great earthquakes centered in New Madrid, Missouri, during the winter of 1811–12, which had moment magnitudes ranging from **M** = 8.1 to **M** = 8.3.

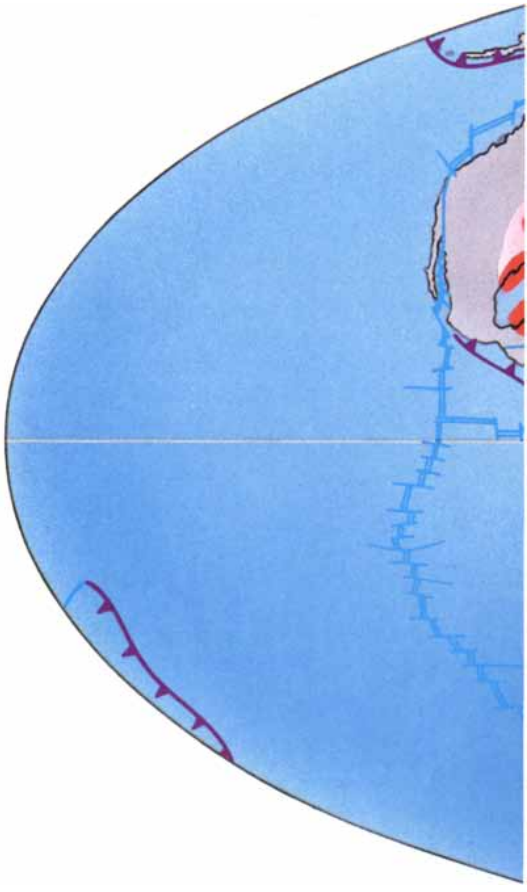
In comparison, the largest earthquake recorded in any setting, the 1960 plate-boundary event in Chile, measured **M** = 9.5, or 63 times larger. (That magnitude represents an energy equivalent to that of an average hurricane but released in one or two minutes rather than over 10 days.) The New Madrid events, however, may have been felt over a wider area than any other earthquake in history; the strong rock of plate interiors transmits seismic waves far more efficiently than the fault-laced crust near plate boundaries. Centered 1,000 miles inland, the New Madrid quakes damaged masonry as far away as the East Coast and collapsed scaffolding erected around the U.S. Capitol [see "A Major Earthquake Zone on the Mississippi," by Arch C. Johnston; SCIENTIFIC AMERICAN, April, 1982].

The Kutch, India, earthquake of 1819 ranks second to the New Madrid events in magnitude: **M** = 7.8, based on sketchy intensity areas. It also yields an independent check of our empirical correlation between intensity areas and seismic moment. In contrast to the faults responsible for most stable-continent earthquakes, which are deeply buried under the layers of sediment that cover much old crust, the fault rupture at Kutch reached the surface.

The effects were dramatic indeed. A scarp between six and nine meters high and at least 90 kilometers long was thrust up; it became known locally as the Allah Bund, or "Wall of God." Land to the north of the Allah Bund was elevated, while to the south what had been a low-lying *rann*—a salt flat that flooded periodically—was further depressed. Sindree Fort, built well before the event, had stood on a small rise in the *rann*. When the motion of the fault depressed the land to the south, the fort was so deeply submerged that soldiers had to escape by boat from the upper turret.

We might note that the juxtaposed uplift and submergence make the Kutch event one of the most vivid illustrations of the elastic-rebound theory of faulting. The theory (not formally elaborated until after the great San Francisco earthquake of 1906) holds that an earthquake is a sudden release of elastic strain built up in the rock; it predicts that rock on opposite sides of the fault will move in opposite directions. At Kutch, absolute sea level acted as a reference plane to dramatize the crust's opposite vertical displacements.

EVENT	HOST STRUCTURE	M
1. New Madrid, 1812	rift	8.3
2. New Madrid, 1811	rift	8.2
3. New Madrid, 1812	rift	8.1
4. Kutch, 1819	rift	7.8
5. Baffin Bay, 1933	margin	7.7
6. Taiwan Straits, 1604	margin	7.7
7. South Carolina, 1886	margin	7.6
8. Nanai, 1918	margin	7.4
9. Grand Banks, 1929	margin	7.4
10. Basel, 1356	rift	7.4
11. Hainan Island, 1605	rift	7.3
12. Exmouth Plateau, 1906	margin	7.2
13. Libya, 1935	margin	7.1
14. Portugal, 1858	margin	7.1
15. So. Tasman Rise, 1951	margin	7.0



STABLE REGIONS make up about two thirds of all continental crust. In defining stable crust, the authors excluded plate boundaries and broad, diffuse regions of active or recent deformation. The 15 largest earthquakes (on a scale of moment magnitude, M) ever recorded in stable crust (*above*) have all occurred where crust has been stretched and thinned over the past 250 million years.

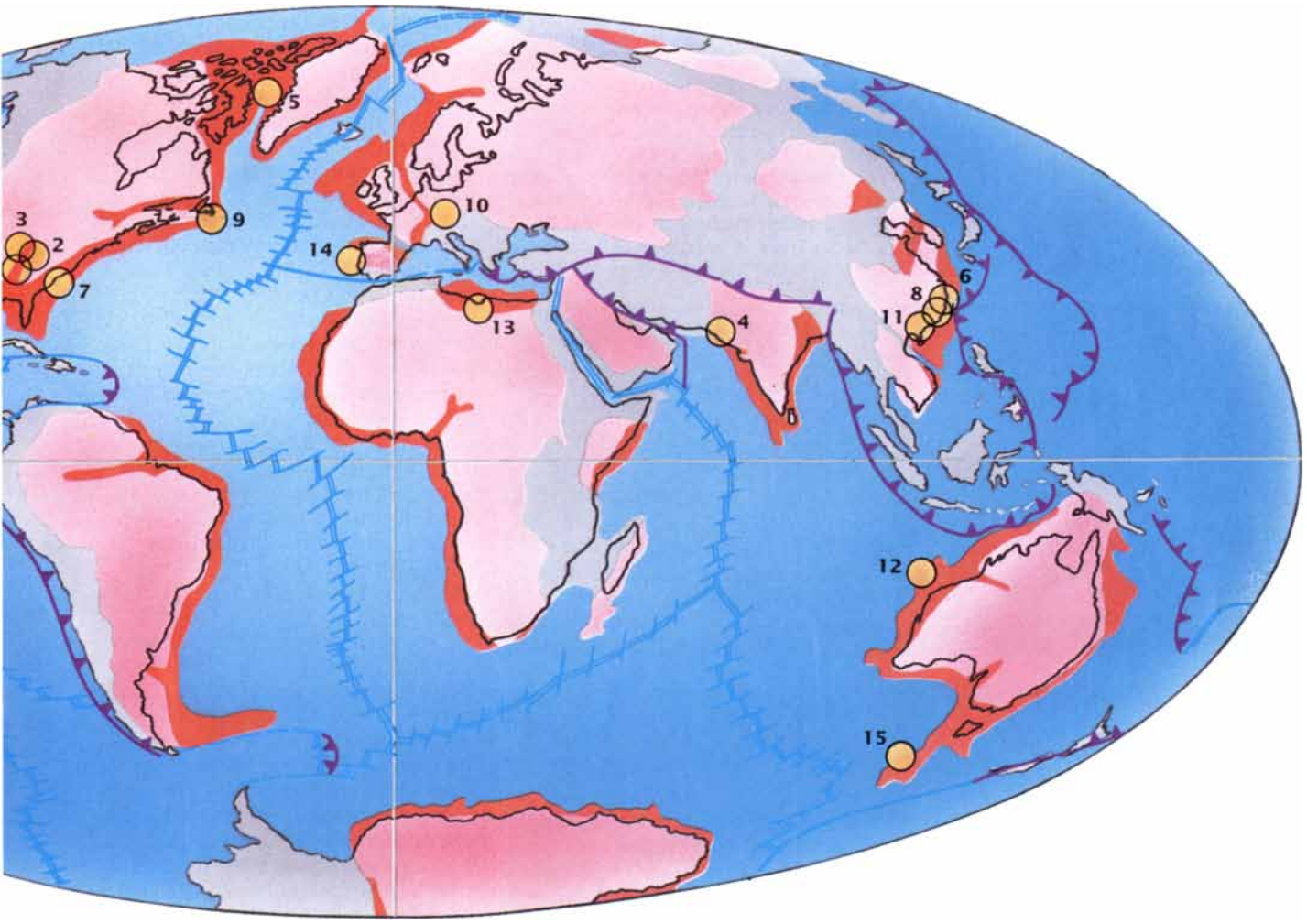
More to the point, the obvious fault movement meant that we could calculate the seismic moment directly. The British had made careful measurements of the fault length and slip. To calculate seismic moment, we also needed to know how deep the rupture zone extended and the rigidity of the rock, but reasonable assumptions can be made for both quantities. The resulting moment was within a factor of two of the moment estimated from the intensity reports, which added to our confidence that we could adequately reconstruct seismic moment from historical accounts alone.

The New Madrid and Kutch earthquakes are only the most prominent elements in a data set of more than 800 stable-continent events of $\mathbf{M} = 4.5$ or more. The number may seem large, but it is the sum of a global record covering centuries or millennia. (Many more earthquakes of $\mathbf{M} = 4.5$ or greater take place along plate boundaries in just a year.) To find out just how seismic-moment release on stable continents compares with earthquake activity worldwide, we constructed a so-called frequency-magnitude diagram. Such a diagram has earthquake magnitude plotted on its horizontal axis and the number of events equaling or exceeding a given magnitude plotted on its vertical axis. If the vertical scale is logarithmic, frequency-magnitude plots for any given seismic zone yield a nearly straight line, sloping down from the smallest, most frequent events to very rare, catastrophic events.

Even if an earthquake catalogue overlooks very small events or covers too short a period to include very large ones, the consistent relation between frequency and magnitude means the plot can be extrapolated to estimate the frequency of earthquakes over a wide magnitude range. It is then straightforward to integrate the results to find the total seismicity—the rate at which seismic moment is being released—for the region covered by the data. Our calculations for stable continental crust yielded a total moment release of about 10^{26} dyne-centimeters (the conventional units of seismic moment) per year, or less than .5 percent of the global total.

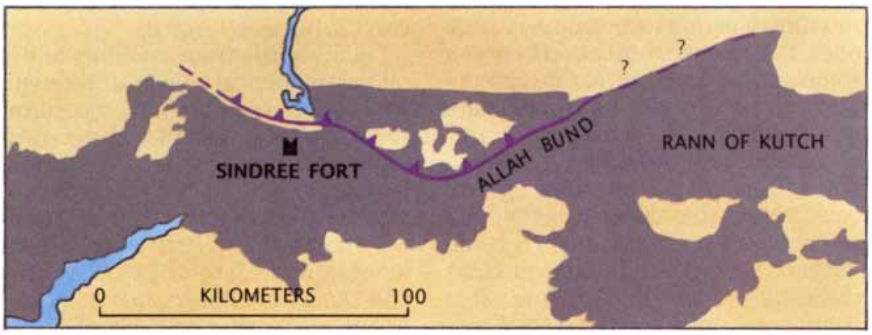
No doubt our data set is incomplete. It probably includes every earthquake larger than $\mathbf{M} = 7$ during the time of the records. Small and moderate-size events in remote regions may never have been recorded, however, which adds uncertainty to the frequency-magnitude diagram. Still, there is reassuring agreement between the seismicity figures and another source of information about processes in stable continental crust: independent measurements of the rate at which stable crust is being deformed.

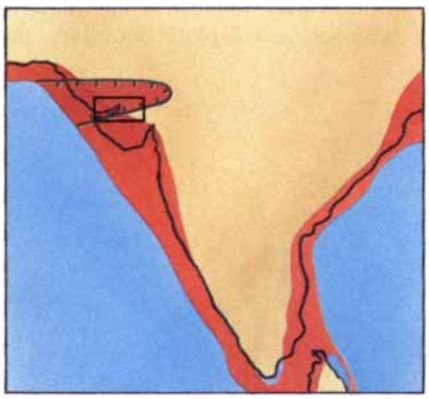
Using signals from satellites or distant astronomical objects as references, geophysicists are now measuring tiny changes in distance between receiving stations thousands of kilometers apart as the intervening crust is stretched or compressed [see "Studying the Earth by Very-Long-Baseline Interferometry," by William E. Carter and Douglas S. Robertson; SCIENTIFIC AMERICAN, November, 1986]. For two stations separated by a zone of active tectonics near a plate boundary, the



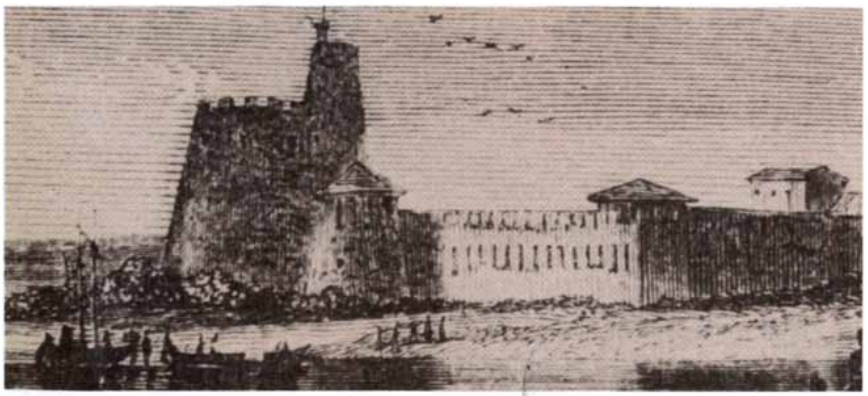
deformation can amount to several centimeters a year. Stable continental crust also deforms in response to the compressive stresses transmitted from plate boundaries, but preliminary measurements suggest the rate is minuscule, about one part in 10^{10} per year—perhaps a millimeter over thousands of kilometers.

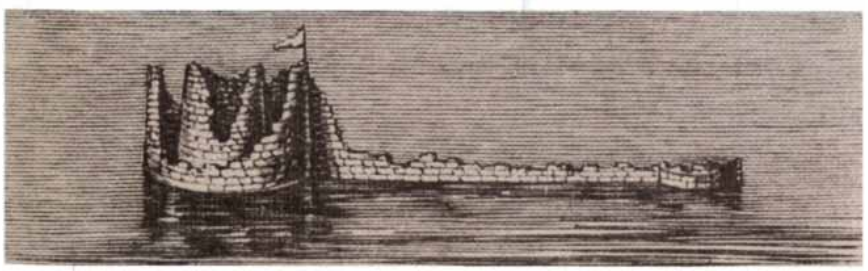
If one assumes that all the stresses built up in this deformation are ultimately released in earthquakes (a reasonable assumption for the strong, brittle crust of continental interiors), it is possible to calculate how much seismic activity should accompany the deformation. B. V. Kostrov of the Soviet Academy of Sciences has shown that the moment-release rate is then proportional to the volume of crust multiplied by the rate of strain, or deformation. A strain rate of 10^{-10} per year yields a predicted rate of moment release in stable continental crust that agrees nicely with what we calculated from our earthquake data set.





KUTCH EARTHQUAKE of 1819 was unusual among stable-continent events in that the fault rupture reached the surface. It uplifted a scarp between six and nine meters high (the Allah Bund) across the salt flats known as the Rann of Kutch (*above*), in western India (*left*). Land to the north was uplifted; to the south the land subsided. Sindree Fort had stood on a rise (*below*); after the earthquake only the turret remained above water, as illustrated 19 years later from another angle (*bottom*). The woodcuts are from 1853 edition of Sir Charles Lyell's classic book *Principles of Geology*.





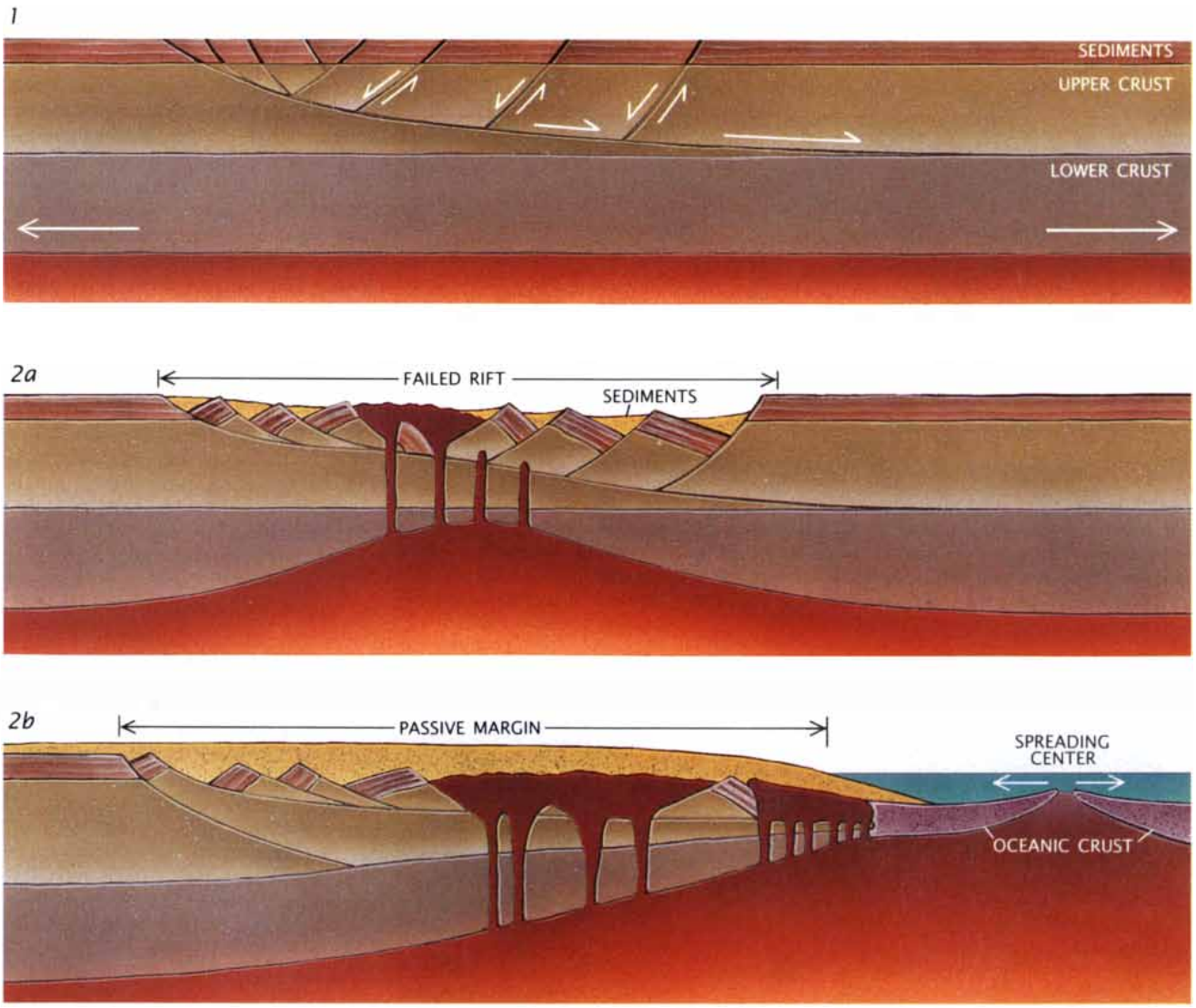
Our figure for earthquake activity in stable crust is actually an average of a rate that varies widely from continent to continent. Antarctica and Greenland are quite devoid of significant earthquakes, perhaps because the massive ice sheets there stabilize any faults, inhibiting slip. The stable parts of South America and Asia (Siberia) are seismically quieter than the remaining stable continental regions. It may be that these variations reflect differences in average, continent-wide levels of stress—a possibility that might be tested by systematic comparisons of deformation rates within different continents.

Within the stable areas of a single continent, too, seismicity varies widely. To find out why, we worked with David B. Bieler, then at Memphis State, to analyze the tectonic history of stable continental regions and correlate it with earthquake distribution. One crustal characteristic stood out as a universal precondition for the largest stable-continent earthquakes.

Like many seismologists, we had expected stable-continent events to be associated with specific zones of weakness in the crust. Ultimately, of course, earthquakes in stable crust are the product of compressive stresses that originate at plate edges. Yet variations in the stress field within a continent seemed unlikely to be the sole explanation of the earthquake distribution. As Mary Lou Zoback of the USGS, Mark D. Zoback of Stanford University and others have shown from worldwide studies of earthquake faulting and of boreholes drilled into bedrock (which deform under the influence of the prevailing stress field), the compressive forces in continental interiors are consistent in orientation across vast areas.

Hence, an earthquake seemed to require a region weakened by past tectonics, where these stresses could be relieved. Crust dissected by old faults, such as an ancient mountain belt or a failed rift, might qualify. The pervasive compression within a continent might sometimes reactivate these old faults, causing them to slip and generate earthquakes.

Our findings largely bear out this reactivated-fault model. Areas that have undergone extension at some time in the past, such as passive margins and failed rifts, are more likely to experience an earthquake of any size than, say, ancient shields are. Of the more than 800 earthquakes we studied, almost half (49 percent) fell within such extended crust, even though it makes



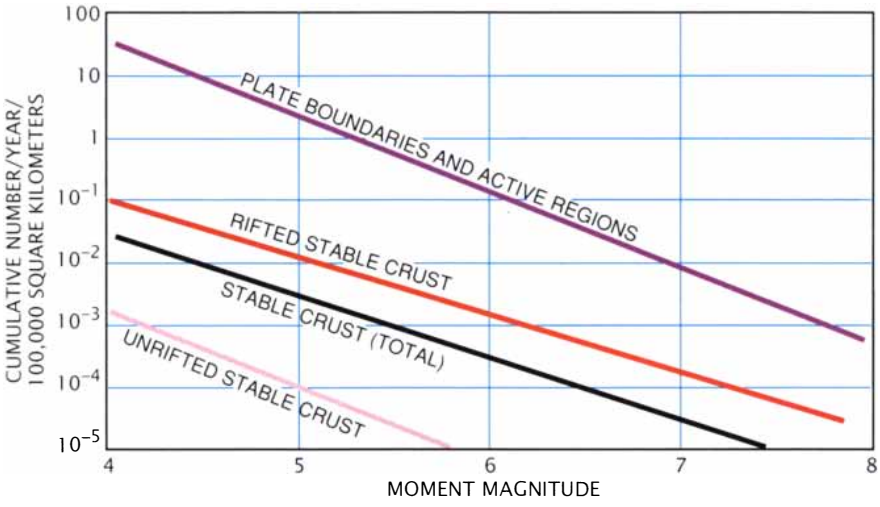
RIFTED CRUST, the site of most large stable-continent earthquakes, forms over millions of years when extensional forces break up the brittle upper crust into blocks separated by active faults (*1*). Continued extension thins the crust, allowing magma to well up. If the stretching stops, the stretched crust remains as a failed rift (*2a*), in which new sediments accumulate. If the stretching continues, the continental crust eventually ruptures, giving birth to a spreading center at which new oceanic crust is produced. The stretched continental crust subsides, resulting in sediment-covered passive margin (*2b*).

up only a quarter of all stable continental crust.

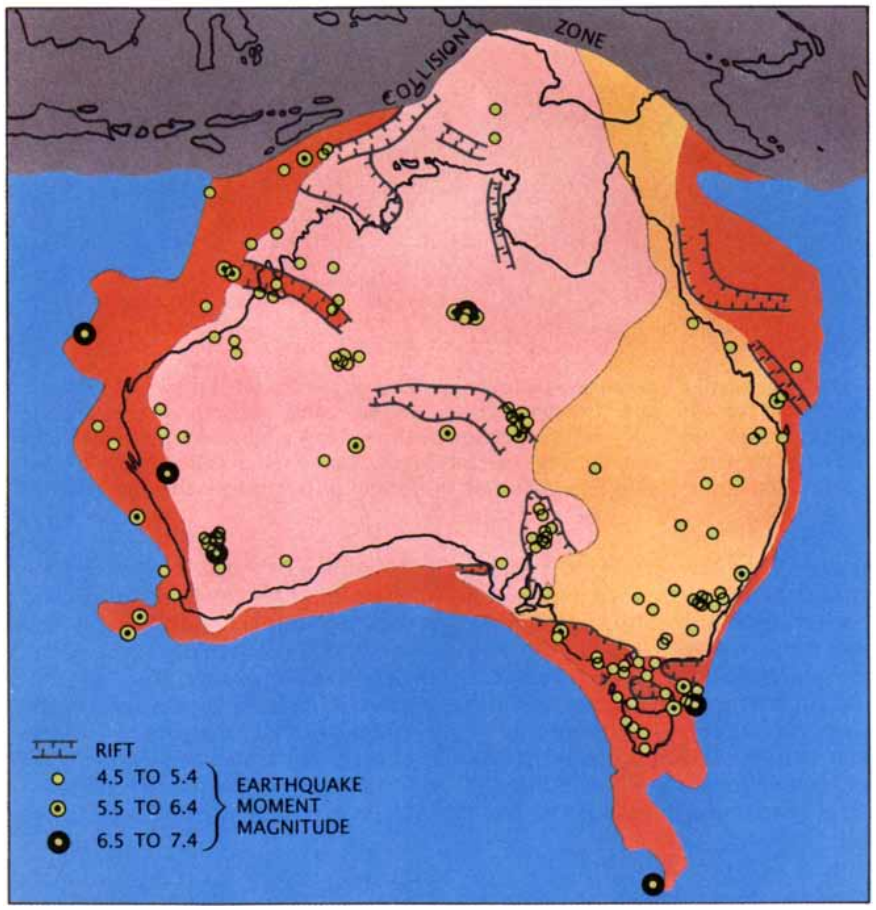
The correlation grows stronger with increasing magnitude; whereas only 46 percent of earthquakes smaller than **M** = 6 took place in extended crust, the figure is 60 percent for events between **M** = 6 and **M** = 7 and 100 percent for the largest events, of **M** = 7 or more. A failed rift underlies the site of the New Madrid and Kutch earthquakes, for example, and the Charleston event took place on an extended passive margin. For all these largest events and most of the smaller ones on rifts or passive margins, the extension was comparatively recent—older than 25 million years (by definition) but younger than 250 million.

The correlation between extended crust and earthquakes of moderate size would have been stronger had it not been for one continent full of exceptions: Australia. The ancient, unrifted crust of western and central Australia has experienced several earthquakes that rank among the largest events not associated with extended crust. The events, which took place in a remarkable 12-hour sequence

FAULT IS REACTIVATED, millions of years after it was formed, in the process thought to generate the largest stable-continent earthquakes. Faults created in continental crust by ancient extensional stresses (*1*) may lie dormant for many millions of years, gradually accumulating a blanket of sediments (*2*). Later the compressive stresses within continental interiors may reactivate a fault, which slips in another direction, either vertically or horizontally, and generates an earthquake (*3*).



SEISMIC ACTIVITY at plate boundaries and the associated active regions is compared with activity in several kinds of stable crust by a frequency-magnitude diagram. Earthquakes of all sizes are several hundred times more common at plate boundaries than they are in stable regions. Within stable crust, earthquakes take place more often in rifted areas; the correlation is the strongest for the largest events.



CONTINENT PEPPERED WITH EARTHQUAKES is Australia, which consists entirely of stable crust—an ancient core (*pink*) surrounded by somewhat younger crust (*orange*) and passive margins (*red*). A region of ongoing plate collision borders Australia to the north. Australia follows the general rule that the largest stable-continent earthquakes occur within crust that has undergone recent extension, but its largest onshore events have taken place in unrifted ancient rock and have produced surface faulting. The map, which shows both recent and historical events, is based on one in a series prepared by the authors for all the stable continental regions on the earth.

at Tennant Creek in 1988 (**M** = 6.2, 6.3 and 6.6), at Meckering in 1968 (**M** = 6.6) and at Meeberrie in 1941 (**M** = 6.8), also had other peculiarities. Like the Kutch earthquake but unlike most other stable-continent events, the rupture in the first two cases reached the surface and created a prominent fault scarp. Study of seismic-wave arrival times and waveforms showed that slip in these Australian earthquakes originated at depths of less than 10 kilometers and often only two or three kilometers—shallower than any other well-documented events in our data set.

Nevertheless, the very largest—and most dangerous—stable-continent earthquakes have taken place within crust that underwent fairly recent extension. One might also expect to find such earthquakes in zones of past continental collision—ancient regions of compressed and folded crust such as the Appalachians. Indeed, small earthquakes are significantly more common in such regions than they are in undeformed crust, and earthquakes of moderate size are slightly more common. No such pattern is seen for larger events, however; it may be that damage to the crust is more pervasive in extended regions than in compressed ones.

Yet if the reactivated-fault model alone were enough to account for all stable-continent earthquakes, one might expect smaller earthquakes to cluster in regions of extended crust as strikingly as larger events do. The much weaker correlation of small earthquakes with such regions raises the possibility that other mechanisms are also at work.

One alternative model for stable-continent earthquakes has been proposed by Mark Zoback, expanding on earlier work by Richard H. Sibson of the University of California at Santa Barbara. Zoback argues that the same compressive stresses found in the brittle upper crust of continents are also present in the lower crust, at depths of more than 20 or 25 kilometers. At those depths, high temperature and perhaps differences in composition make the crust ductile rather than brittle. Instead of cracking or bending elastically under stress, it slowly flows, and by yielding, Zoback suggests, it adds to the stress on the brittle crust above. If this ductile deformation somehow became concentrated in a small region, upper-crustal stresses might build up, leading to earthquakes. Because the properties of lower crust rather than those of shallower rock would determine

earthquake distribution, the events might seem to take place at random, as the smallest earthquakes in our record do.

The most important result of our study is that the largest earthquakes do not take place at random. Knowing that they are concentrated in failed rifts or passive margins is a promising start toward placing constraints on areas of earthquake risk. Yet it is hardly comparable to seismologists' ability to pinpoint earthquake-prone sites near plate boundaries, where many active faults are apparent at the surface and rupture often and regularly. Indeed, for many plate-margin faults it is possible to determine an approximate recurrence interval for major earthquakes, which can be as short as every 40 years or so. As a result, even though such earthquakes cannot yet be predicted with an accuracy of hours or days, they can sometimes be roughly forecast. The fault segment that was responsible for last year's earthquake south of San Francisco, for example, was identified in 1988 as being quite likely to sustain a large event within 30 years.

Forecasting stable-continent events —to say nothing of predicting them— is a difficult problem. The deeply buried faults responsible for most of them cannot be identified at the surface, and the faults do not betray themselves by rupturing frequently. Of the 15 earthquakes of **M** = 7 or greater in our catalogue, none was a repetition of an earlier recorded event of comparable size (if the three New Madrid earthquakes of 1811 and 1812 are counted as one).

Whether or not stable-continent earthquakes obey any periodic pattern at all is an open question. If they do, the recurrence interval must be longer than it is at plate margins—perhaps much longer. Indeed, the space-based measurements of infinitesimally slow deformation within the plates lend support to the idea that great earthquakes could recur at any one location only at intervals of thousands of years. (It is also possible, however, that if this intraplate deformation is concentrated in a few zones of weak crust, earthquakes would recur more frequently there. At New Madrid, for example, some workers estimate that major earthquakes, as large as those of 1811 and 1812, could recur as often as every 500 years.)

The best hope of determining recurrence intervals comes from a developing discipline called paleoseismology: the study of prehistoric earthquakes. Even if the earthquake fault itself is buried, a large event can leave lasting marks at the site, such as landslides and deformed soil layers. Occasionally these features include trapped organic matter—tree roots snapped off in the churning soil, say—which can make it possible to apply carbon dating to the event.

Studies of old soil-liquefaction features in Charleston have suggested that an earlier earthquake preceded the 1886 event by 1,100 years. In Oklahoma, Anthony J. Crone of the USGS and Kenneth V. Luza of the Oklahoma Geological Survey have measured and dated offset stream courses along the Meers fault (a feature marked by a three-to-five-meter-high scarp extending more than 25 kilometers within a failed rift). The Meers fault has been quiet in historical times, but the paleoseismological work suggests that an earthquake of about **M** = 7 took place there some 1,200 years ago.

For sites already identified by an historical earthquake or a visible fault, then, it may soon be possible to say something about the frequency of large events—a first step toward forecasting them and estimating seismic hazard. For the rest of stable continental crust our study provides a broad-brush picture of how seismic risk varies with crustal history. It also provides a wholesome reminder that, even in the quiet hearts of continents, "stable" is a relative term.

SOIL LIQUEFACTION can make it possible to date a large prehistoric earthquake. When a powerful earthquake shakes water-saturated sandy soil, the water pressure rises, and the soil becomes a viscous liquid that can erupt through overlying layers to the surface. The disturbed soil may trap organic matter that can be carbon-dated. In this excavation north of Charleston, S.C., the "sand blow" on the right (dated from bark fragments) is about 950 years old, the left one about 1,900 years old. The earthquakes responsible had a source different from that of the 1886 Charleston event.

FURTHER READING

INTRAPLATE SEISMICITY, REACTIVATION OF PREEXISTING ZONES OF WEAKNESS, ALKALINE MAGMATISM, AND OTHER TECTONISM POSTDATING CONTINENTAL FRAGMENTATION. Lynn R. Sykes in *Reviews of Geophysics and Space Physics*, Vol. 16, No. 4, pages 621-688; November, 1978.

SEISMICITY OF 'STABLE CONTINENTAL INTERIORS.' Arch C. Johnston in *Earthquakes at North-Atlantic Passive Margins: Neotectonics and Postglacial Rebound*. Edited by Søren Gregersen and Peter W. Basham. Kluwer Academic Publishers, 1989.

METHODS FOR ASSESSING MAXIMUM EARTHQUAKES IN THE EASTERN UNITED STATES. Kevin J. Coppersmith, Arch C. Johnston, Lisa R. Kanter, Robert R. Youngs and Ann G. Metzger in *EPRI Report RP-2556-12*. Electric Power Research Institute, 1990.

Likely Litigation

Companies wield lawsuits as a market-development tool

Legal twists, dense pictures, math that pays, creating currency

Last year E. I. Du Pont de Nemours ran into a brick wall. It had begun working on polymerase-chain-reaction (PCR) technology—a revolutionary method for multiplying snippets of DNA, first developed by investigators at Cetus. But when Du Pont tried to license the PCR technology from Cetus, it was rebuffed. In August, Du Pont retaliated: it sued Cetus, claiming that Cetus's PCR patents are invalid.

The Du Pont–Cetus dispute was only one of the thousands of intellectual-property suits launched last year. "Most patent lawyers believe that patents have become more valuable in the last five years," says Richard H. Kjeldgaard, an attorney with Burns, Doane, Swecker and Mathis in Alexandria, Va. As a result, companies are now wielding suits aggressively. Some hope to force competitors to pay royalty fees; others hope to crush competitors by eliminating them from a market altogether. These trends are most prevalent, moreover, in young industries that are testing the limits of property protection for the first time.

Corporate America's interest in intellectual-property rights warmed in the late 1970's as the U.S. began running trade deficits, says Michael W. Blommer, who heads the American Intellectual Property Law Association in Arlington, Va. Worried about competing with manufacturers both at home and abroad, companies took a closer look at their available assets—including their portfolios of patents.

The outcome of litigation, however, was uncertain. Because every court circuit tried both patent cases and appeals, there was a blizzard of conflicting decisions. "Your patent might or might not have been worth anything, depending on what court you were in," Blommer says.

Under the Reagan administration, however, patents became viewed as a "pro-competitive" tool, Kjeldgaard says. About the same time, Congress began clarifying patent laws that had not changed since the 1950's; it has since passed more than a dozen laws. The most important of these, attorneys agree, established a single appeals court for all patent cases in 1982.

The Court of Appeals for the federal circuit quickly began developing a more coherent set of precedents. It also clearly demonstrated the clout of patents with a 1985 decision over a battle between Polaroid and Eastman Kodak. The court upheld Polaroid's patents on instant cameras, cutting Kodak out of the business. Using patents to exclude a company altogether from a market "was a new idea," says Salem M. Katsh of Weil, Gotshal and Manges in New York City. "It made the upside in litigation stronger."

Congress also strengthened the patent protection against foreign infringers with the Trade Act of 1988. Foreign companies had long been prohibited from selling patented products in the U.S. without a license. But the same protection did not cover patented processes unless U.S. patent holders could prove they had incurred significant economic damage. The trade act lifted this requirement, thus providing both product and process patents with the same protection.

Even as patent rights have strengthened, the fledgling industries, such as computer software and biotechnology, have begun flexing their intellectual-property muscles. So far the existing copyrights and patents are proving to be resilient. For instance, in the late 1980's, a few courts began to rule that copyright protected the "look and feel" of computer programs, typically embodied by graphic elements, or icons. Those early decisions prompted a spate of lawsuits, including two launched by Apple Computer against Microsoft and Hewlett-Packard.

According to industry lore, Apple founder Steven P. Jobs first saw computer icons in 1979 when touring Xerox's research center. Apple's later use of icons in its software for personal computers helped it score huge successes; in contrast, Xerox fumbled in its attempts to enter the business.

Now that the courts are upholding copyrights, Xerox is looking for a payback for its early work. In early December, Xerox asked Apple to ante up license fees. When Apple refused, Xerox sued to have Apple's copyrights deleted from the federal register. In addition, Xerox is demanding more than $150 million in damages.

The biotechnology industry is also seething with suits. "The markets are brand-new. What would a recombinant AIDS drug be worth? We don't know," Katsh says. Some biotechnology companies have bet their future on a product or two; as a result, "they have a greater interest in being the sole supplier to a market," Kjeldgaard adds.

The uncertainty has spawned some bitter court feuds—and an inclination to turn to Congress for even stronger intellectual-property laws. One landmark case has been the lengthy dispute between Amgen and Genetics Institute over erythropoietin (EPO), a protein that stimulates the rate at which red blood cells are formed. Amgen is considered the scientific leader, says Robert P. Merges of Boston University School of Law; but a recent court decision that recognized both companies had some rights to EPO has left Amgen at a disadvantage, he adds. The case has been complicated by GI's affiliation with the Japanese firm Chugai Pharmaceutical. Amgen has since taken its case to Congress, arguing that existing patent laws are biased against domestic companies.

With encouragement from other biotechnology firms, notably Genentech, Congress is expected to consider at least two intellectual-property bills this year. One will aim to make it easier to obtain process patents for work based on recombinant technologies; another will try to extend protection against foreign infringers to intermediate biological products used in recombinant technologies.

Meanwhile the biotechnology industry will be watching the dispute over PCR between Du Pont and Cetus. It is a struggle over market share. Early last year Cetus granted an exclusive license to Hoffmann-La Roche to use PCR for diagnostic purposes. Roche hopes to become a major player in the diagnostic-equipment business, says Michael S. Ostrach, Cetus's general counsel; Cetus was eager to have a large share of that future.

Du Pont, on the other hand, did not want to be left out of a business that some have predicted could be worth as much as $1 billion within a decade. When Du Pont uncovered two papers published more than 10 years ago that it claims "clearly anticipated" PCR, the company sued Cetus for invalid patents, says George A. Frank, Du Pont's senior counsel.

Du Pont and Cetus agree on one point: PCR is worth the expense of a fight. The litigation may easily cost each company on the order of $2 million a year and may last more than two and a half years. (This does not include the expense of continuing research; Du Pont, for instance, plans to spend more than $15 million during the next three years on PCR.)

"So say we make up numbers," Ostrach says. "Say the market is worth $500 million." If Du Pont thinks it has a better than one-in-a-hundred chance of winning the lawsuit, "it's an even bet," Ostrach says. "We've done the calculations," Frank replies. Du Pont and Cetus will face off in court on August 7. *—Elizabeth Corcoran*

Not Just a Pretty Face

Compressing pictures with fractals

In a darkened room, as strains of Wagner play in the background, Michael Barnsley is staging a show. The screen of a Sun workstation goes blank for about 60 seconds, then a color portrait of a woman's face bursts into view, filling the display.

That portrait, Barnsley says, was generated from only 5,800 bytes of computer memory—an astonishingly tiny amount of computer space for storing such an image. How? Barnsley says he has discovered a technique for automatically compressing data by ratios of 500 to one or better.

In an age in which information is abundant, finding computer space to store all that data can be difficult. An uncompressed digital version of the woman's portrait would take up about three million bytes of memory; conventional compression techniques can pack these data into several hundred thousand bytes. To add color and video images to electronics products such as computers demands even tighter compression, however.

Among those pushing hard to find new compression schemes are Barnsley and Alan D. Sloan, both mathematicians formerly at the Georgia Institute of Technology. Several years ago Barnsley developed algorithms for drawing exquisite ferns and other images based on fractals—irregular, geometric patterns that repeat themselves on many scales [see "Science and the Citizen"; SCIENTIFIC AMERICAN, February, 1988]. But it was time-consuming work. "Very smart" graduate students spent about 100 hours encoding every picture, Barnsley recalls.

In 1987 Barnsley and Sloan decided to try building commercial techniques for compressing data based on fractals. They formed Iterated Systems, adopted a fern as a logo and hired a team of mathematicians. Then in mid-1988 "I discovered a class of theorems that allow us—with slight modification—to automatically compress any image," Barnsley says.

The woman's portrait, he says, was encoded in about an hour. Precisely how it works, however, is proprietary. People "are just going to have to trust my reputation," Barnsley states. (The company has applied for two patents and is completing a third application.) His insistence on secrecy has irked mathematicians who have some reservations about his claims.

They question whether it is possible to find a generic technique that can rapidly encode any image based on fractals. "Encoding appears to require a lot of new mathematics," says Francis Sullivan of the National Institute of Standards and Technology.

Before an image can be expressed with fractals, it must be organized into an equilibrium state. Investigators must find and apply transformation rules to realize this state. The woman's portrait, for instance, involved on the order of 1,000 transformations, Barnsley says. A description of these transformations is all that is needed to reconstruct the image; since the rules can be expressed with a few parameters, it is possible to achieve high compression ratios.

Many mathematicians doubt that all pictures can be so represented. "We don't have good definitions and theorems of what is efficiently encodeable," says Mehrdad M. Shahshahani, a mathematician at the Jet Propulsion Laboratory. The technique is efficient only if it does not require an enormous number of transformations.

"There are always going to be questions until there's a product out there," Sloan replies. The company plans to ship its first encoding devices in the summer, he says. In March, Iterated Systems will have ready the other half of the system: the decoders. (The decoders, which plug into a personal computer, will decompress fractal images stored at better than 64:1 ratios and display the picture in gray scales at video speeds of 30 frames per second, Barnsley says.)

Until encoders are commercially available, users can work with compressed images provided by Iterated Systems. Customers will be able to

***PORTRAIT OF WOMAN** consisting of some three million bytes of data was compressed into 5,800 bytes with a technique developed by Iterated Systems in Atlanta, Ga. Original photograph by Noriaki Yokosuka/Misako Tanaka.*

send still pictures or videos to the company, which will encode the images and send back code via electronic mail, Sloan says. Among those interested in the technology are telemarketers and real-estate brokers who want to display their wares by means of a personal computer.

"We're at the beginning of a process," Sloan says, in building both the technology and the business. —*E.C.*

One if by Land

Technology-transfer doyen sights a foreign shore

Penicillin earned fame for the two British scientists who developed it in the 1920's and 1930's, but not a shilling besides. Their government gave away, free of charge, the details of the drug made from mold. As World War II raged, the U.S. Department of Agriculture ran with the discovery, developing a fermentation process to manufacture penicillin in large quantities. British firms found themselves paying royalties on the process to their American counterparts. Galled by the experience, the British government created a watchdog agency in 1949 to protect ideas spawned at British universities.

Four decades later that obscure branch of the U.K. civil service, now called the British Technology Group (BTG), is gearing up to play a role in the markets emerging as both Western and Eastern Europe remove old economic and political barriers. The BTG also hopes to represent American inventions in the byzantine international market for intellectual property.

The BTG says it is the largest organization in the world devoted exclusively to technology transfer—the moving of technology from the laboratory to the marketplace. The BTG's staff of about 60 technical specialists tracks the problems that industry is trying to solve and then forages through British academic laboratories for potentially useful ideas. "Technology transfer is a very personal process, not a data-base activity," says Ian A. Harvey, the BTG's chief executive. "You have to have people who know the market and know the technology and can be very creative about linking them."

The BTG's first big hits came in the 1950's, when the agency secured worldwide patents on a second generation of antibiotics called cephalosporins, developed at the University of Oxford, and on the air-cushion suspension system for the Hovercraft. Many efforts have yielded returns only after years of patient funding. The BTG financed pyrethrin, an environmentally benign pesticide, for 20 years before it became a best-seller in 1982. Similarly, the agency spent a decade nurturing research on magnetic-resonance imaging (MRI) before the medical diagnostic tool began to return significant royalties in 1984. Currently, royalties on MRI and pyrethrin patents are 50 percent of the BTG's revenues.

Technological fizzles have naturally outnumbered successes in the inherently risky business of investing in technology. But the BTG made other missteps along the way, observers say. By the late 1970's "it was very much a civil-service organization"—lethargic, difficult to negotiate with and indifferent to the concerns of the market, recalls David Leuellin, a patent lawyer at McKenna & Company in London.

In the early 1980's the British government began putting pressure on the BTG to improve its efficiency. The Thatcher administration considered selling it to the public. It also freed British universities and national laboratories to work with other agencies to patent and market their technology, thus ending the BTG's comfortable monopoly on inventions supported by British taxpayers.

Hardship has galvanized the BTG. Harvey is determined to show that the BTG, with 1,500 patents and annual royalties of about £24 million, is better equipped to enforce patents than smaller organizations.

To prove the point, Harvey, shortly after arriving at the BTG in 1985, put top priority on fighting five outstanding infringement suits—including one against the U.S. firm Johnson & Johnson for violating a licensing agreement for the BTG's patents on MRI. The BTG represented 20 MRI patents from scientists in three British universities in the suit. Johnson & Johnson agreed to settle the case in 1986. The patents generated some £3 million in royalties for the inventors last year. A battle with the U.S. Pentagon over alleged infringement of the BTG's Hovercraft patent continues. Harvey hopes that by demonstrating that the BTG is willing to take on large opponents, it will win a reputation as a tenacious and productive litigator.

The BTG has reason to be particularly concerned about its reputation now; its mainstay patents for MRI and pyrethrin will expire within the next few years. The BTG is eager to bolster its portfolio with U.S. technologies. It won its first major U.S. client in March, 1989, when Johnson & Johnson decided to leave the diagnostic-imaging business and assigned the BTG the task of managing more than 100 imaging-related patents the company held in Europe and the Far East.

The BTG nonetheless faces significant obstacles in its bid to find and represent U.S. technology, says Alan Walton, a venture capitalist in Connecticut. "Universities are increasingly becoming allied with local funding sources," he argues, "and most big companies would rather pay fees to a small company that's brought the technology along." It remains to be seen whether the British agency will make a mark on the American commercial landscape. —*Fred Guterl, London*

PROFILE

Fischer Black

Calculated risks enable mathematician to turn a profit

It looks like a scene from the film *Wall Street:* on the 29th floor of the financial house Goldman Sachs, video terminals blink with data on stocks. Traders frantically shout buy-and-sell orders. On the periphery of this financial melee for the past six years, appropriately enough, has been the office of Professor Fischer Black.

Although Black, a soft-spoken former faculty member at the Massachusetts Institute of Technology, is an oasis of calm deliberation in contrast to the turmoil of the trading floor, he can take much credit for Wall Street's frenetic pace. Less than 20 years ago he and Myron S. Scholes revolutionized the financial world with a mathematical model that described how the price of stock options could be determined by such variables as stock prices and interest rates. (Stock options allow holders to buy or sell a stock at a fixed price at some future date.) "The importance [of the Black-Scholes model] cannot be overstated," says Richard L. Sandor, a director at Drexel Burnham Lambert. "Black-Scholes set a pattern not only for options trading but for trading generally."

Now a partner at Goldman Sachs, Black is still brimming with surprising, often unorthodox ideas. Among his most recent ones is a model that describes the ideal portfolio of international stocks and currencies for an investor. This model has yet to win much acceptance among economists, but Goldman seems willing to give it the benefit of the doubt; the firm re-

cently promoted Black to its asset-management division.

The trek to Wall Street has been a gradual one for Black, now age 52. As a student at Harvard University, he switched majors several times before earning a bachelor's degree in physics in 1959 and a doctorate in applied mathematics five years later. He soon began work as a consultant for Arthur D. Little in Cambridge, Mass., where he met a colleague who had devised a model for pricing securities and other assets. The formula sparked Black's interest in finance.

That model described stocks, so Black turned to options, which were not widely traded at the time. In the late 1960's valuing options was an imprecise art; it required estimating the future worth of the security, discounted by some value—usually an estimate of "risk" and interest rates.

Black began making some simplifying assumptions: he abandoned the former way of estimating the discount rate and instead substituted a rate that depended on the underlying stock price, a more predictable value than was risk. Working with Scholes, who was at M.I.T., Black continued to look for ways to clarify the relations among the variables. They concluded that the theoretical present value of an option was based on a set of parameters that could be estimated: the current stock price, stock volatility (based in great part on past stock-price changes), the date and price at which the option would expire and the prevailing interest rate.

"It was like a technique for manufacturing options," says Robert Merton who is now at Harvard. Plug values into the formula, and it churns out a theoretical value of the security, he explains. Black, Scholes and Merton quickly validated the approach by trying it out in the market. Convincing economists took more time. Two academic journals rejected the paper. The third suggested that the options model was "too narrow," Scholes says, and recommended that the authors discuss other applications of the theory.

The publication of the model in 1973 in the *Journal of Political Economy* roughly coincided with the opening of the Chicago Board of Options Trading. Eager for a way to grapple with the new opportunities, traders latched onto the Black-Scholes model.

The model is not a "rule" for making money, Black emphasizes. (The existence of such a rule—however desirable—would violate economic principles, notably the equilibrium of the market.) Some investors depend on the model to hedge, or protect, their positions in other securities by finding options that will offset changes in stock prices, for instance. Traders exploit the model to take advantage of minute, temporary changes in stock prices that slip past other investors.

FISCHER BLACK helped to invent a formula for pricing options in the early 1970's that now is programmed into traders' computers. Photograph by Jim Marchese.

Black "is nearly divine on options pricing," says Rudiger W. Dornbusch of M.I.T. "That's where one should read him." Black likes to think his work in economics may have more significance than his financial models. He argues that the Federal Reserve's efforts to tweak the money supply have little effect on national output or growth. Instead he believes the economy is primarily driven by business cycles—essentially the waxing and waning of supply and demand in the private sector. "Monetary policy is and must be passive in an economy of well-developed financial markets," he writes in his book *Business Cycles and Equilibrium*.

Such views are far from conventional. Unlike his work in finance, Black's business-cycle theories "are more of a point of view," says Robert Lucas of the University of Chicago. "There are no formulas you can crank out."

Recently, however, a few economists have been taking a hard look at some of the empirical evidence—and finding support for Black's ideas. For instance, Edward C. Prescott, who advises the Minneapolis Federal Reserve Bank, estimates that as much as 70 percent of business-cycle fluctuations are related to economic forces other than the money supply.

Black's formula on the ideal international portfolio also runs counter to accepted doctrine on currency hedging. Many economists think investors make little profit in the long run by hedging the value of their international stocks against changes in exchange rates. (They should instead simply aim to protect the value of their stocks.)

But after spending a week puzzling over differential equations describing currency risks and returns, Black concluded that investors might indeed profit from currency hedging. Furthermore, since the international market gives investors in different countries the same opportunities for risk, they could in theory hedge the same percent of their portfolios, he says.

Theorists have yet to embrace these ideas. But given Black's track record, they might hedge their position. —*E.C.*

One-Step Steel

How steelmakers hope to break the coke habit

By the end of the 1990's steelmakers hope to have developed an innovative steelmaking technique that will make polluting, expensive-to-replace coke ovens unnecessary and so help cut production costs.

The technique, called direct steelmaking, could give North American steelmakers a competitive edge, asserts Ian F. Hughes, a vice president at Inland Steel Industries. Industry observers point out that direct steelmaking—if successful—could also tip the industry's balance of power toward small steel plants.

The first major test of direct steel-

COKE-OVEN BATTERY (in background) burns coal in the absence of oxygen to produce coke. When the vertical panel doors are removed, coke is pushed from the battery into railcars. Scrubbing equipment (in front of battery) helps to control pollution. Photograph courtesy of the American Iron and Steel Institute.

making will take place near Pittsburgh in mid-March. At that time a pilot plant sponsored by the U.S. Department of Energy and the American Iron and Steel Institute (AISI) will begin producing molten steel from iron and untreated coal. The plant, which will eventually produce 120 tons of molten steel a day, is at the heart of a three-year, $30-million direct-steelmaking project initiated in late 1988 by the DOE and AISI. Industry is contributing about $7 million to the project; the DOE is paying the rest.

Direct steelmaking transforms the conventional three-step steelmaking method into a single process. To make steel currently, the large, or integrated, mills begin by heating coal in the absence of oxygen to produce coke, a high-carbon residue. Next coke, iron ore and limestone are blasted with very hot air, which results in liquid iron. This iron is then treated with oxygen and other compounds to produce purified, liquid steel.

But turning coal into coke generates significant carcinogenic emissions. Steel-industry leaders say that existing coke ovens will fall short of the tough emissions standards proposed by environmentalists for the 1990's. Buying new, cleaner ovens is not an attractive option, because these can cost $250 million each. Consequently, "large manufacturers are up against a wall" and looking for alternatives, says Robert W. Crandall, an economist at the Brookings Institution.

Direct steelmaking streamlines production by eliminating the need for coke. Manufacturers start by mixing untreated coal, iron ore and scrap metal in a bath of molten iron and then injecting the mixture with oxygen. The coal serves both as a reducing agent and as a fuel to generate heat. At the same time, the metal is treated and purified to create a "semisteel." This molten metal is then poured into a ladle and refined further into semifinished products, such as slabs that are later rolled into sheets.

That the North American steel industry is only now organizing to fund development of alternative production techniques is "typical of U.S. industry in general in the last two decades," reflects John F. Elliott, metallurgist at the Massachusetts Institute of Technology. Other nations have been pressing ahead with direct-iron-making projects. Japan's Ministry of International Trade and Industry has organized a seven-year, $100-million project; Austria, Australia, South Africa, West Germany and the Soviet Union have similar direct-iron-making projects under way. AISI researchers say their project should leapfrog these efforts by producing semisteel that requires little subsequent processing, rather than crude iron.

Researchers still have a long way to go before they can make the jump to direct steelmaking. Workers at M.I.T. and Carnegie-Mellon University are testing ways to keep smelter temperatures high by harnessing and burning gases from the coal-iron reaction, even as the Pittsburgh facility tests the in-bath smelting process. If, by late 1991, the work in Pittsburgh succeeds, the industry will consider spending $100 million or more to build a demonstration plant incorporating all aspects of direct steelmaking, says Egil Aukrust, who directs the pilot project. The earliest that commercial steel plants could hope to utilize the techniques would be in the late 1990's, he adds.

Who will benefit most from the technology? Crandall predicts that minimills, rather than integrated manufacturers, will be best able to exploit direct steelmaking. These small outfits are less encumbered by bureaucracy and rigid labor rules and so can more readily adopt new methods, he says. Although minimills have prospered by offering relatively low-cost, low-grade steel made from scrap metal, direct steelmaking could make them even more competitive, Elliott says. Unlike conventional processes, direct steelmaking will let the minimills economically produce a much purer metal on a small scale and so enter the lucrative market for thin sheets of high-quality steel, now controlled by the integrated manufacturers. *—Gregory T. Pope*

THE ANALYTICAL ECONOMIST

Changing money in the East bloc

According to some investment advisers, the liberalization of Eastern Europe is the greatest business opportunity since the heady days of the Marshall Plan: economies in shambles and ready to be rebuilt, literate work forces ready to fill the factories so that they can qualify as consumers. Millions of forints, zlotys and koruny can be made in the process. The only problem is that no one really wants forints, zlotys or koruny.

Convertible, or "hard," currencies, such as dollars or yen, can be exchanged for each other freely and so strongly link the economies of their respective nations. Unconvertible currencies (those of the East bloc and of a number of developing nations) can be exchanged only by government permission; they insulate their local economies from the rest of the world.

Unconvertible currencies also discourage foreign investment. If a Western company spends deutsche marks, with which it can buy anything, to

build a plant in Romania, it receives profits in lei, which it can spend only in Romania. Unless the company gets permission to convert the lei back into marks, its money may be as good as lost. Few companies accept such terms, and so governments that have unconvertible currencies set artificially high exchange rates to reap as much hard currency as possible from those who must do business there.

But now East bloc countries are trying to woo foreign investment; they need the money to rebuild their economies and to give their citizens a standard of living comparable to that of the West. The first thing they should do, some economists say, is to make their currencies convertible. A chief exponent of this view is Harvard wunderkind Jeffrey D. Sachs, who earlier wrought improvements in the economy of Bolivia and has now taken on restructuring Poland's.

Other economists contend that making East bloc currencies convertible now is premature and perhaps even dangerous. (Poland may be the exception, says Richard E. Quandt of Princeton University, because its economy is already in such decline.) Before making their currencies convertible, says Edward A. Hewett of the Brookings Institution in Washington, D.C., East bloc nations must negotiate the passage from a cloistered economic system in which money does not matter to the world economy in which it most certainly does. Money in the East has acted as a passive bookkeeping medium for keeping track of transactions, not as a medium of exchange whereby the prices of goods are set at their economic value.

The Soviet Union, for example, explains Ronald I. McKinnon of Stanford University, has at least three kinds of "currency," none freely convertible into another. First, individuals can save paper money and coins issued by the government, but businesses cannot. Next, the surpluses that businesses generate go into accounts with the state bank; different surpluses are earmarked for different purposes, and spending any surplus may require state approval. Last, the state bank lends enterprises "soft" money so that they can buy materials to meet their production quotas. When the government allocates all the goods by decree, the passive system works well, but in a free market its effects are "perverse," McKinnon says. Money-losing enterprises can always borrow more, and profitable ones cannot invest what they earn to increase production.

Adequate taxation is another crucial step. Liberalization of socialist economies, McKinnon explains, has uniformly led to huge government deficits and economic chaos because "the reformers didn't do their homework."

When the state owns nearly everything, the money from "profitable" sectors—those that make things people buy—goes to subsidize "unprofitable" ones, such as national defense. An explicit tax system is unnecessary, because the government can simply set wages and prices at levels that let it skim what it needs. In the absence of explicit taxes, the transition to private enterprise means a precipitous drop in revenues. In the Soviet Union, government revenues declined from 47 percent of the gross national product in 1985 to 38 percent in 1989; the budget deficit climbed by almost precisely the same amount, from 2.7 percent of the GNP to 13.1 percent.

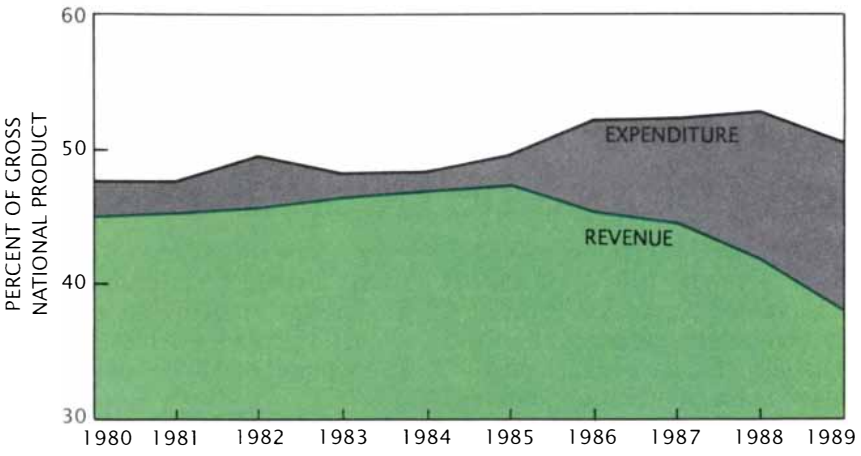
Without appropriate monetary and tax systems, McKinnon says, making East bloc currencies freely convertible could make those countries' economic picture even bleaker than it is now. In a free market the value of Eastern currencies would plunge, raising the price of imported goods and potentially triggering hyperinflation. Foreign investors might find it so cheap to buy local assets that they would squeeze out local entrepreneurs.

Yet that leaves the East bloc countries in a dilemma. If making the forint or the koruna convertible would swamp the local economy, where will the money come from for needed investment? Some economists, such as McKinnon and Janos Horvath (who served briefly as chairman of Hungary's National Economic Reconstruction Council in 1956), suggest internal sources: enterprises should be cut off from all credit and forced to finance expansion out of retained earnings. If this draconian solution works as it did in West Germany in the late 1940's, companies will raise prices, and consumers will finance the investment by drawing down savings that would otherwise have remained idle. Once the economy is thriving, according to this view, there will be time to talk about convertible currencies.

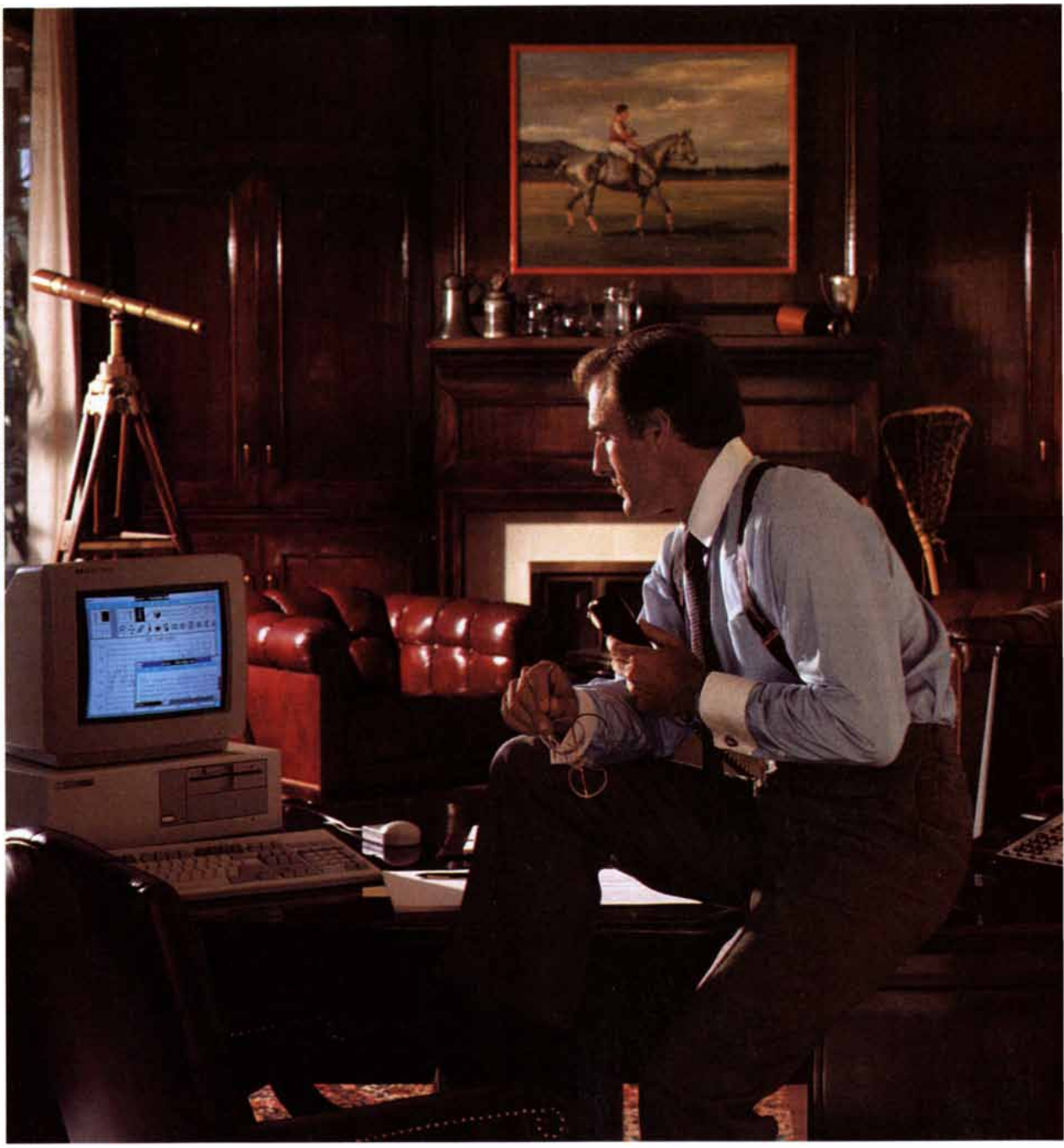
Then again, foreign investment may be a necessary part of rebuilding the East bloc, as it was in Western Europe under the Marshall Plan. If so, Hewett says, measures short of full convertibility will still let Westerners see their profits in hard currency. Some companies have practiced barter for years; probably the best example is the lucrative trade of soft drinks for vodka. Another alternative might be hard-currency auctions, he says. Enterprises could convert local money to yen or marks regularly but not on the fluid and volatile basis that characterizes currency trading in the West.

With such alternatives available, a focus on fully converting currencies may be misleading, Hewett argues. There are other barriers to investment; issues of property rights, for example, have yet to be settled. Once such problems are resolved, Eastern Europe may be at least as attractive to foreign investors as any other developing area; its roads and other infrastructure are better than most, its workers are better educated and its effective wages will be just as low.

—Paul Wallich and Elizabeth Corcoran



***SOVIET BUDGET DEFICIT** has grown sharply since the inception of* **perestroika.** *According to Ronald I. McKinnon of Stanford University, the plunge in revenues* **(green)** *is a natural consequence of liberalization: the "profits" of previously state-controlled enterprises no longer flow to Moscow. Source: Planecon.*



Tom bought Hewlett-Packard long before he bought PCs.

When he got his first job on Wall Street, Tom's HP calculator was his most valued possession. It was the combination of innovation and reliability that gave him an edge. So when his Information Systems Manager recommended they go with Hewlett-Packard personal computers, Tom needed no convincing.

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. 142G.**

There is a better way.

HEWLETT PACKARD

The Road to the Global Village

by Karen Wright

...after more than a century of electric technology, we have extended our central nervous system itself in a global embrace, abolishing both space and time as far as our planet is concerned.

A quarter of a century has passed since an obscure scholar at the University of Toronto penned those words. In a treatise called *Understanding Media,* Marshall McLuhan prophesied the worldwide coalescence of human awareness into a single community that he would ultimately style the "global village." Who today would quarrel with McLuhan's prophecy? The world watched in real time as the peoples of Eastern Europe tossed out decades of history. An earthquake in San Francisco shook millions of baseball fans. A sprawling computer network joins users in 35 countries. Trading on the Tokyo stock exchange can ruin a New York broker's breakfast.

And these notes may be only the overture to an even more powerful suite of changes. The developed world is experiencing a transforming convergence of computing and communications technology whose impact, experts say, will rival that of the replacement of muscle power by machines. The hybridization is already apparent in a roster of implements such as facsimile machines, automatic bank tellers, VCR's, answering machines, compact discs, cellular phones, video games and high-definition TV.

Industrial and academic laboratories are incubating technology that will make the artifacts used by today's villagers look decidedly primitive. To find out what's in store, I visited the workers who are building the road to the global village. They took me beyond the touch screen and the mouse, through the looking glass of the computer terminal to the technological wonderland of tomorrow.

Trade à la Modem

Like most revolutions, this one has its roots in economics. The romance between computing and communications technology is both driving and being driven by a burgeoning services sector. Nearly 75 percent of U.S. workers are employed in services, up from 55 percent in 1948. And those services have, in the words of one analyst, "a voracious appetite for information technology."

"Service providers now own about 85 percent of the total U.S. stock of information-technology items," says Stephen S. Roach, a principal and senior analyst at Morgan Stanley. Computing and communications equipment are smack in the middle of Roach's definition. Whereas manufacturing industries spend 10 percent of their capital on information technology, he says, the services sector's allotment doubled in the past decade to almost 20 percent.

It might not be surprising that services gobble up information technology; many of them would never have come into being without it. Yet some experts, Roach included, maintain that the notion of the service economy masks a more profound transition from an economy based on material resources to an economy in which most wealth is generated by an immaterial essence: information.

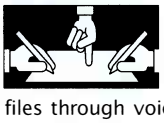
"You can break services into two very different parts: personal services and information services," says Marc Porat of Apple Computer, whose 1976 doctoral dissertation at Stanford University crystallized the notion of the "information economy." The percentage of the U.S. work force employed in personal services (such as hair styling) has remained constant at between 15 and 20 percent since the turn of the century, Porat says. During the same period the percentage of information workers—bankers, insurance agents, lawyers, science journalists—has gone from a trickle to a flood. Information work currently accounts for 55 percent of employment in the U.S., according to Roach.

The vendors of such technology have taken Porat's analysis to heart. "A few years ago we asked ourselves what business we're in," says Robert W. Lucky, executive director of AT&T Bell Laboratories. Eventually, Lucky says, AT&T decided it was in the "IM&M business: information movement and management. See, we're not communications anymore. We're information movement and management." He chants the phrase like a mantra.

Lucky's boss, Arno A. Penzias, vice president of research at Bell Labs, calls it a movement "from things to thinking." Not everyone is convinced. "Be careful using the phrase 'information

GLOBAL-VILLAGE PHRASEBOOK

COOPERATIVE WORK: Technology that allows people in remote places to interact with each other and with the same documents and files through voice, data and video links; also known as computer-supported cooperative work (CSCW).

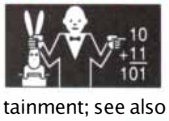


FLOPS: Stands for "floating-point operations per second," a measure of how fast a computer can perform calculations involving numbers that are not integers. Such calculations are critical for applications that describe the physical world.

HUMAN BEING: An analog processing and storage device with a bandwidth of about 50 bits per second. Human beings excel at pattern recognition but are notoriously slow at sequential calculations.

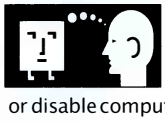
HYPERMEDIA: Media whose contents are arranged in a nonlinear fashion or stratified to represent varying levels of detail. Compare with "multimedia."

INFOTAINMENT: Computing and communications technology designed to inform or educate while entertaining. Sometimes called edutainment; see also "multimedia."

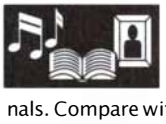


ISDN: An abbreviation for "integrated-services digital network," a communications pipeline that can carry both voice and data signals. Basic ISDN standards allow for two voice channels and one 16,000- bit-per-second data channel.

KNOWBOT: A software servant that acts, without requiring explicit instruction, to prioritize or scout for information, protect confidential data or disable computer "viruses."



MULTIMEDIA: Computer-based presentations combining two or more media, such as text, graphics, writing and video and audio signals. Compare with "hypermedia."



VIDEOTEX: Any of several on-line information and retail services targeted at the home consumer and mediated by personal computers or dedicated terminals.

VIRTUAL REALITY: A surrogate or metaphysical environment created by communications and computing systems. "Virtual" often denotes the computer-generated counterpart of a physical object: a "virtual room," a "virtual circuit" and so on.

economy,'" says James Brian Quinn of Dartmouth College. "Information is interesting only insofar as it affects other things." Quinn stands by "service economy" as the more meaningful rubric because, he says, information and information technology represent just one subset of a host of factors that contribute to the productivity of the services sector [see "Technology in Services," by James B. Quinn, Jordan J. Baruch and Penny C. Paquette; SCIENTIFIC AMERICAN, December, 1987].

A Bolt from the Blue

Recent events suggest that—information economy or no—people are not prepared to swallow information technology whole. Some attempts to weave computer and communications technology into the fabric of daily life have fallen on very hard times. The videotex consumer-information services, for example, have what Lucky describes as a "sad and tortured" history: Knight-Ridder sank $50 million into its Viewtron project before folding it in 1986, while the IBM-Sears joint venture Prodigy has gathered only 200,000 subscribers. Scholastic, Inc., lost millions in its first bout with educational software in the early 1980's. No one has cashed in on electronic banking either; years after its introduction, fewer than 100,000 people nationwide bank by computer.

Observers say such misfirings betray an insensitivity within the computer and communications industries. Mitch Kapor, founder of Lotus Development Corporation and chairman of ON Technology, Inc., calls it the "absence of a design sensibility. The problem of making computers useful to people as communications and information devices is not an engineering problem," Kapor expounds, "it's a design problem. Engineers are trained to eliminate the subjective factors. But it's exactly the subjective factors that are critical here."

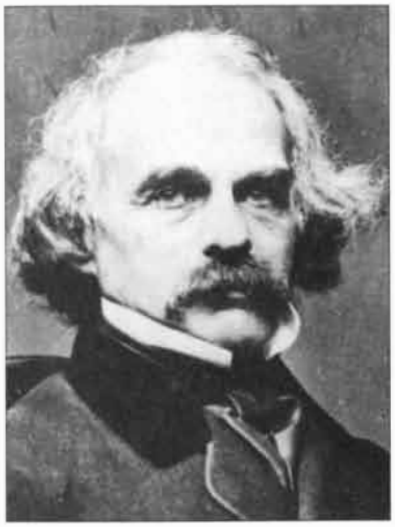
Madison Avenue adman Arthur Einstein, who presided over the creation of IBM's memorable "little tramp" advertisements, says he has defined one of those factors: "Terror. You have to confront the documentation," he says. "You have to learn a whole new language. Did you ever use the word 'interface' before you started using a computer?"

Computer companies have spent many years and many millions enhancing user-friendliness with windowed graphics interfaces (pardon the expression) and other contraptions, enabling a whole population of computer-naive

They saw it coming. . .

1851, Nathaniel Hawthorne
The House of the Seven Gables

"Is it a fact . . . that, by means of electricity, the world of matter has become a great nerve, vibrating thousands of miles in a breathless point of time? Rather, the round globe is a vast head, a brain, instinct with intelligence! Or, shall we say, it is itself a thought, nothing but thought, and no longer the substance which we deemed it!"



1925, Aldous Huxley
Those Barren Leaves

"It's comforting to think," said Chelifer, "that modern civilisation is doing its best to re-establish the tribal régime, but on an enormous,

national and even international scale. Cheap printing, wireless telephones, trains, motor cars, gramophones and all the rest are making it possible to consolidate tribes, not of a few thousands, but of millions. . . . In a few generations it may be that the whole planet will be covered by one vast American-speaking tribe, composed of innumerable individuals, all thinking and acting in exactly the same way, like the characters in a novel by Sinclair Lewis. . . ."

Mr. Cardan nodded and puffed at his cigar. "That's certainly a possibility," he said. "A probability almost; for I don't see that it's in the least likely that we shall be able to breed a race of beings, at any rate within the next few thousand years, sufficiently intelligent to be able to form a stable non-tribal society. . . ."

"Perhaps a slightly lower standard would be necessary," suggested Chelifer.

Marshall McLuhan coined the term "global village" in the 1960's, but futurists and visionaries were toying with the concept of an interconnected, worldwide society long before then. McLuhan told his friend and collaborator Bruce R. Powers that the phrase came to him while he was contemplating Hawthorne's passage. When McLuhan published *Understanding Media* in 1964, his view of such a world was sanguine, but as the years passed he grew leery of the communications "implosion." In *The Global Village,* which Powers shepherded through production after McLuhan's death in 1980, the authors devote nine pages to the "satisfactions" of the "new tribalism"—and 38 pages to its "dissatisfactions."—*K.W.*

To get ahead light conversation.

Something astonishing is happening on Wall Street.

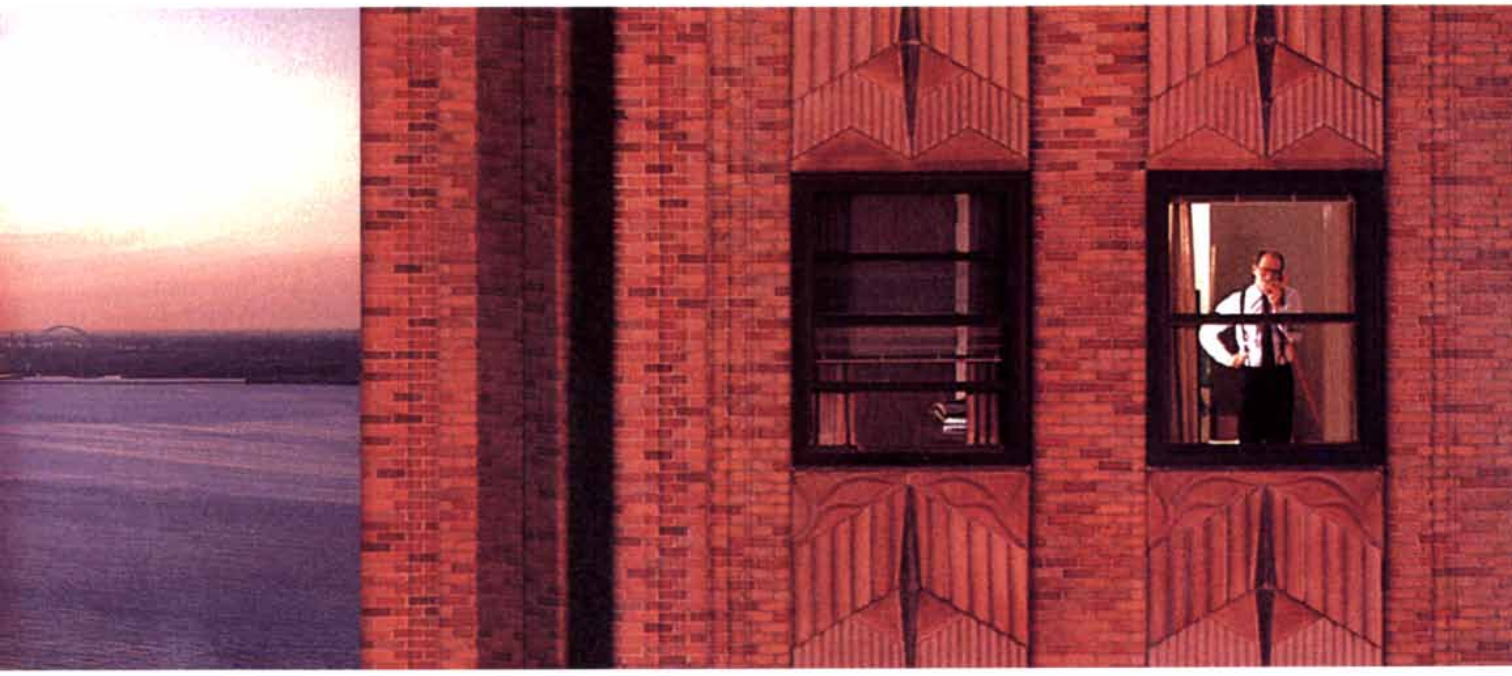
Twenty-six fierce competitors are joining together on a private communications network from NYNEX.®

Top firms like Kidder, Peabody, Cowen & Company, and Nomura Securities.

The new fiberoptic highway links the firms directly to their trading partners throughout New York's financial community. Carrying voice and data by light waves.

And it's going to save the members of the Securities Industry Association about $15 million over the next five years. Which even to these giants is no small change.

on Wall Street, try



The NYNEX family of companies would like to work with you, too. We offer you everything from computers, networks, and software to the services of New England Telephone and New York Telephone, the NYNEX company that created this innovative network. Call us at 1 800 535-1535.

When we say the answer is NYNEX, it's not just a lot of talk.

Need to communicate? Need to compute? The answer is

NYNEX®

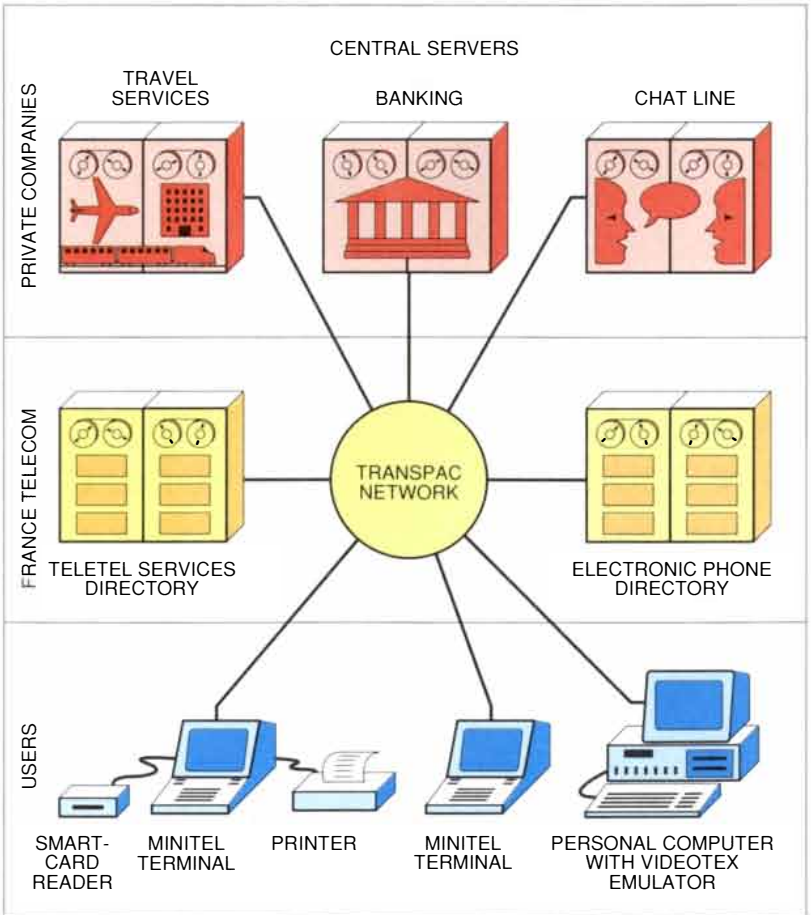
Vive le terminal

The French government-owned telecommunications monopoly has put in place what most experts agree is the only videotex network in the world that is a truly popular success. The Teletel network provides a gateway to 12,000 electronic information services in 18 percent of French households, a claim no other national network can make. Through Teletel's "Minitel" terminals, which France Telecom distributes free of charge, five million telephone users can manage their finances, schedule airline and hotel reservations, order items from computerized catalogues and chat with each other on the *messagerie* (a service that has nurtured a *messagerie rose* of hired escorts and erotic messages).

"Nothing in Europe can compare with the scale of Teletel," says Graham Thomas of the science policy unit at the University of Sussex. Great Britain and West Germany, which started their own videotex programs at the same time France did, have each garnered fewer than 200,000 subscribers. The younger Prodigy videotex system in the U.S., run jointly by IBM and Sears, has about the same number, and Japan's 10-year-old Captain product has fared even worse: fewer than 100,000 terminals have been sold.

Why has France succeeded where others have foundered? Perhaps because it is so easy to become a Teletel initiate. Potential subscribers need only show up at a France Telecom store and request a terminal; they forfeit nothing more than their right to receive a thick telephone directory once a year. For 37 centimes (about six cents) a minute, France Telecom provides a directory-assistance data base that gets the reluctant user's feet wet while saving the utility the expense of printing and delivering telephone books.

To attract business and higher-income home users to the network, France Telecom is promoting new products and services, such as a smart-card reader that makes Teletel less susceptible to forgery. Yet the utility's efforts to lure corporate subscribers could falter. There are no international standards for Teletel, and with the unification of the European common market looming, Teletel runs a risk of getting caught in a technological cul-de-sac.—*Fred Guterl, London*



individuals to partake of the machines at work and play. Yet computers still daunt educated, adult consumers, and even sophisticates can find them maddeningly slow. So estimable sponsors of research such as IBM, Apple, Xerox and Bell Communications Research (Bellcore, for short) are enlisting anthropologists, sociologists and psychologists in the effort to build information technology users can relate to.

Lucy Suchman, staff anthropologist of Xerox's Palo Alto Research Center (better known as PARC), watches workers in an airline-operations room at San Jose International Airport to learn how they extract particular information from a chaotic assortment of radio, telephone, computer, text and video feeds. In her 10 years of research she has identifed two ingredients of technology that can win people's enthusiasm: familiarity and flexibility.

Suchman points out that all technology requires learning; driving an automobile, for example, is not an intuitive act. But because people are familiar with that technology, she says, "people's experience tells them that the payoff is going to be worth the effort. With computer technology, the payoff is unknown." Often, Suchman says, the technology is too inflexible to provide any payoff at all. "Users need to be able to accommodate a design to their specialized requirements," she observes. Suchman thinks the design process would improve if engineers and programmers observed how their wares were used after the devices left the lab. As it is now, she says, the technology just "appears—like a bolt from the blue."

The Whole-Person Paradigm

PARC has a history of creative thinking; the personal computer, laser printer and local-area network were born there, and it counts among its consultants a former Benedictine monk. A growing concern for human behavior is evident in a variety of projects in other labs as well. Researchers industry-wide are building interfaces that exploit natural forms of human expression: speech, gesture and handwriting, to name a few. Workers at Apple call this approach the "whole-person paradigm."

"My dream is to open up computer power to the majority of the people in the world," says psychologist Richard Bolt, who runs the interface group at the Massachusetts Institute of Technology's rarefied Media Lab. "Most people don't know how to type or how to log in, but most people do know

how to talk and how to point." Bolt directed the creation of the Media Room, a kind of digital environment that recognizes speech and gesture. "We're still trying to live up to promises George Lucas and Walt Disney made a long time ago," gripes one denizen of the Lab. Then he grins. "But that's okay."

Nicholas Negroponte founded the Media Lab in 1982 on the conceit that computing, publishing and broadcasting will someday be sublated in a grand "computer presence." Negroponte has so far made believers of 60 corporate supporters, who together donate $13 million a year to the cause.

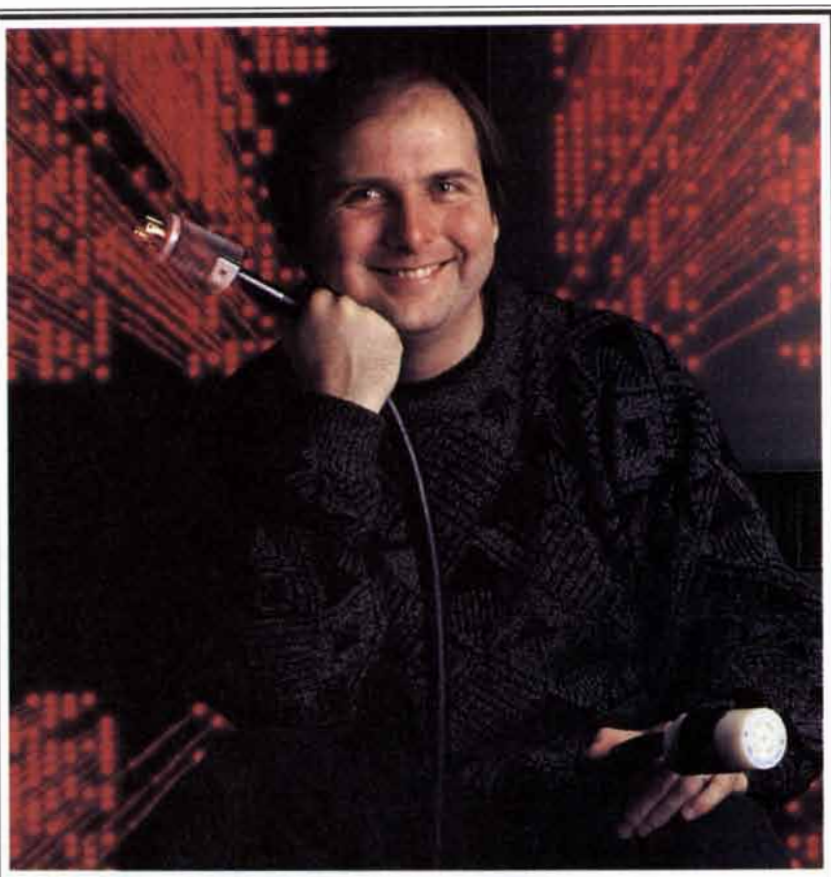
In recent years major suppliers of hardware, software, consumer electronics and communications equipment have also launched or beefed up their own programs investigating speech processing, synthesis and recognition as well as handwriting and gesture recognition. Speech synthesizers already answer telephones and tell some car owners when their oil is running low; Canon makes a $500 stylus and digitizing screen that lets you write on a computer display. Later this year Sony plans to market a handwriting-recognition system that can interpret the standard set of 1,800 Chinese characters the Japanese use.

The Room You Can Talk To

Computers themselves are already becoming more verbally adept. Machines with vocabularies of several hundred words can recognize isolated words just about as well as human beings can. Getting a computer to understand what those words mean, of course, is the trick. Consider that every word in the salutation of Hamlet's famous soliloquy is a homonym: "two bee oar knot too bee."

For 10 years Bell Labs has been teaching a computer how to take airline reservations over the phone. Progress has been impressive. The Airline Information System has a 132-word vocabulary and understands more than 12 billion sentences. AIS learns voices by listening to one trainee say 200 sentences (the trainee has to say each of the 132 words). Like most speech-recognition systems, it relies on parallel processing to attain practical speeds: it breaks a sentence into pieces it can analyze separately and simultaneously.

The AIS computer consists of 24 32-bit digital signal processors, each of which can perform eight million floating-point operations per second (industry calls them FLOPS). Don't try this at home—you'd need the equiv-



Thinking Machines' thoughtful man

Danny Hillis keeps a running list of the things he hopes to accomplish during his lifetime. "Write a mystery novel. Teach kindergarten. Build a starship." He pauses, suddenly self-conscious. "You know . . . the kinds of things *everybody* thinks about." A computer scientist by trade, Hillis has made a smashing success of his life's work while preserving, at age 33, the guileless curiosity of someone who doesn't know what he wants to be when he grows up.

Hillis is the brains behind the Thinking Machines Corporation, the Cambridge, Mass., company he helped to found in 1983 to sell his iconoclastic supercomputer, the Connection Machine. One of the fastest computers in existence, a typical Connection Machine can do several billion floating-point operations per second (FLOPS) and costs between $2 and $4 million. Thinking Machines has sold about 50 machines since the company built the first prototype in 1985; sales have been growing by 50 percent a year. Last year they totaled $45 million.

Not a bad return on inspiration. Hillis lit on a new philosophy of computer architecture as a graduate student at the Massachusetts Institute of Technology, when he became interested in designing a computer that could simulate commonsense reasoning. Instead of using one souped-up processing unit over and over again, as conventional computers do, such a computer could solve problems by marshalling the strength of many smaller processors operating in parallel, or simultaneously. A Connection Machine has 64,000 so-called parallel processors [see "The Connection Machine," by W. Daniel Hillis; SCIENTIFIC AMERICAN, June, 1987].

About 10 percent of all the supercomputers worldwide are Connection Machines, and many manufactures are incorporating elements of parallelism in their standard designs. Experts say that parallel processing will become increasingly important as computers tackle the kinds of chores human beings do best, such as image and speech recognition. Hillis thinks that by 1995 half of all supercomputing will be done on massively parallel machines. Arrogance? His associates call it "boyish enthusiasm." Playfulness seems to be one way Hillis realizes his ambitions. "There was a time when everybody thought we were crazy. So we tried to appear as normal as possible," he reflects. "Maybe it's time to act crazy again."—*K.W.*

alent of 2,000 personal computers.

Bell Labs' David Roe gave me a demonstration. He tells AIS the destination and the day and time he wants to leave, the computer makes a note of what it thinks he said and then comes up with an appropriate spoken response. Personally, I like to get vegetarian meals when I fly, and so I ask Roe to order one. He hesitates; apparently "vegetarian" isn't one of those 132 words.

"I would like a vegetarian meal," Roe says stiffly into the receiver.

I would like flight zero eight to Seattle, types AIS.

"It seems to like eights," Roe apologizes. Well, we're all entitled to our preferences.

AIS doesn't know when it doesn't know something. As a research project it is a laudable achievement, but Roe points out that in the real world it wouldn't get away with offering vegetarians flights to Seattle. It would have to be able to tell you it didn't understand what you said.

Down the hall from AIS there is a conference room called HuMaNet (for *hu*man-*ma*chine *net*work) that knows enough to ignore you when you're not talking to it. That's important, because the system controls the lighting and the audiovisual materials as well as the phone lines connected to the room. Its various components answer to "facility control," "data-base control" and "ISDN control." (ISDN stands for integrated-services digital network, a communications pipeline that can carry both voice and data signals.)

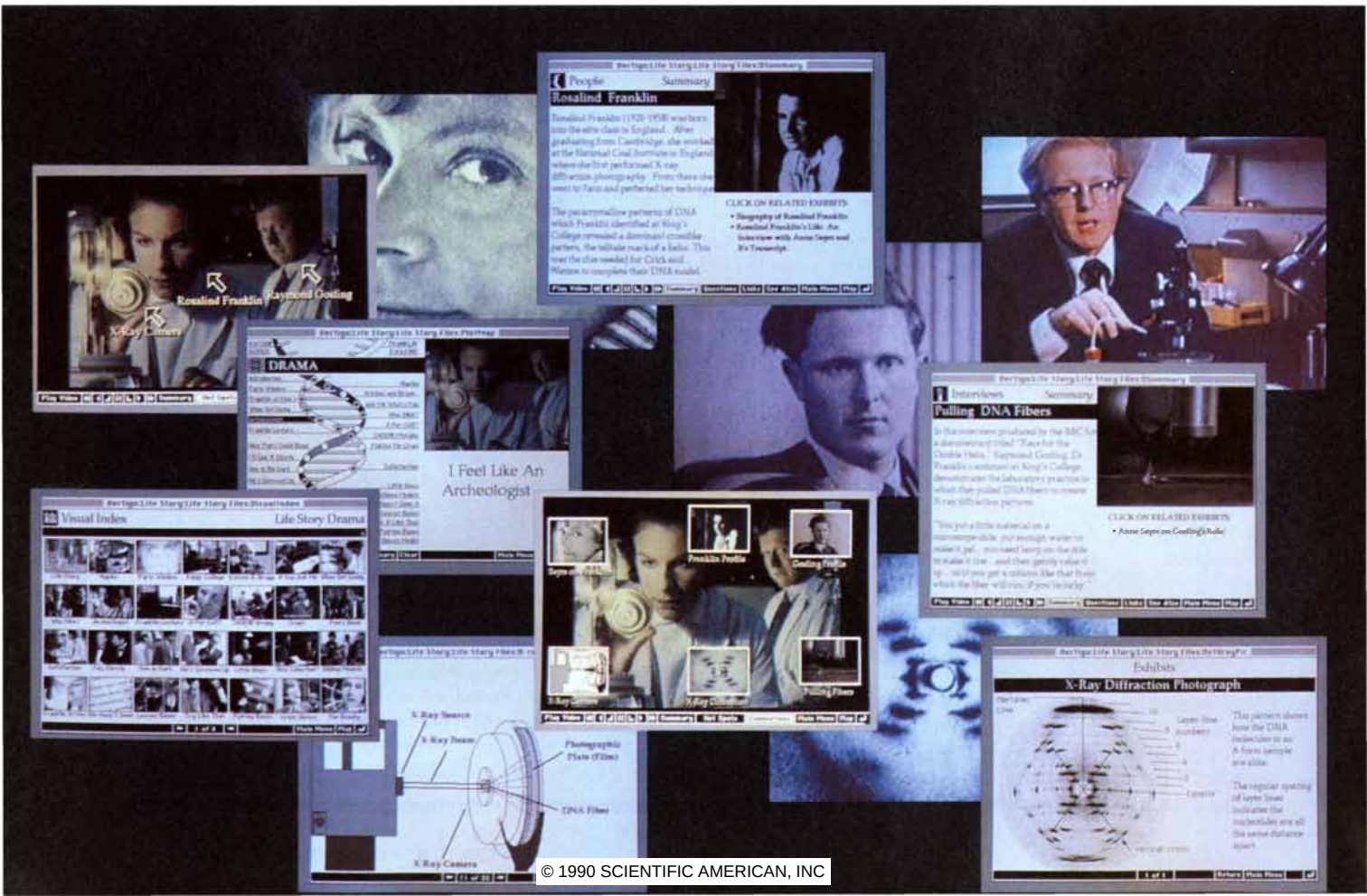
David Berkley is HuMaNet's trainer. "Facility control, lights down," says Berkley, and down they go. The room's "speech seeking" microphones are arranged unobtrusively above the projection screens; the couches are comfortable. "Data-base control, viewgraph one," Berkley says, and a picture of a tree frog appears on one of the video screens. "Data-base control, next image"; another frog. "ISDN control, call room two," and HuMaNet dials. The room tunes out as Berkley explains how it works.

HuMaNet has an 80-word vocabulary that includes the word "please" because, Berkley says, people can't seem to kick the habit of being polite. The system's components are PC-based and integrated by a Sun workstation. It rides on a 144,000-bit-per-second ISDN line. That, by the way, is the going rate for first-generation ISDN service, which is already installed in more than 100 companies in the U.S.

The whole-person paradigm is also behind industry's fascination with multimedia—a combination of text, graphics, audio and video orchestrated by computers. Companies such as Apple, IBM, Intel, AT&T and Bellcore, as well as entrepreneurs such as special-effects expert Robert Abel and smaller institutions such as the Institute for Research on Learning in Palo Alto, are experimenting with interfaces that accommodate varying degrees of user interaction. The Information Workstation Group, a consulting firm in Alexandria, Va., estimates that multimedia will have a worldwide market of $16.6 billion by 1994.

Adventures in Multimedia

Multimedia programs are usually billed as "infotainment": education in the guise of a video game. One of the more sophisticated prototypes is Life Story, configured by the Smithsonian Institution, Lucasfilm Ltd., Adrian Malone Productions and the five or six people who make up Apple's Multimedia Lab in San Francisco. "Life Story" is actually the name of a 1986 BBC drama about the discovery of the structure of DNA. The multimedia prototype uses the drama as a springboard



for investigating the people and practices involved in the discovery.

One of the chief architects of the program, a quiet zealot named Fabrice Florin, showed me how to freeze a scene and pick it apart. I quickly got to a sound bite in which biographer Anne Sayer discusses Rosalind Franklin's upbringing; then I got a video clip of the real James Watson talking about how Franklin was portrayed in the scene I had watched; then I segued to an animated explanation of the procedure the scientists had used to define DNA structure; and finally I landed in an interview with Linus Pauling, who grudgingly acknowledged that Watson and Francis Crick's discovery might be the greatest of 20th-century biology (the two beat Pauling to what would have been his third Nobel prize).

I had a great time, learned a lot and forgot to ask Florin penetrating questions. Then something on one of the "cards" reminded me of something else I had read a couple of cards back. I wanted to look at it again, but I had no idea where that card was in relation to me. Panic seized me. I was lost.

"I want to go back," I said to Florin.

"Back where?" he asked. How to describe it? Forget the card.

"Back to the beginning!" I yelped. But wait—this is hypermedia. There *is* no beginning. The place at which I began is not necessarily The Place at Which One Must Begin. I'm adrift in a sea of information, with no point of reference. Florin is nonplussed. Help!

Maybe I'm a prisoner of linear Westernized left-brain prejudices, but I need something like a trail of bread crumbs to keep me oriented. Tom Landauer, steward of Bellcore's "SuperBook," told me later that many people get lost in such programs. SuperBook is a hypertext medium, meaning it has information stratified into varying levels of detail. It puts a path tree on the screen indicating the user's stratum: a "you are here" for the hypertext voyager. I was comforted.

Multimedia has also been incorporated in computer-supported cooperative work (CSCW), the aim of which is to allow people in remote locations to interact with each other and with the same documents and files in a virtual space that could span feet or continents. Some of the projects, such as Bell Labs' Rapport and the Slate prototype crafted by BBN Systems and Technologies Corporation, rely on terminals. Others, such as Bellcore's Video Window and PARC's Colab, use screen projection as well [*see illustrations on next page*]. The generic sobriquet for CSCW software is "groupware."

MULTIMEDIA PROTOTYPE Life Story lets a user explore at will different aspects of the discovery of the structure of DNA. The prototype, which takes its name from a 1986 BBC drama, incorporates sound bites, video clips, computer-generated diagrams and animation, as well as scenes from the drama itself. Life Story is the product of a collaboration among Apple's Multimedia Lab, the Smithsonian Institution, Lucasfilm Ltd. and Adrian Malone Productions.

Let's Get Virtual

As far as engaging the *whole* person goes, the technology known as virtual reality takes the trophy. Pioneered by the National Aeronautics and Space Administration and VPL Research, Inc., virtual reality envelops one in a surrogate existence. The user dons a helmet that presents each eye with built-in graphics screens; a glove or an entire suit fitted with position sensors and laced with fiber-optic threads tells the computer how the user is moving. You can walk around in a "virtual space," or "fly" with the point of a finger, or grab and rearrange objects in the space. Two people can enter the same reality and interact with each other, cloaked in virtual disguises.

Like AT&T's talking machines and Apple's multimedia show-and-tell, virtual reality seeks legitimacy in practical applications. Someday it could aid "visualization for complex-system monitoring," "architecture, city planning and product-design previewing" or "simulation of three-dimensional assembly," says VPL's promotional material. Right now it's just plain fun.

"Careful," says Jaron Lanier, founder and CEO of VPL Research, as he fastens the helmet of the virtual world on my head. "Many a science writer has never come out of here." "Here" being, in this case, a virtual reconstruction of the Mad Hatter's tea party from *Alice in Wonderland*. I'm Alice.

"Do you mind being Alice?" Lanier asks. "Because you can be the Mad Hatter if you want. Sometimes," he adds, "men get mad when they find out they're Alice." I can hear Lanier, but all I can see is a colorful vista of geomorphic shapes, mountains on the horizon, a picnic table with the fixings for tea, the Mad Hatter hovering at a safe distance and, hovering closer, a hand, which I realize abruptly is mine.

"If you point your index finger, you can fly," says Michael A. Teitel, VPL's senior optical engineer. But first I want to look around. I turn my head slowly to the right and the left; slowly, because Lanier's system is operating at about a quarter of its capacity on this particular day, and if I turn my head quickly, the graphics interface that is supplying my world won't be able to keep up. Then I'd get dizzy and fall down, which turns out to be part of the fun, too.

I point my finger and zoom to the foot of the mountain range; I turn around, slowly, to see whence I came. It's a long way away and I feel lonely. Back to the party.

"You can pick things up, too," says Teitel. He is monitoring my reality on a standard graphics terminal. I make a few unsuccessful grabs for a teacup as he and Lanier shout encouragement. When I finally get the hang of using my new hand, I'm possessed by the urge to throw things. This must be what it feels like to be a baby.

The fun has its price. VPL sells a one-user version of its RB2 ("reality built for two") system for $225,000, including the Silicon Graphics IRIS computers that control the input to each eye. If you want to pull out all the stops and get the RB2 for two people, VPL will knock $20,000 off the second set for a total of $430,000.

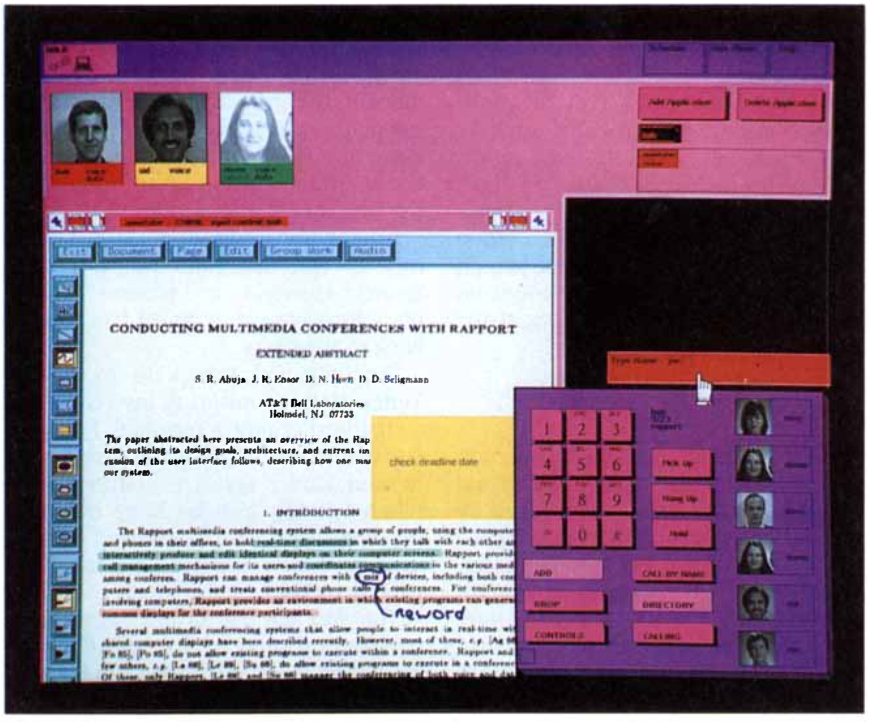
Later in the session Lanier, Teitel and I build a custom reality that includes gnats shaped like Stealth bombers. The gnats are supposed to buzz around my head, never straying more than a few feet, in the time-honored gnat tradition. But when I get into the new reality, I can't see any gnats. Teitel tells me all four of them are there, but that for some reason they are lurking behind my head, just beyond my peripheral vision. I make several vain attempts to turn around fast and catch the insects off guard. I get dizzy. I fall down.

Teitel succumbs to the obvious pun. "Must be a bug in the program."

The Party Line

Projects such as HuMaNet, Life Story, SuperBook and virtual reality are valiant attempts by companies large and small to dissolve the resistance that keeps today's computer from enjoying the comfortable ubiquity of the television and the radio. Almost all such prototypes demand extraordinary amounts of data capacity. Many also presuppose the existence of digital pipelines capable of handling it.

The digital transmission of television-quality full-motion video and audio signals requires 45 million bits per second. Even a still TV image hogs 24 million bits of memory in a computer file. Phone lines carry voice signals at



COOPERATIVE WORK projects such as Bell Labs' Rapport let people in remote locations work together on the same computer files. The pictures at the top of the screen indicate voice and video links (the "absent" Doree has temporarily left her terminal). Users can highlight and edit text with an electronic stylus (*colors at center*) and attach memos to the body of a document (*yellow square*). A telephone keypad is on the right.

TELECONFERENCING WINDOW at Bell Communications Research links groups of people with full-motion video and "hands free" microphones. Conferees can also project transparencies onto one half of the window for joint scrutinization.

the rate of only 64,000 bits per second, and data transmission is even slower: a standard modem sends data over phone lines at 2,400 bits per second. It would take almost three hours to transmit a single, soundless TV still by modem.

Fiber-optic cables have been reported to carry as many as one *trillion* bits of information per second. In the past decade the rates achieved by optical fibers increased 100 times, while the cost of fibers fell from $3 to 15 cents per meter, according to Kessler Marketing Intelligence. AT&T and other carriers in the U.S., as well as the post, telegraph and telephone ministries in other countries, are laying down millions of miles of fiber every year. "The capacity's out there," says Bob Lucky. "The only question is can we agree on how to use it."

That question has prompted much political and economic maneuvering. "I can't think of anything that requires architectural vision with interworking standards more than international networking," says Irwin Dorros, executive vice president for technical services at Bellcore. Dorros says the community of suppliers, users and regulators has to arrive at some kind of consensus about where the fiber is going to go, who will pay to install it, who will own the resulting network and what services it will supply. All those questions, in turn, bear on the technical standards of the network: bit rates, switching mechanisms, software protocols and so on. "Until those issues are resolved," says Solomon Buchsbaum, vice president of customer relations at AT&T, "your global village is a mirage."

Here's where the chumminess between the computer and communications industries breaks down. The phone companies aren't used to transporting data, and the computer companies aren't used to the idea of universal service. "There's a real cultural problem between us and the phone company," says David Clark of M.I.T.'s laboratory of computer science.

Clark presided over the growth of the computer-network leviathan Internet. Internet arose "higgledy-piggledy," as he puts it, from the four-node network instituted in 1969 by the U.S. Department of Defense. Today it has become the intangible eighth wonder of the world, connecting 936 networks, at least 175,000 computers and countless users in 35 countries. People talk about it as if it were alive. Clark describes its growth as "anarchic democracy at its best."

Many experts think a public version of something like Internet will be vital

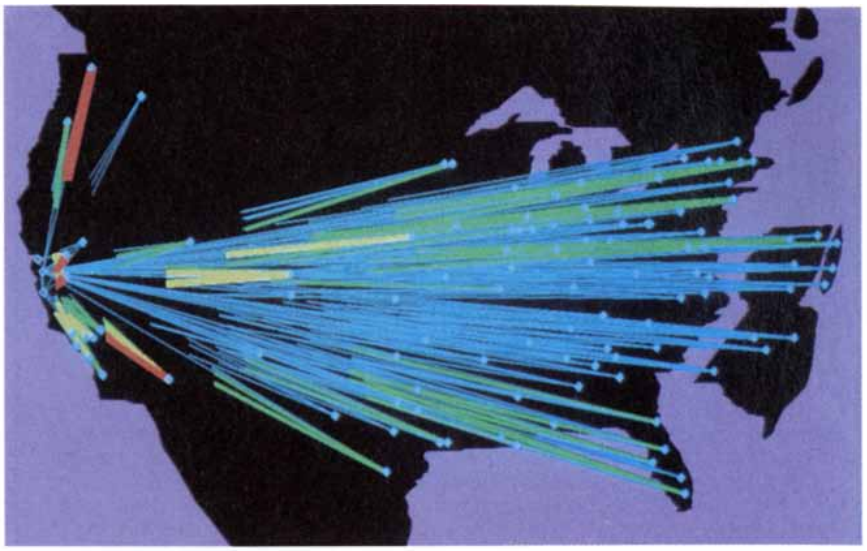
to international competitiveness. "It could make a substantial difference in our economy," says Michael L. Dertouzos, who heads M.I.T.'s laboratory of computer science and chaired the university's commission on industrial productivity [see "Toward a New Industrial America," by Suzanne Berger, Michael L. Dertouzos, Richard K. Lester, Robert M. Solow and Lester C. Thurow; SCIENTIFIC AMERICAN, June, 1989]. In the commission's report Dertouzos advocates the creation of a national information infrastructure. But, he emphasizes, unlike Internet, such a coordinated infrastructure will not come about by itself.

The administration's Office of Science and Technology Policy has in fact thrown its weight behind a high-performance computing bill, introduced by Senator Albert Gore of Tennessee last May, that would set up a three-billion-bit-per-second National Research and Education Computer Network. NREN would link the nation's centers of higher education with facilities in government and industry. The bill also includes provisions for the creation of a directory of users and resources, a feature natives of computer networks say is desperately needed.

While Congress ponders Gore's initiative, the International Telegraph and Telephone Consultative Committee has already hammered out the international standards for the unified voice and data channel, ISDN. Initially, ISDN lines would provide two 64,000-bit-per-second voice channels and one 16,000-bit-per-second data channel; these rates are the upper limit for standard copper cable. In Europe and Japan, government-owned monopolies have grabbed the standards and run. But in the competitive context of the U.S. market, the standards do not provide the economic incentive for building public data "superhighways" that would bring high-speed transmission to every individual.

"The business is not in global networking but in private networks," Lucky explains. "Every big corporation has its own network, because that's where the densest traffic flows. About 95 percent of all calls are made within those small islands." To establish a universal network, on the other hand, "you'd have to set up a network for millions and millions of people to route 5 percent of the calls."

ON Technology's Kapor faults applications developers for failing to provide a conceptual impetus that would pull capital investment and political consensus toward more universal service. "You can't expect Washington to get excited about things in the abstract," he says. "Nobody's thinking about the content side of this."



TELEPHONE TRAFFIC JAM following the earthquake in the San Francisco Bay Area illustrates the crises that can beset network-management systems. The lines show phone calls that were not completed 90 minutes after the earthquake on October 17, 1989. The width and color of the lines indicate the extent of blocking; blue represents the lowest number of uncompleted calls, red the highest. Crowding also occurs during call-in contests or when tickets to Rolling Stones concerts go on sale.

Knowbots

Of course, some people *are* thinking about the content side of the computing and communications revolution. There is little consensus, save for one point: that the salient issue of the information age will be the need to exclude information. The average global villager will be on the receiving end of a torrent of media traffic. And McLuhan himself warned that "the more information one has to evaluate, the less one knows."

"The major problem is that our brain can absorb only 50 bits per second. Technology won't change that," says AT&T's Buchsbaum. The figure refers to the amount of information a person can register, which is modest compared with the amount the sensory organs harvest. "You're not going to re-engineer the human brain," Buchsbaum says. "We are analog beasts." So what is the solution? One possibility would be to tune out. Alvy Ray Smith of Pixar, a San Rafael, Calif., graphics firm, predicts the blossoming of a counterculture that would simply exempt itself from mainstream society. Most of us, however, are already wired.

One could pay other people to stem the information tide. As Nick Negroponte of the Media Lab points out, that is what consultants, journalists and secretaries already do. In fact, Negroponte maintains, there is no reason such exclusion needs to be performed by a living, breathing human being with a meager bandwidth. Artificial-intelligence experts at Xerox PARC, IBM, AT&T, M.I.T. and other industry and academic labs are designing programs that can ferret out things you'd like to know and banish things you aren't interested in. How does a computer program know what you like? It learns your preferences by watching what you do, "just like a secretary does," one researcher told me.

Such programs are called computational agents, or "knowbots." Thomas Malone of M.I.T.'s Sloan School has an embryonic knowbot program that prioritizes his electronic mail according to parameters such as whether or not it's from his boss. Eventually knowbots could provide systemic defense against computer "viruses" and protect sensitive or confidential information from intruders. They could also act as personal emissaries, scouts or staunch defenders of one's privacy.

"Machines talking to machines—on your behalf" is how Negroponte describes it. Negroponte says he would be willing to pay $100 a year to have the information he needed from each issue of *Scientific American* condensed to a summary less than one page long. "Computers could do that. In the future it'll be machines at both ends. I'm not saying that we'll be obsolete—they'll be talking on our behalf. But it won't be you, or me."

Kapor isn't willing to go that far. "I

Japan says *hai* to fiber optics

If the plans disclosed this past October by Nippon Telegraph and Telephone (NTT) are realized, the Japanese will by 2015 be knit together in a fiber-optic telecommunications system that runs into virtually every home and every office. The "integrated services digital network" (ISDN) would bring the capacity for features such as TV phones and customized electronic newspapers to the company's 50 million customers and link the archipelago by submarine cable and satellite to similar networks in the West. NTT says the service will have 7,000 subscribers in more than 200 cities by its two-year anniversary in April.

NTT's network, like the ISDN available in other countries, typically operates at 64,000 bits per second and has a maximum transmission rate of 1.5 million bits per second. The lower narrow-band capacity exceeds by a factor of 50 the normal rate of data transmission over phone lines, which is about 1,200 bits per second by modem. Yet it still cannot carry full-motion video, high-definition TV or traffic between local-area networks. Consequently, NTT is preparing for broadband ISDN services that could accommodate transmission speeds of up to 150 megabits per second. Iwao Toda, a veteran NTT executive who heads the company's research and development headquarters, told SCIENTIFIC AMERICAN that NTT's labs are investigating another leap to 10 billion bits per second.

Recent government proposals to break up NTT have chilled some ISDN enthusiasts, but from NTT's point of view the only serious questions are when and how the company will make the huge investments—estimated at between 25 and 30 trillion yen (from $174 to $208 billion)—required to equip the archipelago with a contiguous network. "Certainly they can do it," says Nobuyuki Nakahara, president of Tonen, an oil company that has diversified into *hikari fiba* ("optical fiber") manufacture. Indeed, with 280,000 employees and more capital than any other company in the world (11,559 billion yen, or $80.3 billion, in assets), NTT should be able to rise to the occasion.

Not everyone in Japan is brimming with enthusiasm for the electronic future. A hundred years ago sword-bearing samurai who regarded the country's new telephone lines as a threat passed beneath them holding white fans over their heads to ward off impurity. To this day extremists occasionally sabotage NTT cables. But NTT is already looking ahead to the time when domestic-network proprietors will cooperate on a global scale. In summarizing his thoughts, Toda plucks a line from a favorite musical, *The Man of La Mancha:* "To surrender one's dreams is madness."—*Henry Scott Stokes, Tokyo*

don't think you can automate totally, nor do I think you want to. People will always do better than strictly mechanical methods." For the time being we'll have to. Knowbots are still the stuff of fantasy—or nightmare, depending on your point of view.

The Village Revisited

As it stands now, the convergence of computing and communications technology is characterized by a wealth of possibility and a dearth of direction. "The delivery of information will change the way people, businesses and countries operate," Buchsbaum says. The question is how. "Clearly, power and wealth will be redistributed," says Peter Schwartz of Global Business Network, a consulting firm that specializes in future scenarios. "But not necessarily in predictable ways."

Much of the uncertainty in the information marketplace stems from the fact that no one has figured out how much information, and the shuffling thereof, is worth. "It's hard to know how you price information in any realistic sense," observes sociologist Daniel Bell. "Information is not consumed and destroyed the way you consume and destroy food."

Bob Lucky calls information "economic ether." The obvious example is software piracy. You can give software away over and over again, and you still have it. This loaves-and-fishes quality of information has no place among the parables of capitalism.

"Our whole economic system gets its axioms from the notion of property," Marc Porat says. Whereas property is by nature scarce, information has no inherent scarcity, he asserts; consequently, "neo-classical economics doesn't work." The information content of a slice of pizza (advertising, legal expenses and so on) accounts for a larger percentage of its cost than the edible content does, according to Henry Kelly and Andrew W. Wyckoff of the Congressional Office of Technology Assessment.

While economists and managers struggle to redefine value, the information economy is erasing another classic landmark: the national border. "The absolute imperative of the information age is the need to establish a global economy," says Penzias of Bell Labs. "If you go to a party and some one says, 'What's your globalization strategy?' and you say, 'What's that?' you become invisible."

The global scale of economic activity does not necessarily spell homogeneity for the consumer; in fact, it bodes just the opposite. "Just talk to the advertising people," Dan Bell says. "The key word is 'niches.' " The Paris-based ThinkNet Commission says that vast networking will enable "mass customization" by imparting the ability to gauge customers' preferences individually. Shoppers from Seattle to Singapore to Stuttgart could have cars, clothing and curtains made to their specifications because suppliers could precisely ascertain demand.

What is liberating in one respect might be restrictive in another. In his 1983 book *Technologies of Freedom,* the late Ithiel de Sola Pool of M.I.T. warned that the gradual digitization of all forms of expression could provide an avenue of invidious attack on First Amendment rights. Pool pointed out that radio and television broadcasters have already accepted the dilution of those rights; if the government chooses to regard all digitized information in the same category, he warned, it could attempt to exert control over any material that becomes available through electronic venues.

In any event, publishers, film companies and broadcasters will have to find new ways to cope with a distinctly different environment than the one that exists today. Consumers themselves will be infringing on the territory of those industries. With high-speed networks and multimedia interfaces in place, Kapor and other observers suggest, writers could disseminate their work without publishers, composers and musicians could have an audience without signing a recording contract and anyone with a video camera could produce and distribute a sitcom.

It is too early to tell which, if any, of these speculations the future will validate. In a 1987 article in *Daedalus,* Bell likened the "world society" wrought by information technology to "a set of giant Calder mobiles, shifting in uneasy balances.... The exact configurations are difficult to capture."

In other words, stay tuned.

Make it easy on yourself.

Easy access to AT&T.

Most public phones automatically give you AT&T. But sometimes you'll find a phone where the location owner has chosen another long distance carrier. Even then it's still easy to reach AT&T in most cases. Simply dial our easy access code, **10 + ATT + 0**, before dialing the area code and number you're calling.*

Calls more places.

Whether you're calling across the country or across the world, the *AT&T Card* makes it easy. AT&T connects you to over 250 countries and areas worldwide—more than any other long distance card.

Most reliable.

You can count on the *AT&T Worldwide Intelligent Network* to get your call where it needs to go. So you have less to worry about. And when calling from overseas, you can rely on this network to access *AT&T USADirect® Service* —the only service that connects you to an operator back in the States from over 60 countries. That's worldwide service no one even comes close to.

So make it easy on yourself.

Use the *AT&T Card*, the only calling card you'll ever need—and it's free. To order an additional card or a new one, **call toll-free 1 800 222-0300, Ext. 500.** You'll also receive the latest information on how best to use your *AT&T Card*.

Fast connections.

With the AT&T network you get fast connections for all your long distance calls. And since all local telephone companies honor the *AT&T Card*, it's convenient for all your calls.

Best service. Best value.

This card delivers the quality service you've come to expect from AT&T. With more operators, consistent low prices and innovative calling plans, AT&T gives you the best value.

AT&T
The right choice.

*For international calls, please dial 10 + ATT + 01 + the country code + the local number.

Suspension-Feeding Vertebrates

Animals that filter their food out of the water can reap the abundance of plankton and grow in huge numbers or to enormous size

by S. Laurie Sanderson and Richard Wassersug

Tadpoles, herring, basking sharks, flamingos, mallards and blue whales form an unlikely family. The group ranges from the smallest free-living vertebrates to the largest and includes amphibians, fishes, birds and mammals. These organisms are all suspension feeders: they eat by processing massive volumes of water through their feeding apparatus and filtering out small organisms and other fragments of organic material.

Vertebrate suspension feeders subsist on minute plants and animals too small to be sensed and captured individually. Typically they engulf many food particles during each feeding bout, and they are unselective about what they eat. For example, a basking shark 10 meters long can ingest more than 540 liters of zooplankton a day—millions of tiny animals, each no more than one millimeter long. At least 1,850 cubic meters of water pass through the shark's mouth in an hour. Smaller suspension-feeding fishes such as menhaden and bighead carp routinely consume phytoplankton (microscopic plants) as small as 10 microns in diameter. Some tadpoles can remove particles the size of bacteria from water; to subsist on such a diet, they may need to process a volume of water equal to their own body volume every few minutes.

Suspension-feeding fishes are important in both ecological and economic terms because they occupy a low position in the trophic pyramid (the food chain). Plants—both phytoplankton and larger forms—make up the base of the pyramid; herbivores form the second level, carnivores that feed on them the third and so on. Each successive level of the pyramid accounts for less biomass than the level below it. Only about 10 percent of the food eaten by individual predators is incorporated into their biomass; the rest goes to basic metabolism and to work such as pursuing and catching prey. Most suspension feeders are either herbivores or carnivores that consume herbivorous zooplankton, and so their populations can account for an extraordinary amount of biomass. Suspension-feeding species such as anchovies, herring and carp account for more than a third of the annual catch of ocean and freshwater fish. The vast food resources available to suspension feeders can also support enormous individual organisms; the largest animal that has ever lived, the blue whale, subsists on krill, a herbivorous zooplankton.

Suspension feeding is not only one of the most widespread feeding methods but also among the oldest. R. Glenn Northcutt of the Scripps Institution of Oceanography and Carl Gans of the University of Michigan have hypothesized that the earliest vertebrates were suspension feeders. The larvae of some ancestral vertebrates may have subsisted on plankton (as tadpoles do today) and then metamorphosed into actively swimming carnivores. In this scenario, characteristics such as special sense organs and efficient patterns of aquatic locomotion evolved as animals switched from suspension feeding to the active capture of larger prey.

Although modern vertebrate predators that bite or engulf individual prey most likely evolved from suspension-feeding ancestors, this does not mean that modern suspension feeders are relics somehow overlooked by evolution. The diversity of suspension feeders suggests that the practice has reevolved many times because it is particularly suited for capturing the large quantities of food available at low concentrations in many different aquatic environments. In addition to the complex filtering structures that have evolved to trap food particles, suspension feeders have developed specialized bodies, necks, heads and jaws to implement their particular eating habits. The convergent evolution of suspension feeding in many separate lineages offers an exciting opportunity to analyze the particular structures that correspond to different modes of suspension feeding.

We recently undertook a survey of the feeding mechanisms of all known suspension-feeding vertebrates with the goal of recognizing correlations between structural features and methods of feeding. In particular, we concentrated on identifying common features shared by suspension feeders in evolutionarily divergent lineages; such similarities may provide clues to the physical mechanisms involved in suspension feeding.

The most important distinction we make among suspension feeders is one between ram feeders, which rely on forward motion to deliver water into the mouth, and suction feeders, which suck in water while remaining essentially stationary. This distinction cuts across taxonomic lines; its focus is the differing ways that organisms transport water to their filtering surfaces. Although suspension feeding usually calls to mind cranial adaptations such as oversize jaws, locomotive specialization is equally important. An animal that delivers water to

S. LAURIE SANDERSON and RICHARD WASSERSUG met 12 years ago when Wassersug taught an introductory biology course in which Sanderson enrolled as an undergraduate. Sanderson is now a National Science Foundation postdoctoral research fellow in environmental biology at the University of California, Davis, where she studies the hydrodynamics and food intake of suspension-feeding fishes. In 1988 she spent two weeks living 14 meters underwater in the U.S. National Oceanic and Atmospheric Administration's Caribbean research habitat. Wassersug is professor of anatomy at Dalhousie University in Halifax, Novia Scotia. He investigates the functional morphology of tadpoles with a view to elucidating their evolution.

its feeding structures by forward motion has no need for a complex muscular oral pump, and so its cranial specializations for suspension feeding—other than the proliferation of the filters themselves—may be relatively minor.

Both the ram feeders and the suction feeders are further divided into continuous feeders, which pass water through the mouth continuously, and intermittent feeders, which essentially take one gulp at a time and extract the food from it before refilling the mouth. Continuous suction feeders remain still while feeding. This requires radical cranial adaptations—not only a pumping mechanism but also valves to ensure that water flows in only one direction through the filters. Not surprisingly, the vertebrates best designed for sustained aquatic locomotion are ram feeders, whereas those that do not swim continuously, such as birds, tadpoles and lamprey larvae, have developed suction-feeding mechanisms.

The continuous ram feeders constitute the largest category of vertebrate suspension feeders: more than 20 different species of fish, including the paddlefish, manta ray, basking shark, whale shark, megamouth shark and certain anchovies, herrings, sardines and mackerels. Fishes in this category swim forward with the mouth completely open, allowing water to escape from gill slits at the back of the head through finely branched filter structures. They have been described as tow-net filter feeders because their method of feeding appears similar to the workings of the plankton nets towed by oceanographers.

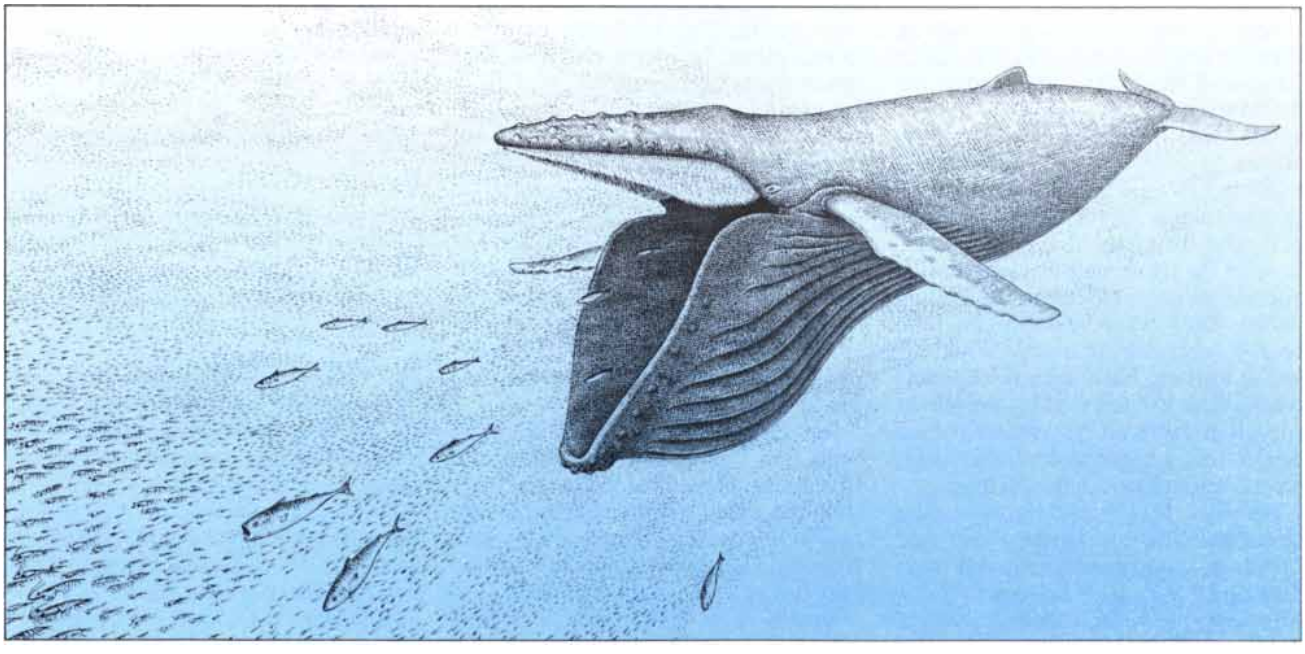
Intermittent ram feeders use their forward velocity to force their mouth open and rapidly fill the oral cavity in a single pulse. Only some of the largest whales, the rorqual whales, are massive enough to practice intermittent ram feeding. For this strategy to work, inertial forces must be very high—that is, the water containing the prey must remain in place while the whale moves forward and its mouth opens around both water and prey.

Ram feeders are all roughly similar in body design, even though they may vary in size by orders of magnitude. Suction feeders are a much more diverse group. The category includes lamprey larvae (ammocoetes), tadpoles, dabbling ducks, flamingos and certain bony fishes. Ammocoetes, tadpoles, ducks and flamingos all have an oscillating pump in the throat and the floor of the mouth that generates feeding currents while the animals themselves remain stationary.

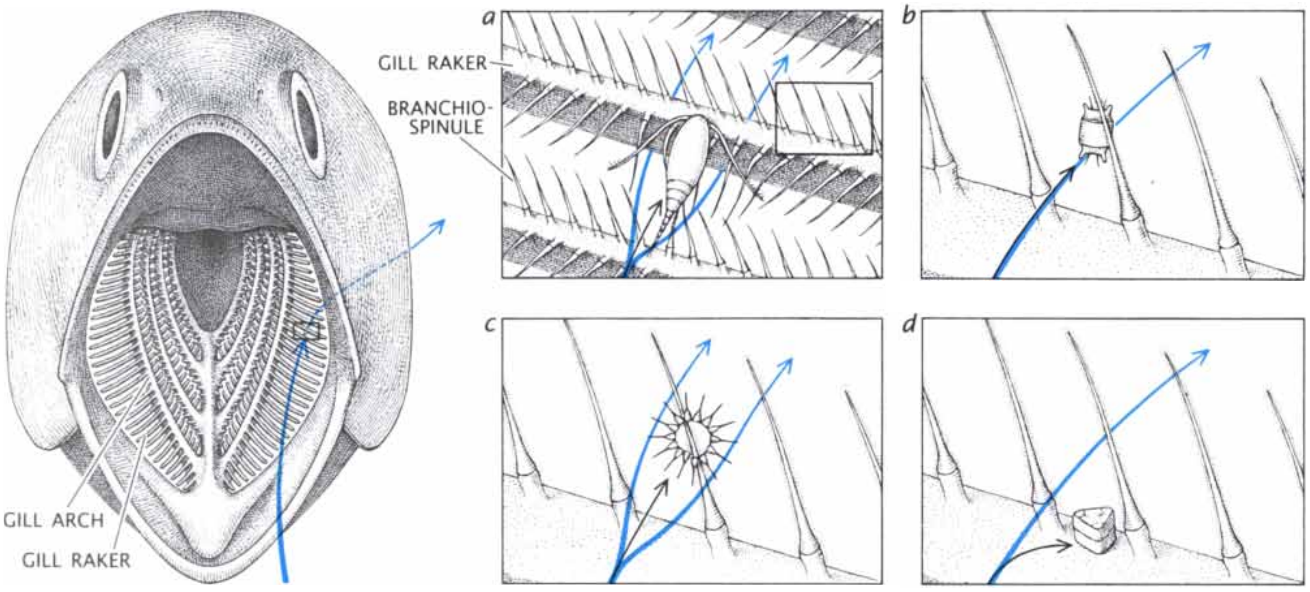
The feeding systems of ammocoetes and tadpoles are tightly linked to their breathing apparatus: their internal entrapment surfaces are part of the gill system. For suspension-feeding birds, in contrast, the entrapment surfaces are anatomically independent of respiratory surfaces; ducks and flamingos rely on fine, comblike structures called lamellae at the edges of the beak to filter food from the water. Both groups feed on particles that are minute relative to their own size. One of us (Wassersug) has found that certain tadpoles can grow to metamorphosis on single-cell organisms less than 10 microns in diameter.

Intermittent suction feeders, the final group in our categorization, are relatively unspecialized; they differ little from the myriad nonsuspension-feeding planktivorous fish that visually locate individual prey and suck them in. What distinguishes the intermittent suction feeders is that they do not alter their swimming speed or direction to focus attention on individual plankton.

Intermittent suction feeding is intermediate between continuous ram feeding and suction feeding on individual prey. Depending on feeding conditions, an intermittent suction feeder may tend more toward either end of this range. A predator whose jaws gape wide enough to engulf more than one typical prey at a time may engage in intermittent suction feeding when prey are too sparsely scattered or too large for continuous ram feeding. Light can also play an impor-



RORQUAL WHALE feeds on a school of krill and on filter-feeding mackerel. This humpback, an intermittent ram feeder, rushes forward to drive a single enormous gulp of water containing many fish and crustaceans into its mouth. It then closes its mouth and expels the water while the baleen fringes hanging from the roof of its mouth retain the prey. The mackerel, in contrast, are continuous ram feeders; they filter zooplankton from the water as they swim through it.



GILL RAKERS of a suspension-feeding fish extract food particles from water and direct them to the fish's esophagus. The water flows through the gill rakers and out between the gill arches. Three mechanisms may account for the capture of food particles too small to be trapped by simple sieving (*a*): direct interception by sticky surfaces (*b*), inertial impaction on structures that divert the water flow (*c*) and gravitational deposition as dense particles settle out of the water flow (*d*).

tant role: threadfin shad, for example, can feed by intermittent suction in complete darkness, but in bright moonlight they visually seek out and attack single large prey.

How does a suspension feeder separate its prey from the enormous volumes of water that pass through its mouth? Probably the best-known separation mechanism of any vertebrate suspension feeder involves whalebone, or baleen. Prey become entangled in the rough fringes of the baleen plates that descend from the margins of the roof of whales' mouths. Baleen is flexible, and so the porosity of its fringes can be expected to vary depending on such hydrodynamic factors as water pressure and direction of flow. The stiffness of the fringes also differs among species depending on their feeding habits. Gray whales, which feed on coarse, abrasive material from the ocean bottom, have stiff, thick, short plates with a coarse, short fringe. At the other extreme, Sei whales, which prefer tiny copepods to larger prey, have a much finer, denser fringe.

Knowing that baleen and similar biological filters separate water and prey is not nearly enough. Just how do they do it? Michael C. LaBarbera of the University of Chicago, basing his arguments on the morphology of feeding structures, has suggested that the filters of suspension-feeding fishes and baleen whales may function essentially as sieves: the spacing of filter elements is smaller than the size of many of their prey. Yet many vertebrate suspension feeders retain particles that are small enough to pass between the elements of their filtering arrays.

Moreover, static sieves offer high resistance to flow and clog rapidly. Industrial engineers have found that other mechanisms may be more effective than sieving in many cases; in vertebrates such mechanisms may operate separately or in conjunction with sieving. The chief ones are direct interception, in which particles that touch a sticky (mucus-covered) feeding structure will adhere to it; inertial impaction, in which particles that are slightly denser than the surrounding fluid strike the feeding structure while the fluid bends to pass around it; gravitational deposition, in which objects denser than water settle out of a flowing stream; and electrostatic entrapment, in which charged feeding surfaces preferentially retain particles of a certain charge or degree of hydrophobic attraction.

Daniel I. Rubenstein of Princeton University and Mimi A. R. Koehl of the University of California at Berkeley have theorized that inertial impaction may come into play for whales and large fishes that engage in continuous ram feeding where food particles are relatively large and dense and flow velocities are high. The mucus found on the filters of lamprey larvae, tadpoles and some species of fish also points to capture mechanisms other than sieving. In tadpoles, for example, mucus-coated organs called branchial food traps, which lie above the gills, collect suspended food particles even though they are not porous.

One clear pattern that has emerged from our survey is that filters in vertebrate suspension feeders are invariably flexible; their pore size is not fixed. Even relatively rigid filters, such as the gill rakers of fishes and the lamellae of birds, vary their porosity as the gill arches move toward or away from the midline of the body or as the beaks open and close. Filters in organisms such as whales and tadpoles are arranged so that the pressure of the water passing through changes their porosity. (The pressure, of course, is affected in turn by the density of any prey and other material on the filters.) All these factors change the flow at the microscopic level and make it difficult to establish what entrapment mechanisms are effective at any instant.

Each of our four groups of suspension feeders has evolved different anatomical adaptations. Continuous ram feeders have an enormous head; it accounts for from 25 to 28 percent of the total body length of the basking shark, megamouth shark and whale shark, for example. The head of baleen whales can be as much as one third the length of the body.

Despite their large head, continuous ram feeders such as the sharks and whales have proportionately smaller brains than their nonsuspension-feeding relatives. The orbits and eyes are

also reduced, and the eyes are directed laterally rather than forward; vision, after all, is not an important factor in capturing their prey. (Some smaller suspension feeders such as anchovies and mackerels have large orbits, however; the orbit occupies as much as half of the length of the Atlantic mackerel's skull. These species switch from ram feeding to feeding on individual prey depending on prey size, density and location.)

Baleen whales furnish an example of just how far an organism can become adapted to continuous ram feeding. The head is enormous in comparison with that of toothed relatives. At the back of the skull there is a large sloping area that serves as an attachment point for the back muscles, which must contract to resist the downward torque exerted by water pressure when the mouth is open. The skull's shape reduces the torque the muscles must counteract; the bones arch up and back, so that when the mouth is open it is in line with the long axis of the spinal column. Furthermore, baleen whales have proportionately the shortest neck of any mammal. The cervical vertebrae are fused, greatly limiting head movement.

In addition to changes in the shape of the skull, continuous ram feeders have either fewer and smaller teeth than their relatives or no teeth at all (as in the case of the baleen whales). The bones of the upper and lower jaws are also adapted to suspension feeding; they may be elongated and broadened compared with those of nonsuspension-feeding relatives but are usually not thickened or heavily ossified. This quality is consistent with their less demanding role as regulators of water flow into the mouth rather than as active elements in the capture or chewing of prey.

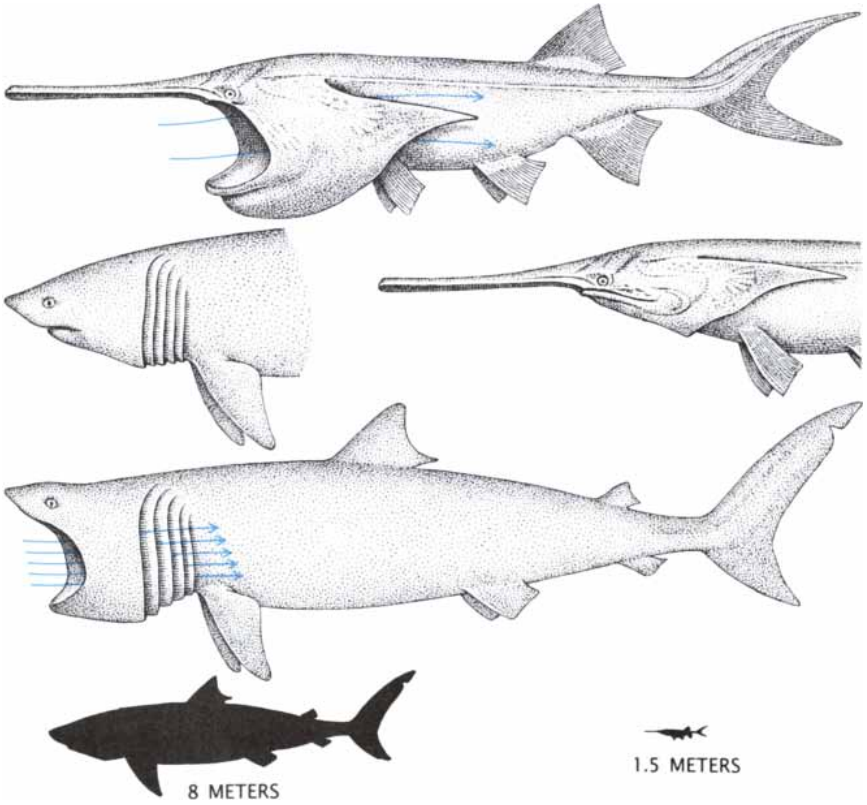
Continuous ram-feeding fishes are unique in their extensive elaboration of the branchial (gill) apparatus. In the Indian mackerel, for example, the comblike gill rakers (elements that retain food particles as water passes through the gill arches) are so long that they protrude past the corners of the mouth when the mouth is open.

In addition to the proliferation of gill rakers and the loss of teeth, the development of accessory organs known as epibranchial organs has been a dominant evolutionary theme in suspension-feeding fishes such as herring, shad, sardines, menhaden and alewife. These organs consist of a pair of pockets at the top rear of the pharynx, just above the esophagus. The gill rakers appear to channel small food particles into the pockets, where they apparently coalesce into a lump that the fish then swallows. These organs are particularly associated with the ingestion of small food particles.

Right and bowhead whales are the largest continuous suspension feeders. In these whales the filtering structures inside the mouth present high resistance to water flow; the resulting pressure drag slows the animals down when the mouth is open. As a result, the whales are unable to capture large, evasive prey that can swim out of the path of the open mouth. Instead they consume relatively small, slow-swimming prey. They have been observed, for example, capturing copepods by skimming the surface of the water with the mouth open.

A prominent body characteristic of rorqual whales, the intermittent ram feeders, is a set of longitudinal grooves in the throat region. These grooves flatten out and allow the skin to expand when water is taken into the mouth. The lower jaw of a blue whale may rotate back as much as 90 degrees, and in a matter of seconds the whale can engulf 60 cubic meters of water—approximately half of its body volume. The floor of the whale's mouth is compliant, so that the whale can engulf a large volume of water rapidly; otherwise the pressure in front of the swimming whale would push food and water away from the mouth. Rorqual whales ingest not only relatively slow-moving prey such as krill but also schools of fish or squid.

It is, of course, the large size and consequent high inertia of the swimming whale that makes this feeding method possible. According to models developed by Lisa S. G. Orton of Duke University and Paul Brodie of the New Bedford Institute of Oceanography in Nova Scotia, the pressure on the snout of a fin whale is great enough at normal swimming speeds to expand the mouth completely once it begins to open. The soft tissues surrounding the mouth have become specialized to present little resistance to this expansion. The tongue, for example, has a cleft that runs down its midline; under the impetus of inrushing water, the cleft allows the tongue to fold and invert completely into an intermuscular cleft in the floor of the mouth, thus making room for the



WIDELY GAPING JAWS of continuous ram suspension feeders enable them to process enormous volumes of water as they swim forward through food-bearing regions. The paddlefish (*top*) feeds mostly along muddy river bottoms. Although its primary diet is freshwater plankton, the paddlefish may inadvertently ingest smaller suspension-feeding fish as well. The basking shark (*bottom*), in contrast, feeds in midocean.

enormous volume of water and prey.

The whales' jaws must close quickly around the mouthful of water to keep prey from swimming out. The jaws, their joints and the muscles that cross them have been specialized to accomplish this feat. The jawbones of rorqual whales are not fused at the chin like those of most other mammals but are connected loosely by soft tissue. The temporomandibular articulation between the jawbones and skull is freer than that of any other mammal; the mandibles can rotate about their long axis as well as swing up, down, left and right. Moreover, in one species, fibrocartilage tissue spans the temporomandibular joints. Such tissue will store energy when the joints are forced open by the oncoming water and then release it like a loaded spring to help snap the jaws shut.

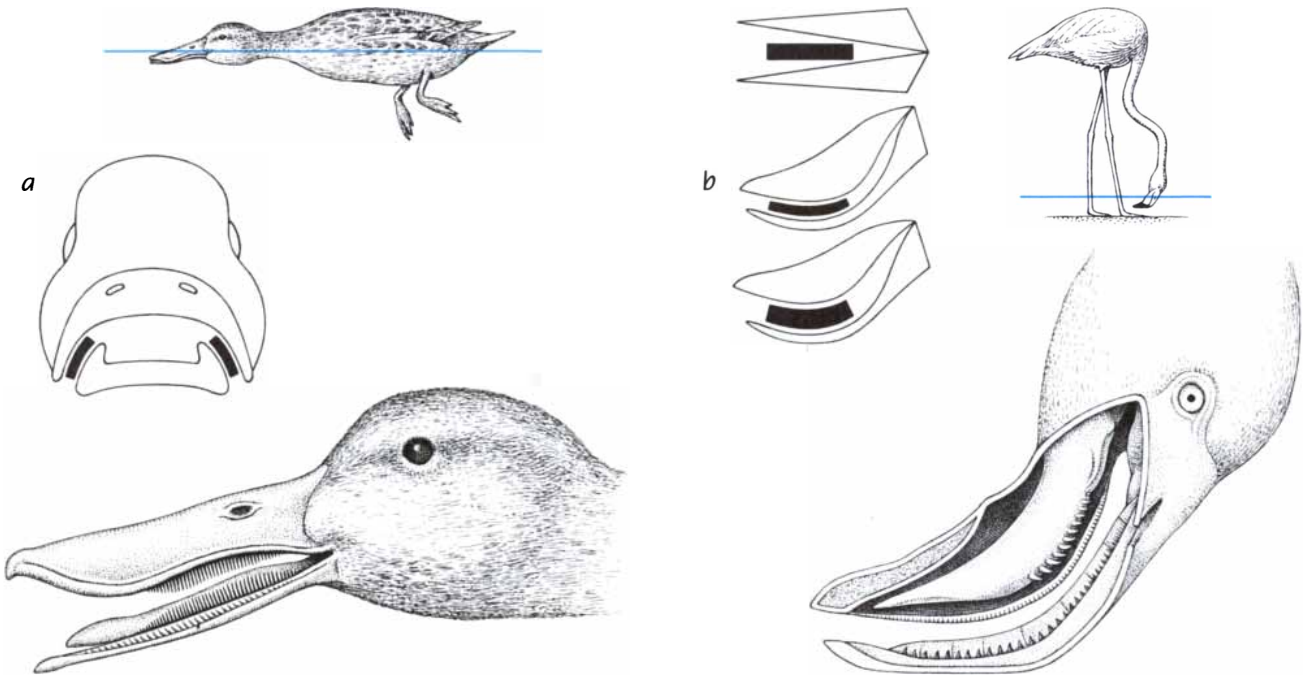
In contrast to the ram feeders, which open their mouth as wide as possible to engulf prey, the continuous suction feeders have a reduced oral gape. The narrow mouth opening excludes large particles from the mouth—which may be particularly important when these animals feed near the bottom, where there is a lot of nonfood material that could clog their entrapment systems.

Tadpoles show the widest variation of feeding structure among the continuous suction feeders. Some are grazers and have sharp beaks with which they tear up and suck in large food items, such as bottom-growing plants. Others can eat only microscopic prey; they live in midwater regions and lack hardened mouthparts. Both kinds of species employ essentially the same water-transport and particle-entrapment mechanisms, but species that feed on microscopic plankton have smaller pumping muscles than those that live near the bottom.

Once food particles have entered the tadpole's mouth, rows of papillae on the floor and roof of the mouth act as both sieves and funnels; they shunt large particles directly into the esophagus and deflect smaller ones into the pharynx. Farther back in the pharynx are the gill filters and the mucus-covered branchial food traps; small food particles stick to the surfaces of the traps and are shuttled down to the gill filters by a mechanism not completely understood. The size and porosity of the gill filters in different species of tadpoles correlate roughly with the size of the particles the species ingests, but tadpoles commonly trap particles much smaller than the pore size of their filters. One of us (Wassersug) has suggested from study of the anatomy that direct interception of food particles on mucus-covered surfaces may be as important as sieving.

Birds do not possess any of the filtering structures that capture food for whales, fishes and tadpoles. Instead the bills of suspension-feeding birds have rows of keratinized projections called lamellae, which can strain small particles from the water. The bills of mallard ducks, for example, are relatively long and broad; when the mouth is closed, the large upper bill completely covers the lower bill and the tongue, and when the mouth is slightly ajar, the adjacent lamellae of the upper and lower bills can act together as a sieve.

The tongue of ducks is also specialized for suspension feeding; it is broad, with large, swollen areas in front and back. The anterior swelling moves away from the roof of the mouth when the tongue moves forward, allowing water to enter the mouth from the front; the jaws also



SUSPENSION-FEEDING BIRDS filter food particles from water by means of comblike lamellae at the edges of the beak. They draw water into the mouth and push it out again by oscillating the tongue rapidly forward and back. The shoveler duck (*left*) has lamellae along the outer edges of its lower beak and the inner edges of its upper beak. Front and rear swellings on the duck's tongue enable it to draw water in through the front of the beak and expel it to the sides and rear through the lamellae; the tongue acts as both piston and valve. The flamingo's tongue acts as a simple piston (*right*); the bird controls the flow of water by opening and closing its beak slightly in synchrony with the motion of the tongue. The sharp bend in the beak makes the distance between the upper and lower jaws from the bend to the tip of the beak more uniform than it would be if the beak were straight (*b*). This uniformity persists regardless of how far the jaws are open. The sides of the duck's bill (*a*) show the same effect because of the curvature of the jaws and the overlap between them.

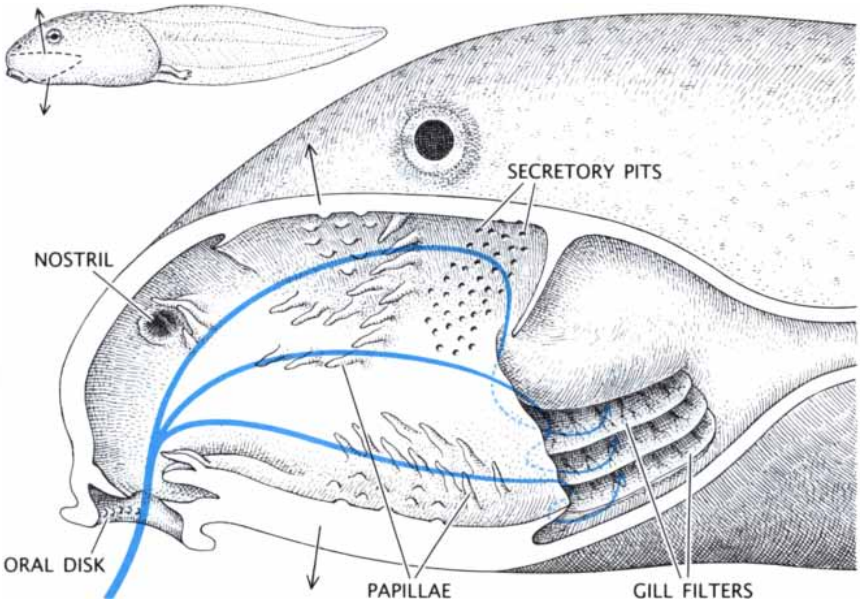
open slightly to suck water in. When the tongue is drawn back, the front swelling presses against the roof of the mouth, forcing the water backward and out the sides of the bill through the lamellae. The rear swelling has projections called lingual scrapers on both sides; as the tongue moves back and forth, these scrapers act to dislodge food particles from the lamellae and draw them toward the esophagus. In mallard ducks, the tongue tip moves forward about 11 millimeters and then back again every 60 to 70 milliseconds. This rapid oscillation moves the water at high velocity, and so the lamellae may capture food particles by inertial impaction as well as by sieving.

The oral anatomy of the flamingo—the ultimate suspension-feeding bird—is even more specifically adapted to pumping water and filtering food from it. The flamingo's tongue is narrower and more cylindrical than that of a duck; it lies in a bony trough formed within the deep lower jaws and typically oscillates from four to 17 times a second when the flamingo is feeding. The upper beak has a keel that descends toward the tongue and forms a cap over it. This tight confinement suggests that the tongue acts as a simple piston. There are spines toward the rear of the tongue that may help move food particles toward the esophagus, but the region of contact between the tongue and the multiple rows of fine lamellae on the upper and lower beaks is far smaller than the equivalent region in ducks.

As in ducks, the mouth of the flamingo opens and closes slightly with each tongue cycle. These subtle jaw movements ensure that water flows through the mouth only in one direction: in through the front of the beak and out through the lamellae at the sides. Opening and closing and lateral movements of the jaws can change the spacing between the lamellae significantly and thus control their resistance to flow.

Just how the lamellae of birds capture food particles is still not clear. Until recently researchers presumed that only sieving was operative. As with other suspension-feeding vertebrates, there is a general relation between the fineness, density and number of filters and the typical size of the particles ingested.

The correlation is not absolute, however; the diet of some ducks cannot be predicted from the morphology of their lamellae. Jan G. M. Kooloos, Gart A. Zweers and their colleagues at the University of Leiden have shown that ducks continuously adjust the porosity of their filtering mechanism to the size of the available food particles by altering their gape and the elevation of the upper beak. Furthermore, some ducks, such as mallards and tufted ducks, can feed efficiently even when their lamellae have been damaged extensively. Mechanisms other than simple sieving must be operating, but full understanding of their details awaits further study.



TADPOLE GILL ELEMENTS act as filters for feeding. Muscles in the floor of the tadpole's mouth and at the side of its head pump water through the gills. Secretions of sticky mucus on surfaces above the gill filters catch food particles as they flow by. Some tadpoles tear up bottom-growing plants and ingest the fragments; other species can subsist entirely on particles the size of bacteria filtered from open water.

In many ways, vertebrate suspension feeders are a black box: food particles go into the mouth, but they don't come out again. Although many anatomical elements of suspension feeding have been identified, the detailed hydrodynamics involved is still unclear. The basic musculature of the tongue and pharynx of many suspension feeders also remains unexamined; workers have studied the stomach contents of many whales but not the cranial muscles that help get food into the stomach.

Some progress has been made; for example, one of us (Sanderson), along with Joseph J. Cech, Jr., of the University of California at Davis, recently managed to place an endoscope inside the mouth of a freely swimming suspension-feeding fish and watch as food particles interacted with filtering elements. We also inserted a probe to record the velocity of the water. Preliminary analysis indicates that the patterns of water flow are not consistent with predicted mechanisms of particle capture. Only additional experiments will determine whether these results apply to other vertebrate suspension feeders. Other approaches to analyzing suspension-feeding mechanisms are now being taken, ranging from high-speed X-ray movies to supercomputer models of fluid flow through fishes' mouths; with luck some of them will bear fruit.

FURTHER READING

THE FEEDING MECHANISMS OF BALEEN WHALES. August Pivorunas in *American Scientist*, Vol. 67, No. 4, pages 432–440; July/August, 1979.

A COMPARATIVE STUDY OF THE BUCCAL PUMPING MECHANISM OF TADPOLES. Richard J. Wassersug and Karin Hoff in *Biological Journal of the Linnean Society*, Vol. 12, pages 225–259; November, 1979.

THE SUSPENSION FEEDING MECHANISM OF THE LARVAL LAMPREY *PETROMYZON MARINUS*. Jon Mallatt in *Journal of Zoology, London*, Vol. 194, Part 1, pages 103–142; May, 1981.

A REVIEW OF PLANKTIVOROUS FISHES: THEIR EVOLUTION, FEEDING BEHAVIOURS, SELECTIVITIES, AND IMPACTS. Xavier Lazzaro in *Hydrobiologia*, Vol. 146, pages 97–167; 1987.

COMPARATIVE MECHANICS OF FILTER-FEEDING OF *ANAS PLATYRHYNCHOS, ANAS CLYPEATA* AND *AYTHYA FULIGULA* (AVES, ANSERIFORMES). J.G.M. Kooloos, A. R. Kraaijeveld, G.E.J. Langenbach and G. A. Zweers in *Zoomorphology*, in press.

Unconscious Mental Functioning

Rigorous, quantitative studies of psychotherapy are challenging certain widely held views of how the unconscious mind works and how patients in therapy make progress

by Joseph Weiss

How does psychotherapy work? In the past several years my colleagues and I of the Mount Zion Psychotherapy Research Group in San Francisco have found some surprising answers to this question and to an inseparable, and equally profound, question—namely, how does the unconscious mind function?

It is generally assumed that human beings cannot carry out unconsciously the same kinds of intellectual activities they perform consciously, such as making plans and assessing risks. Yet our studies of patients in psychotherapy indicate that, in fact, people can unconsciously think, anticipate consequences and make and carry out decisions and plans. What is more, patients enlist these abilities in the service of working to become well—in the service of gaining control over their irrational beliefs, feelings and behaviors.

Most ideas about unconscious mental functioning and the therapeutic process have been developed by psychotherapists on the basis of clinical impressions, recorded as notes or recalled from memory. This approach has been fruitful for producing new ideas, but it cannot assess their relative value. Hence, the Mount Zion Psychotherapy Research Group, which I co-direct with my colleague Harold Sampson, has gone beyond the clinical method, choosing instead to depend on reliable data and to carry out rigorous, quantitative investigations that are planned in advance to test specific hypotheses.

Our studies have focused on comparing the merits of two distinct psychoanalytic (Freudian) hypotheses about the nature of unconscious mental functioning. The two hypotheses can be evaluated empirically because they make distinct, testable predictions about how patients will behave during therapy.

Psychoanalytic theory assumes that, beginning in early childhood, powerful "mental contents" (thoughts and feelings) that are not tolerable to the conscious self become buried beneath what may be called a repression barrier, which consists of forces that prevent repressed material from reaching awareness. Nevertheless, the buried contents—which Sigmund Freud initially thought consisted mainly of sexual and aggressive impulses and later concluded also included beliefs, judgments and such emotions as shame and guilt—continue to influence mood and behavior. They thereby contribute to the symptoms that can propel people into therapy, such as inexplicable depression, unfocused anxiety and maladaptive behaviors that seem to be resistant to conscious control.

Because the repression barrier limits people's conscious knowledge of why they act and feel as they do, thus limiting their control over parts of their personality, a major focus of psychoanalytic therapy is to help patients weaken their repressions and confront repressed material. The therapist asks patients to free-associate: to put into words any thoughts, images, memories and feelings that enter the mind. Such associations provide clues to the patients' unconscious motivations and concerns. The therapist may then interpret the patients' statements, pointing out what appear to be the underlying wishes, fears, beliefs, guilts or the like. Presumably the therapist's interpretations help patients to gain insight into the effects of the unconscious mind on their conscious thoughts, feelings and behaviors.

Although both hypotheses we examined are compatible with the basic tenets of psychoanalytic theory, they make profoundly different assumptions about the degree of control people can exert over unconscious mental functioning. The first, and more influential, hypothesis, which I call the dynamic hypothesis, basically derives from Freud's early writings and assumes that people have little or no control over their unconscious mental life. It proposes that the unconscious mind consists predominantly of two kinds of forces. On the one hand, sexual and aggressive impulses seek gratification and push toward consciousness; on the other hand, repressive forces oppose the impulses. The impulses and repressions interact dynamically, much as forces interact in the physical world. For instance, two

JOSEPH WEISS is a training analyst at the San Francisco Psychoanalytic Institute, clinical professor of psychiatry at the University of California Medical Center in San Francisco and co-director (with Harold Sampson) of the Mount Zion Psychotherapy Research Group. He got his M.D. from the University of Cincinnati Medical School and was trained in psychoanalysis at the San Francisco Psychoanalytic Institute, where he himself became a training analyst in 1962. Weiss began to develop his ideas about how the unconscious mind functions, and to test those ideas informally, in 1958. He and Sampson founded the Mount Zion Psychotherapy Research Group in 1972 to evaluate the ideas by formal, quantitative research methods; the group now has about 30 active members, most of whom combine research with private practice. A number of the group's projects have been supported by the National Institute of Mental Health and by grants from the Mount Zion Hospital in San Francisco.

equal and opposite forces may nullify each other, and a strong force may overwhelm a weaker one. Through their interactions, these forces determine behavior.

The other hypothesis, which I call the unconscious-control hypothesis (or, more simply, the control hypothesis), elaborates on ideas Freud put forth briefly in some of his later writings. It assumes that people can exert some control over unconscious functioning. According to this idea, people keep impulses and other mental contents repressed not because the repressive forces are necessarily more powerful than the unconscious impulses but because individuals can unconsciously decide (by extrapolating from the past and by assessing current reality) that experiencing or expressing certain repressed material would be dangerous. For example, patients may decide that expressing love to another person would lead to rejection or humiliation.

The control hypothesis further assumes that patients in therapy have a strong unconscious wish to get better. They therefore desire unconsciously (as well as consciously) to bring forward repressed mental contents and explore their significance. Hence, they may unconsciously decide to lift repressions and permit hidden material to come forth when doing so no longer seems dangerous.

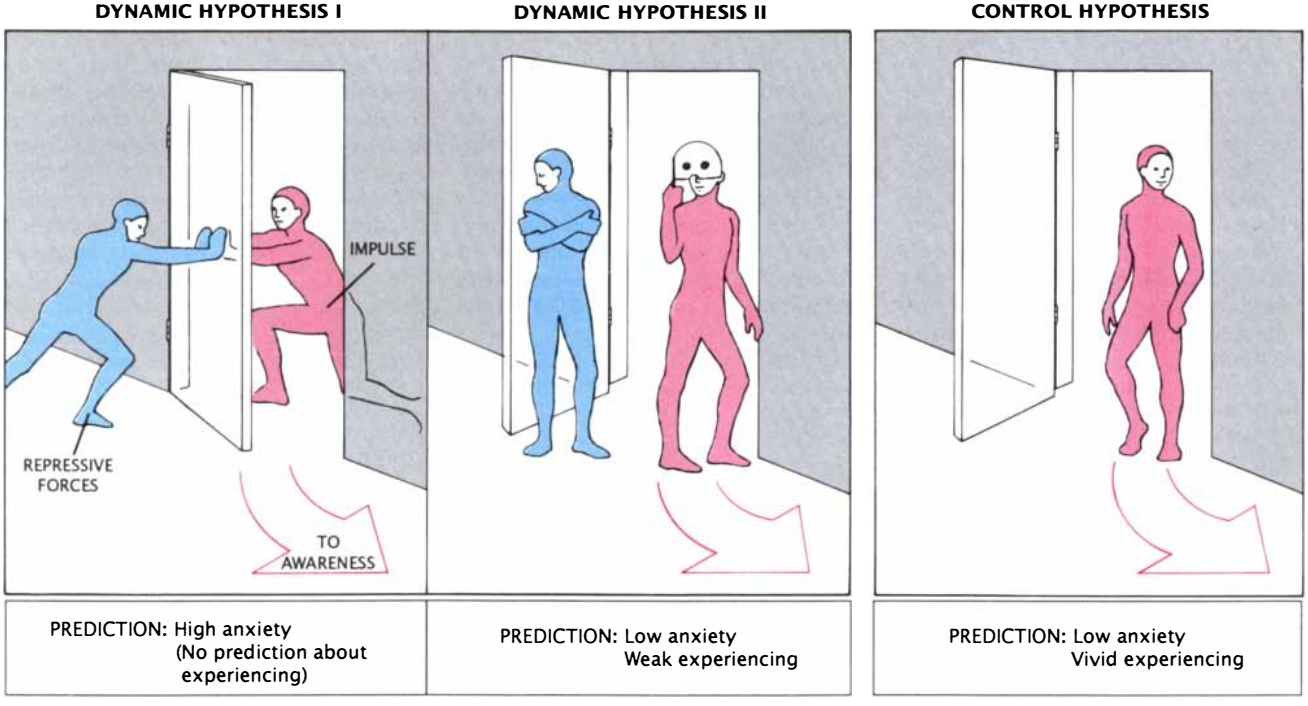
We compared the dynamic and control hypotheses by examining the different explanations they offer for certain events that commonly occur in therapy. In one such event, patients spontaneously become aware of unconscious material (such as hostility toward a sibling) without the therapist's having mentioned it previously.

In explaining this phenomenon of therapy, both hypotheses emphasize certain special features of the relationship between patient and therapist. The relationship is strictly professional and confidential and is confined to the office; also, the therapist maintains an impersonal, investigatory attitude toward patients. Each hypothesis, however, focuses on a different aspect of the relationship.

The dynamic hypothesis emphasizes the frustrations evoked by the therapist's attitude. It assumes that during treatment the therapist becomes the object of patients' unconscious impulses; for instance, a patient's repressed hostility toward a rivalrous sibling may become transferred onto the therapist, so that the patient becomes unconsciously angry at the therapist. Patients' unconscious impulses may then be frustrated by the therapist's detachment and so may become more intense—just as hunger becomes more gnawing when it goes unsatisfied. Frustrated impulses push with intensified pressure toward consciousness but, because of the counterpressure of the repression barrier, are in most instances prevented from reaching awareness.

An intensified impulse may spontaneously become conscious in spite of this barrier in two ways, according to the dynamic hypothesis. In one instance the impulse may become so powerful relative to the forces of repression that it overwhelms those forces and breaks through the repression barrier. For example, a patient may react with rage to a slight provocation by a therapist.

The frustrated and intensified impulse may also enter consciousness if



COMPETING HYPOTHESES about unconscious mental functioning were assessed by testing the predictions they make about how patients will feel and act while becoming spontaneously aware of an unconscious impulse not previously mentioned by the therapist. The "dynamic" hypothesis holds that an intolerable repressed impulse can become conscious spontaneously in two ways. It can overwhelm the repressive forces (*left*), in which case the hypothesis predicts that conflict between the impulse and the repressions it arouses will make patients anxious. The impulse may also become disguised (*center*) and reach consciousness by escaping notice by the repressive forces. Presumably patients will then feel little anxiety (because the impulse does not come into conflict with the repressions) but will also fail to experience the undisguised impulse vividly. The "control" hypothesis (*right*) holds that an unconscious impulse will become conscious when patients remove repressions after determining by unconscious reasoning that they can safely experience the impulse. Patients are predicted to feel little anxiety and to experience the undisguised impulse vividly. The author's findings support the control hypothesis.

it is disguised, so that what becomes conscious is a much milder, derivative version of the impulse. Because the disguise hides the original impulse's intensity and impetus to action, the milder impulse escapes the usual repressive forces. For example, someone's anger at the therapist may enter consciousness in the form of an apparently isolated, "silly" fantasy of tripping the therapist.

The dynamic hypothesis predicts that if an impulse emerges because it overwhelms the forces of repression, it will be in conflict with these forces as it becomes conscious; consequently, patients will be tense and anxious. The hypothesis also predicts that if an impulse emerges because it has been disguised, it will emerge without producing anxiety. Because in the instance of disguise patients will have little awareness of the power of the original impulse, the hypothesis further predicts that they will not experience the impulse vividly.

I should note here that the word "experience" has a specific meaning, which is based on a standardized instrument known as the Experiencing Scale. People are said to be vividly experiencing a thought or feeling if they articulate it clearly and focus on understanding its significance. Individuals who talk in vague terms and whose words seem detached from their feelings are deemed to be weakly experiencing a thought or feeling.

In contrast to the dynamic hypothesis, the control hypothesis holds that, far from frustrating the patient, the therapist's noncritical attitude and pledge of confidentiality create an atmosphere of security. Patients therefore may conclude unconsciously that they can safely bring to consciousness certain repressed material. For example, a man who has consistently repressed hostile impulses out of fear of provoking retaliation may decide in therapy that he can express his anger at the therapist without risking punishment. He may then lift the repressions opposing that anger.

The control hypothesis postulates that because patients bring forth repressed material only after they have unconsciously overcome their worry about the consequences, they will not feel especially anxious as they become aware of the material. Moreover, having no need to disguise their thoughts and feelings, they will be able to face and reflect on them; in other words, they will experience them vividly.

In sum, then, the competing predictions are quite distinct. The dynamic hypothesis predicts either great anxiety or, in the case of a disguised impulse, low anxiety coupled with low experiencing of the emerging material. The control hypothesis, in contrast, predicts low anxiety coupled with vivid experiencing.

To evaluate which prediction is correct, we examined the mental functioning of one patient, Mrs. C, during psychoanalysis, an intensive form of therapy in which a patient sees the analyst four or five times a week. (Analysts are psychotherapists who have undergone several years of special training in Freudian theory and technique.)

Our basic task in this study was conceptually quite simple: to determine Mrs. C's levels of anxiety and experiencing at the moment she became aware of previously repressed material. This undertaking was easier said than accomplished, however, as a much simplified description of our careful and time-consuming methods will demonstrate.

In all of the studies we conduct, our data are collected from verbatim transcripts of therapy sessions that have been tape-recorded with the patient's consent. The transcripts provide accurate, permanent data and permit work done by one team of investigators to be double-checked by other teams.

In this particular study we examined the transcripts of Mrs. C's first 100 sessions with her analyst, who was male. To begin identifying material that was repressed early in the analysis but later emerged spontaneously, we first isolated all mental contents—ideas, attitudes, memories and feelings—that appeared in the transcripts of later sessions (sessions 41 to 100) but not in earlier transcripts. For instance, we found that in one of Mrs. C's later sessions she recalled wanting to kill her brother and that neither she nor her therapist had made any mention of such a desire earlier.

Mrs. C might actually have been aware of some of the new items even though she had not discussed them. Such items would not qualify as having been repressed, and so we had to eliminate them from consideration. To do so, we asked some 20 analysts or analysts-in-training to act as judges. We gave each judge our list of new contents, along with condensed versions of the transcripts of the first 10 analytic sessions. The judges then determined, on the basis of their own understandings of the patient's problems, which items they thought had been repressed during the first 10 sessions. (As is good scientific practice, our judges are always "blind" to extraneous information that could influence their determinations.) The judges also indicated on a five-point rating scale the degree of confidence they had in each assessment. We accepted only those items that were judged reliably to have been repressed—items that received a mean confidence rating of 4 or 5.

Now we had a new list of items that had been repressed and had then become conscious. Our next task was to eliminate contents that were in any way suggested to Mrs. C by her therapist. We gave independent judges the new list, together with a list of everything the analyst had said during the first 100 sessions. The judges found only one content that had been discussed by the therapist; that content was eliminated.

Only then were we in a position to measure Mrs. C's anxiety level and degree of experiencing during the emergence of previously repressed material. We gave transcripts of five-minute segments, or excerpts, of her speech during therapy to two sets of judges, who independently rated them on either of two scales designed to measure anxiety. (We included every segment assumed to contain previously repressed material, together with many segments chosen at random. The judges were blind as to which segment was which.) One of the scales, the Mahl Scale, assesses anxiety based on the frequency of speech disruptions; more disruptions reflect higher anxiety. The other, the Gottschalk-Gleser Scale, assesses anxiety based on the frequency of references to certain topics—for instance, mutilation, shame and death.

From the mean values of the judges' ratings, we were able to conclude that Mrs. C was no more anxious when the repressed material was emerging than at other times. Indeed, ratings made with the Mahl Scale indicated that she was much less anxious than usual when the repressed material was becoming conscious. (All of the findings I discuss are highly statistically significant unless otherwise noted.)

The Experiencing Scale served for rating how vividly Mrs. C experienced the emerging contents. Based on an analysis of the same segments as before, we found that when previously repressed contents were reaching awareness, Mrs. C experienced them more vividly than she experienced randomly selected contents.

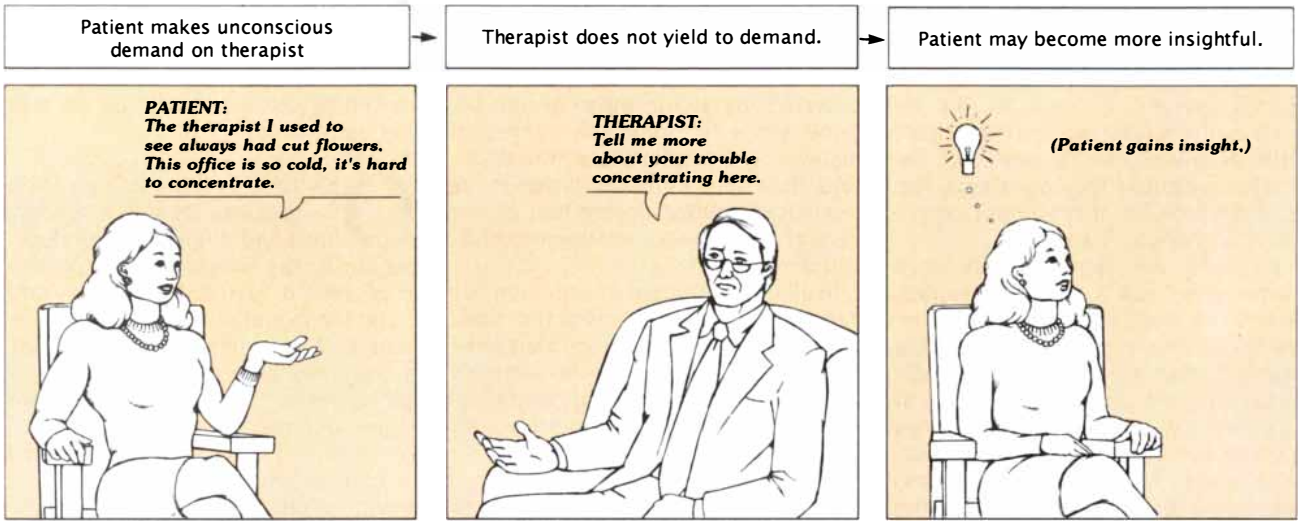
Mrs. C's low level of anxiety and her vivid experiencing are compatible

with the control hypothesis but not with the dynamic hypothesis. This result indicates that patients most often bring forth repressed material spontaneously because the therapeutic setting makes doing so seem safe, not because repressed impulses have become intensified by frustration.

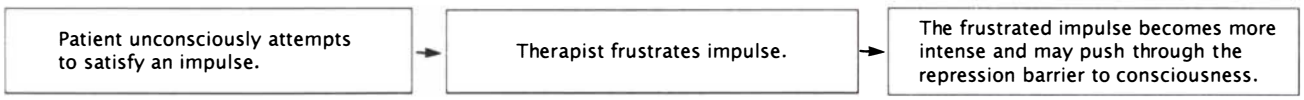
In another test of the two hypotheses, we examined why a particular sequence of events happens frequently in many analyses. In this sequence a patient unconsciously makes a powerful demand on the analyst. (For example, a male patient may, by describing a sexual fantasy involving a woman who resembles his female therapist, unconsciously express a wish to have a sexual relationship with the therapist.) The analyst, in responding, does not yield to the demand but instead reacts in a nonyielding but noncritical way—perhaps by interpreting the patient's statement or asking a question. After receiving such a nonyielding reply, certain patients may gain insight into their unconscious motivations and make progress in treatment.

The dynamic hypothesis explains the sequence by assuming that the demands are usually attempts to gratify repressed impulses. By not giving in to the demand, the analyst frustrates the impulse, causing it to become more intense. If the impulse becomes powerful enough, it may break through the repressive barrier to consciousness; if it remains repressed, it may, by its unconscious pressure, facilitate the therapeutic process.

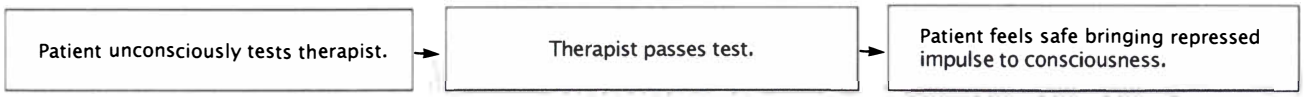
a **THE OBSERVATION:**



b **HOW THE DYNAMIC HYPOTHESIS EXPLAINS THE OBSERVATION:**



c **HOW THE CONTROL HYPOTHESIS EXPLAINS THE OBSERVATION:**



d **THE RESULTS**

CHANGE IN PATIENT'S BEHAVIOR	PREDICTED BY DYNAMIC HYPOTHESIS	PREDICTED BY CONTROL HYPOTHESIS	RESULTS
Change in anxiety	Increase	Decrease	Decrease
Change in relaxation	Decrease	Increase	Increase
Change in boldness	———	Increase	Increase
Change in loving	———	Increase	Increase

DYNAMIC AND CONTROL HYPOTHESES offer different explanations for the observation that patients sometimes gain insight into themselves after they make an unconscious demand on the psychotherapist (*a*)—such as a veiled demand for an apology (*cartoon*). The dynamic hypothesis (*b*) assumes that patients make such demands because of an unconscious desire to satisfy their impulses (a hostile impulse, in the cartoon). If the therapist frustrates an impulse by not yielding, the impulse intensifies and may break through to consciousness. (The woman in the cartoon might realize she wants to hurt the therapist.) The resulting prediction is that as patients become aware of their repressed impulses, they will become more anxious than usual and less relaxed. The control hypothesis (*c*) assumes that patients make demands as a test, to determine whether the therapist can comfortably tolerate hearing about repressed material. It holds that patients will be reassured by seeing that the therapist is comfortable and will then lift their repressions. (The woman in the cartoon might realize she does not have to worry that her hostility will hurt the therapist easily.) Hence, they will feel less anxious than usual and more relaxed; they will also feel bolder and more loving. Again, a test of the predictions (*d*) was supportive of the control hypothesis.

The control hypothesis, in contrast, assumes patients often have quite a different motivation when they unconsciously make demands. Recall that the control hypothesis assumes patients unconsciously want to become aware of repressed material. Yet they also fear that expressing certain of their repressed thoughts and feelings might endanger them by damaging their relationship with the therapist. They therefore make unconscious demands as a means of indirectly testing the therapist's tolerance of such thoughts and feelings. If patients see that the therapist does indeed comfortably tolerate their demands (thereby passing their tests), they may gain the sense of safety needed to lift their repressions. For example, the man who unconsciously attempts to seduce his female therapist may feel reassured by finding that the therapist is not seducible and also is not angered by his sexual demands; he may then gain confidence in her and become less afraid of confronting his repressed sexual interest in her. (The control hypothesis also assumes that for some patients, "passing the test" actually requires the therapist to yield to certain unconscious demands. Nevertheless, for the sake of study, we focus only on instances in which the therapist passes the test by not yielding.)

What predictions do the two hypotheses make about patients' responses to the therapist's nonyielding behavior? The dynamic hypothesis predicts that after the therapist's response, the patients' frustrated impulses will intensify and come into increased conflict with repressive forces, and so the patients will become more anxious and less relaxed. In contrast, the control hypothesis predicts that because certain patients feel reassured by the therapist's nonyielding response, they will become less anxious and more relaxed.

The control hypothesis also makes other predictions about the patients' response, ones the dynamic hypothesis does not address. It predicts that after a test is passed, patients will feel more loving toward the therapist, whom they will view as helpful. Moreover, because patients will feel safer, they are likely to become emboldened—free to express themselves more directly.

To test these predictions, we once again studied the transcripts of the first 100 sessions of Mrs. C's analysis. As our research design required, Mrs. C is a person who seemed to benefit more when her therapist did not accede to her unconscious demands than when he did.

We identified a large number of interactions in which Mrs. C made a powerful unconscious demand on the analyst. To compare the predictions of the competing hypotheses directly, we then honed this list so that it would include only demands that fit the criteria of both hypotheses—demands that could simultaneously be construed as attempts to gratify an unconscious impulse and as important unconscious tests.

We did so by asking judges who believed in the dynamic hypothesis to identify instances in which Mrs. C was trying to gratify an important unconscious impulse and by asking judges who believed in the control hypothesis to select instances in which Mrs. C was unconsciously posing an important test of the analyst. We then picked out the interactions that both groups of judges selected: the overlap set of interactions.

Next we rated the interventions, or responses, made by the therapist in these selected interactions. Much as we did before, we asked judges accustomed to thinking in terms of the dynamic hypothesis to rate each intervention according to how well it frustrated Mrs. C's unconscious impulses. At the same time, we asked judges accustomed to thinking in terms of the control hypothesis to decide how well the interventions passed her unconscious tests. This procedure enabled us to distinguish interventions that frustrated Mrs. C's impulses or passed her tests from those that acceded to her demands or failed to pass her tests. (It also enabled us to correlate the ratings of the two sets of judges, to assure ourselves that, as our research plan required, the interventions one set of judges saw as frustrating Mrs. C's unconscious impulses were the same ones the other set of judges saw as passing her unconscious tests.)

Now we asked judges to rate segments of Mrs. C's speech (before and after the intervention) on separate scales designed to assess anxiety, relaxation, boldness and loving feelings. A different set of judges applied each scale, and all judges were blind as to whether a segment came before or after an intervention. We then calculated the changes in Mrs. C's ratings on each scale by means of a statistical method that yields what is called a residualized gain score.

Our compilation of the data revealed that when the analyst did not accede to Mrs. C's unconscious demands (and therefore either passed her tests or frustrated her wishes), Mrs. C became less anxious, bolder, more relaxed and more loving. Taken together, our results provide support for the control hypothesis. They indicate that patients make unconscious demands on therapists as a way of assuring themselves that they can safely confront the thoughts, feelings and memories that are blocked by repression.

How can therapists best help patients gain insight? I have developed a version of the control hypothesis that addresses this issue. My version specifies the kinds of interpretations that will have an immediate beneficial effect on patients. Moreover, it predicts that such interpretations will contribute to the overall value of the therapy.

My hypothesis, which my colleagues and I have now tested, assumes that psychological problems are rooted not in repressed impulses that maladaptively seek gratification (as the dynamic hypothesis would say) but in painful ideas known as "pathogenic" beliefs. These unconscious, irrational ideas cause, and help to maintain, psychological disturbance. They are maladaptive in that they prevent people from seeking certain highly desirable goals; the beliefs warn people that if they do try to attain such goals, they will endanger themselves and suffer fear, anxiety, guilt, shame or remorse. Pathogenic beliefs can vary from person to person. For instance, one man may suffer from the belief that he should not seek independence, lest he make those close to him feel rejected and unhappy; another may suffer from the belief that if he seeks a sexual relationship with a woman to whom he is attracted, he will be severely punished.

Patients in psychotherapy are highly motivated to disconfirm—disprove—their pathogenic beliefs so that they will feel safe in moving toward the goals those unconscious beliefs warn them against. They work to disprove the beliefs in part by unconsciously testing them in relation to the therapist, sometimes by making unconscious demands and sometimes in other ways. For instance, a woman who feared she would hurt her parents and her male therapist by becoming independent might experiment with independent behavior in her sessions by disagreeing with the therapist's opinions and then unconscious-

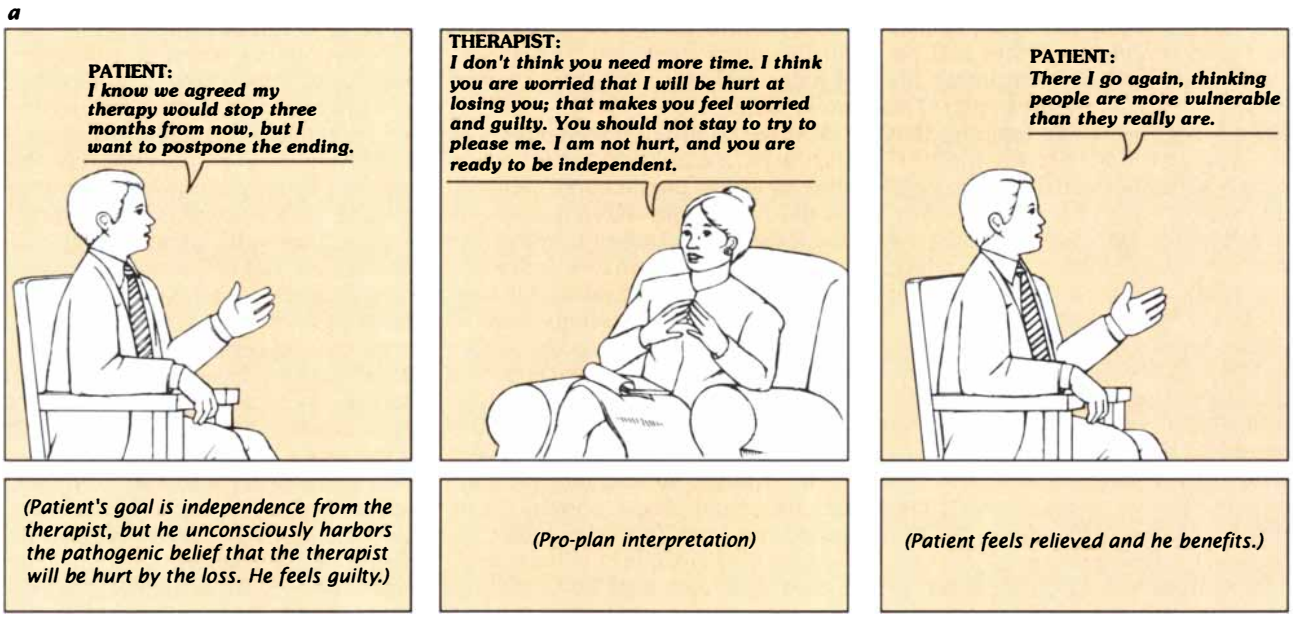
ly monitoring him to see if he feels hurt. Moreover, patients use the interpretations made by the therapist to gain insight into their unconscious beliefs and to realize that the dangers their beliefs warn against are not real.

Because the behavior of patients in treatment is directed unconsciously at disproving their pathogenic beliefs and moving toward certain goals, it can be said to be planful. I therefore label as "pro-plan" those interpretations that would be expected to help patients carry out their unconscious plans and as "anti-plan" those interpretations that tend to hinder such progress.

To understand the concept of planful interpretations better, consider the hypothetical case of a young man who feels guilty about wanting to become independent from his parents because he unconsciously believes that if he weakens his ties, his parents will be devastated. One of his unconscious plans during therapy would be to gather evidence against this belief so that he can feel comfortable, say, moving out of the house. He might begin carrying out his plan by testing unconsciously the therapist's ability to tolerate a bid for independence. For instance, he might discuss his fantasy of taking a job in another city. He might find helpful—pro-plan—the interpretation that he is reluctant to consider seriously such a move for fear of hurting his therapist and family, whereas he might find hindering—anti-plan—the suggestion that his fantasy reflects an unconscious desire to avoid facing his dependence on his therapist and family.

My hypothesis gives rise to the prediction that patients will react differently to pro-plan interpretations than to irrelevant or anti-plan interpretations. The hypothesis predicts that when a pro-plan interpretation is offered, patients will gain insight into their beliefs and will experience their feelings vividly. When an anti-plan interpretation is offered, however, patients will feel conflicted; they will also become less insightful and experience their feelings less vividly.

My colleagues and I have tested these predictions in studies of three short-term psychoanalytic therapies, each lasting 16 weeks. Our methods for correlating pro-plan and anti-plan interpretations with levels of insight and experiencing were straightforward. First we had clinician judges determine a patient's pathogen-



"PRO-PLAN" INTERVENTION by a therapist is meant to help a patient disprove an unconscious belief (*a*). The author's version of the control hypothesis holds that patients harbor unconscious "pathogenic" beliefs that warn of dire consequences (such as guilt or shame) if certain goals are sought. In therapy, patients unconsciously plan to disprove these handicapping beliefs (*left*). The hypothesis predicts that pro-plan comments by the therapist—those that support the plan (*center*)—will help patients, who will gain insight into their unconscious motivations and will experience their feelings vividly (*right*). Comments that are "anti-plan"—that support pathogenic beliefs—are predicted to be unhelpful and to be followed by a decline in insight and experiencing. In a test of the hypothesis, clinician judges were given transcribed segments of a patient's speech before and after the therapist's interpretations (*b*). The judges scored the patient's levels of insight and experiencing on standard scales (*c*). Other judges independently evaluated the "planfulness" of the interpretations. Correlations of planfulness with changes in insight and experiencing suggest that pro-plan interpretations are indeed beneficial.

ic beliefs and the goals those beliefs warned against. The judges did this by examining the transcripts of the intake interview (which preceded the actual start of therapy) and the first two therapeutic sessions. On the basis of their understandings of the transcripts, the judges then generated what they considered to be pro-plan insights: ones that would be expected to help disprove the patient's pathogenic beliefs.

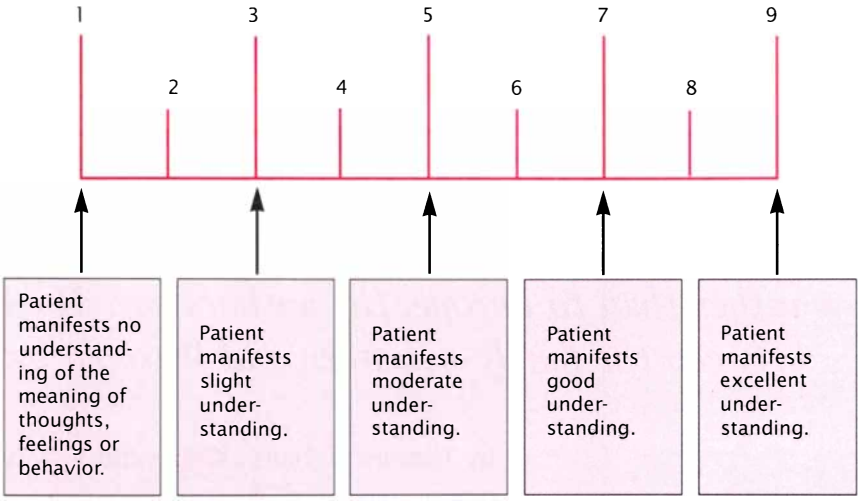
Next we gave a second set of clinicians the lists of beliefs, goals and proposed helpful insights, along with a list we had compiled of the therapist's actual interventions, which consisted of all comments intended to convey insight. On the basis of these lists, the judges (who were blind to the patient's responses) rated the comments on a scale ranging from strongly anti-plan to strongly pro-plan.

We still had to determine the extent to which the patient's insight and experience changed in response to the therapist's interventions. This we did by isolating stretches of the patient's speech immediately before an interpretation was made and immediately after. One set of judges, who were blind both to the order in which the patient's comments were made and to the therapist's interventions, rated each segment on what is called the Morgan Patient Insight Scale. Another set of judges rated the segments on the Experiencing Scale. Then residualized gain scores were calculated.

For each patient we found a strong correlation between pro-plan interpretations and improvements in both insight and level of experiencing. For instance, when the mean level of planfulness of all interpretations in a given therapy session was correlated with the mean value of the patient's level of experiencing in that session, the correlation was very high: .78 for one patient, .54 for the second and .57 for the third. Hence, the research supported my hypothesis.

Does the strategy of making pro-plan interpretations actually help to improve the mental health of patients in the long run, or are the positive effects generally restricted to the session during which the interpretations are made? In an attempt to find out, we interviewed the patients and assessed their mental health (by a battery of so-called outcome measures) six months after they parted ways with their therapists.

One patient, it turns out, did very well, another did moderately well and the third did poorly; these outcomes correlated well with the kinds of interpretations the patients received. When we calculated the fraction of interpretations that were pro-plan, anti-plan and ambiguous, we found that in the first case, the percentages were 89, 2 and 9, respectively; in the second, they were 80, 2 and 18, and in the third, they were 50, 6 and 44. Although the findings must be replicated in a larger number of cases to be convincing, the data do suggest that patients who receive a high percentage of pro-plan interpretations will do better than patients who do not.



SCALE for rating patients' insight into their thoughts, feelings and behaviors is one of many tools by which the author and his colleagues quantify observations of patients in their research. The rating scale shown here is much simpler than the Morgan Patient Insight Scale mentioned in the text but produces comparable results.

How, then, does psychotherapy work? Our studies suggest some answers. By demonstrating that pro-plan interpretations are helpful and anti-plan interpretations are not, we have supported the hypothesis that patients suffer from unconscious pathogenic beliefs and that they make and carry out unconscious plans for disproving these beliefs—with the aim of breaking down the obstacles to attaining their conscious and unconscious goals.

Our other studies say something about how patients go about carrying out such plans. The findings are consistent with the assumption that patients bring forward their pathogenic beliefs as well as other unconscious thoughts and feelings only when they decide they may do so safely—that is, without risking internal dangers such as guilt or shame or external dangers such as loss of love. We also found evidence that patients actively seek assurances of safety by unconscious testing of the therapist.

It seems that the cognitive capacities of the unconscious mind have been underappreciated and that human beings can unconsciously carry out many intellectual tasks, including developing and executing plans for reaching certain goals. The implications for therapy are obvious: good therapists will carefully infer their patients' unconscious goals and strive to offer interpretations designed to advance movement toward these goals.

FURTHER READING

THE EMERGENCE OF WARDED-OFF CONTENTS. Suzanne Gassner, Harold Sampson, Joseph Weiss and Suzanne Brumer in *Psychoanalysis and Contemporary Thought,* Vol. 5, No. 1, pages 55–75; 1982.

THE IMMEDIATE EFFECTS OF TRANSFERENCE INTERPRETATIONS ON PATIENTS' PROGRESS IN BRIEF, PSYCHODYNAMIC PSYCHOTHERAPY. Polly Fretter. Doctoral Dissertation, University of San Francisco, 1984. *Dissertation Abstracts International,* Vol. 46, No. 6. University Microfilms No. 85-12112.

INSIGHT, THE MIND'S EYE: AN EXPLORATION OF THREE PATIENTS' PROCESSES OF BECOMING INSIGHTFUL. Jessica Broitman. Doctoral Dissertation, Wright Institute, 1985. *Dissertation Abstracts International,* Vol. 46, No. 8. University Microfilms No. 85-20425.

HOW DO INTERPRETATIONS INFLUENCE THE PROCESS OF PSYCHOTHERAPY? George Silberschatz, Polly B. Fretter and John T. Curtis in *Journal of Consulting and Clinical Psychology,* Vol. 54, No. 5, pages 646–652; October, 1986.

THE PSYCHOANALYTIC PROCESS: THEORY, CLINICAL OBSERVATIONS AND EMPIRICAL RESEARCH. Joseph Weiss, Harold Sampson and the Mount Zion Psychotherapy Research Group. Guilford Press, 1986.

The Early History of Indo-European Languages

The common ancestor of these languages has been traced to Asia rather than to Europe, the authors say. The once-clear distinction between the family's Eastern and Western branches is now blurred

by Thomas V. Gamkrelidze and V. V. Ivanov

Linguistics, the scientific study of language, can reach more deeply into the human past than the most ancient written records. It compares related languages to reconstruct their immediate progenitors and eventually their ultimate ancestor, or protolanguage. The protolanguage in turn illuminates the lives of its speakers and locates them in time and place.

The science developed from the study of the Indo-European superfamily of languages, by far the largest in number of languages and number of speakers. Nearly half of the world's population speaks an Indo-European language as a first language; six of the 10 languages in which *Scientific American* appears—English, French, German, Italian, Russian and Spanish—belong to this superfamily.

Over the past 200 years, linguists have reconstructed the vocabulary and syntax of the postulated Indo-European protolanguage with increasing confidence and insight. They have tried to unravel the paths by which the language broke into daughter languages that spread throughout Eurasia, seeking at the origin of those paths the homeland of the protolanguage itself. The early investigators placed the homeland in Europe and posited migratory paths by which the daughter languages evolved into clearly defined Eastern or Western branches. Our work indicates that the protolanguage originated more than 6,000 years ago in eastern Anatolia and that some daughter languages must have differentiated in the course of migrations that took them first to the East and later to the West.

The reconstruction of ancient languages may be likened to the method used by molecular biologists in their quest to understand the evolution of life. The biochemist identifies molecular elements that perform similar functions in widely divergent species to infer the characteristics of the primordial cell from which they are presumed to have descended. So does the linguist seek correspondences in grammar, syntax, vocabulary and vocalization among known languages in order to reconstruct their immediate forebears and ultimately the original tongue. Living languages can be compared directly with one another; dead languages that have survived in written form can usually be vocalized by inference from internal linguistic evidence. Dead languages that have never been written, however, can be reconstructed only by comparing their descendants and by working backward according to the laws that govern phonological change. Phonology—the study of word sounds—is all-important to historical linguists because sounds are more stable over the centuries than are meanings.

Early studies of Indo-European languages focused on those most familiar to the original European researchers: the Italic, Celtic, Germanic, Baltic and Slavic families. Affinities between these and the "Aryan" languages spoken in faraway India were noticed by European travelers as early as the 16th century. That they might all share a common ancestor was first proposed in 1786 by Sir William Jones, an English jurist and student of Eastern cultures. He thus launched what came to be known as the Indo-European hypothesis, which served as the principal stimulus to the founders of historical linguistics in the 19th century.

In their reconstruction of the ancestral Indo-European language, the early linguists relied heavily on Grimm's law of *Lautverschiebung* ("sound shift"), which postulated that sets of consonants displace one another over time in predictable and regular fashion. The law was posed in 1822 by Jacob Grimm, who is more widely famed for the anthology of fairy tales he wrote with his brother, Wilhelm. Grimm's law explained, among other things, why in the Germanic languages certain hard consonants had persisted despite their universal tendency to yield to soft ones. The set of softer, "voiced" consonants "b," "d," "g" (followed by momentary vibration of the vocal cords), posited in the protolanguage, had apparently given way to the corresponding hard set "p," "t," "k." According to Grimm's law, this had come about by "devoicing" those consonants ("p," for example, is unaccompanied by vocal vibration). Thus, the Sanskrit *dhar* is seen as an archaic form of the English "draw," which is itself more archaic than the German *tragen* (all of which mean "to pull").

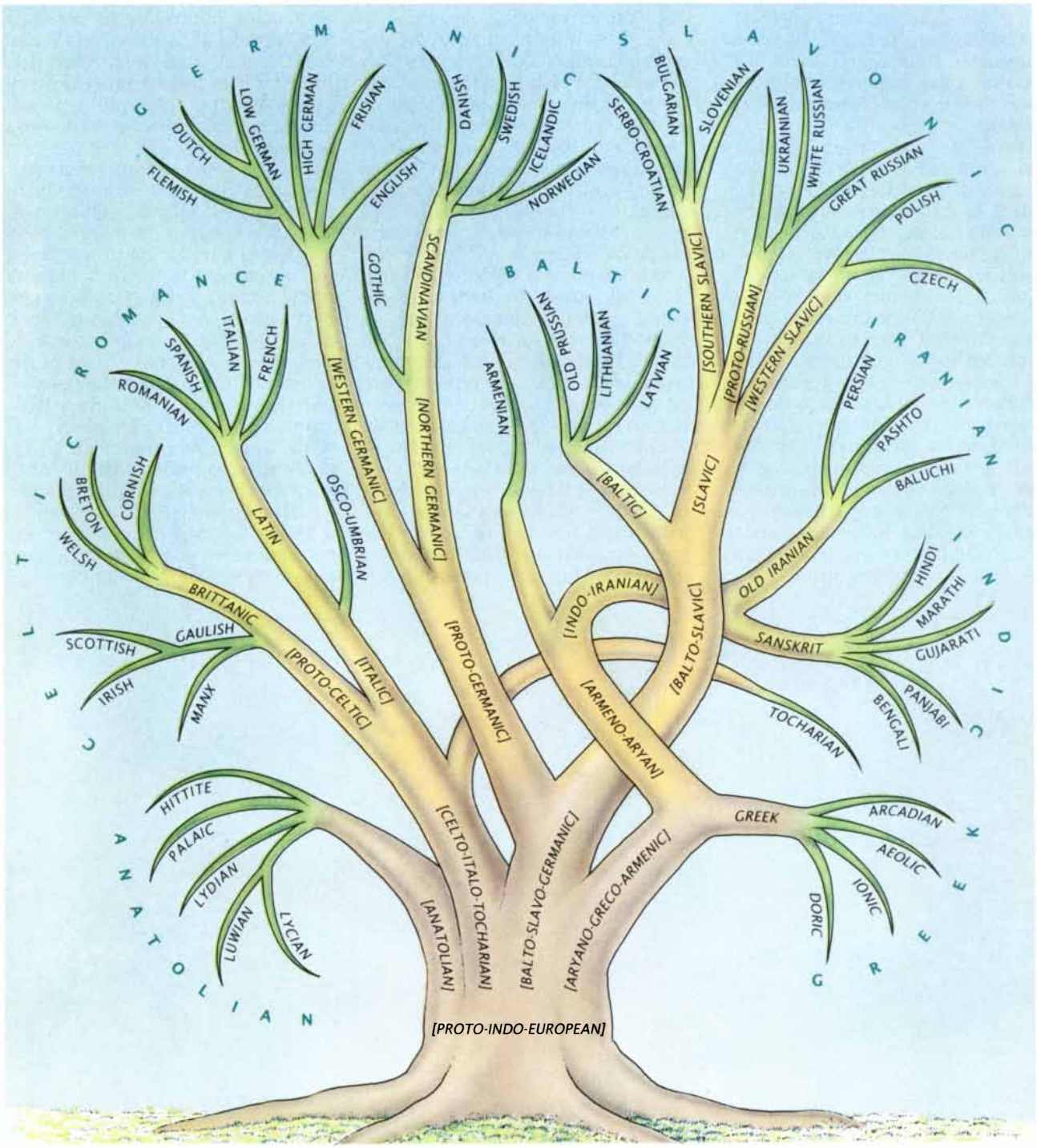
These rules were used to reconstruct an Indo-European vocabulary that implies how its speakers lived.

THOMAS V. GAMKRELIDZE and V. V. IVANOV are the authors of *The Indo-European Language and the Indo-Europeans,* a two-volume work published in Russian in 1984; an English version will be published this fall by Mouton de Gruyter. Gamkrelidze directs the Tsereteli Institute of Oriental Studies in Tbilisi and is a professor of linguistics at Tbilisi State University. Ivanov is professor of linguistics and chair of the department of Slavic languages at the Institute for Slavic and Balkan Studies in Moscow.
The authors wish to thank Gerard Piel, chairman emeritus of *Scientific American,* for helping to prepare this article for publication.

The words described a landscape and climate that linguists originally placed in Europe between the Alps in the south and the Baltic and North seas in the north [see "The Indo-European Language," by Paul Thieme; SCIENTIFIC AMERICAN, October, 1958].

More recent evidence now places the probable origin of the Indo-European language in western Asia. Three generations of archaeologists and linguists have thus far excavated and deciphered manuscripts in close to a dozen ancient languages from sites in modern Turkey and as far east as Tocharia, in modern Turkestan. Their observations, together with new ideas in pure linguistic theory, have made it necessary to revise the canons of linguistic evolution.

The landscape described by the



FAMILY TREE of the Indo-European languages can be traced back to a protolanguage that flourished more than 6,000 years ago. The protolanguage split into dialects, which evolved into distinct languages; these then fissioned into generations of daughter languages. Tocharian, a dead language of Asia, has ties to Celtic, an ancient European tongue. Similarities between the Balto-Slavic and Indo-Iranian families indicate that they influenced each other before their speakers moved north and south, respectively. Dead languages are shown in italics; languages that left no literary remains are enclosed in brackets.

protolanguage as now resolved must lie somewhere in the crescent that curves around the southern shores of the Black Sea, south from the Balkan peninsula, east across ancient Anatolia (today the non-European territories of Turkey) and north to the Caucasus Mountains [*see illustration below*]. Here the agricultural revolution created the food surplus that impelled the Indo-Europeans to found villages and city-states from which, about 6,000 years ago, they began their migrations over the Eurasian continent and into history.

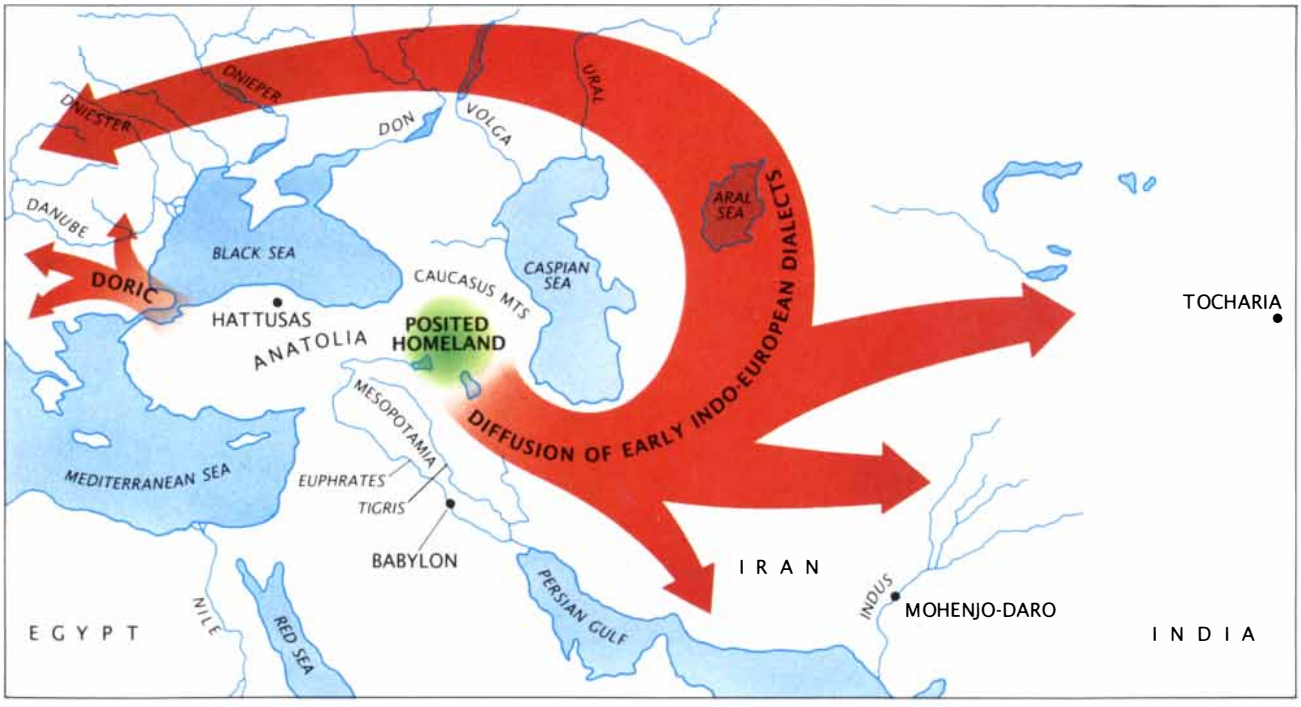
Some of the migrants invaded Anatolia from the East around 2000 B.C. and established the Hittite kingdom, which held all of Anatolia in its power by 1400 B.C. Its official language was among the first of the Indo-European languages to find its way into writing. Early in this century, Bedřich Hrozný, a linguist at Vienna University and later at Charles University in Prague, deciphered Hittite inscriptions (written in cuneiform, the ancient writing system based on wedge-shaped symbols) on tablets that had been found in the library of the capital at Hattusas, 200 kilometers east of modern Ankara. The library also contained cuneiform tablets in two related languages: Luwian and Palaic. The evolution of Luwian could be traced in later hieroglyphic inscriptions made around 1200 B.C., after the fall of the Hittite Empire. To this emerging family of Anatolian languages linguists added Lydian (closer to Hittite) and Lycian (closer to Luwian), known from inscriptions dating back to late in the first millennium B.C.

The appearance of Hittite and other Anatolian languages at the turn of the third to the second millennium B.C. sets an absolute chronological limit for the breakup of the Indo-European protolanguage. Because the Anatolian protolanguage had already fissioned into daughter languages by that point, investigators estimate that it departed from the parent Indo-European no later than the fourth millennium B.C. and possibly much earlier.

This inference is supported by what is known about the portion of the Indo-European community that remained after the Anatolian family had broken away. From that community came the languages that persisted into written history. The first to branch off was the Greek-Armenian-Indo-Iranian language community. It must have begun to do so in the fourth millennium B.C. because by the middle of the third millennium B.C. the community was already dividing into two groups, namely, the Indo-Iranian and the Greek-Armenian. Tablets in the Hattusas archives show that by the middle of the second millennium B.C. the Indo-Iranian group had given rise to a language spoken in the Mitanni kingdom on the southeast frontier of Anatolia that was already different from ancient Indian (commonly called Sanskrit) and ancient Iranian. Cretan-Mycenaean texts from the same eras as Mitanni, deciphered in the early 1950's by the British scholars Michael G. F. Ventris and John Chadwick, turned out to be in a previously unknown dialect of Greek. All these languages had gone their separate ways from Armenian.

Tocharian was another language family that diverged from the Indo-European protolanguage quite early. Tocharian is one of the more recently discovered Indo-European languages, first recognized in the early decades of the 20th century in texts from Chinese Turkestan. The texts were comparatively easy to decipher because they were written in a variant of the Brahmi script and were mainly translations from known Buddhist writings.

Not long ago, the British scholar W. N. Henning suggested that the Tocharians be identified with the Gutians, who are mentioned in Babylonian cuneiform inscriptions (in Akkadian, a Semitic language) dating from the end of the third millennium B.C., when King Sargon was building the first



MIGRATIONS AND CULTURAL DIFFUSION carried the Indo-European protolanguage from the homeland, which the authors place in the Transcaucasus, and fragmented it into dialects. Some spread west to Anatolia and Greece, others southwest to Iran and India. Most Western languages stem from an Eastern branch that rounded the Caspian Sea. Contact with Semitic languages in Mesopotamia and with Kartvelian languages in the Caucasus led to the adoption of many foreign words.

great Mesopotamian Empire. If Henning's views are correct, the Tocharians would be the first Indo-Europeans to appear in the recorded history of the ancient Near East. Lexical affinities of Tocharian with Italo-Celtic give evidence that the speakers of the two language families had associated in the Indo-European homeland before the Tocharians began their migration eastward.

The diverging pathways of linguistic transformation and human migration may now be traced back to a convergence in the Indo-European protolanguage and its homeland. This has followed from the revision in the canons of phonology we mentioned above. An uncontested peculiarity of the sound system of the protolanguage, for example, is the near absence, or suppression, of one of the three consonants "p," "b" or "v," which are labials (consonants sounded with the lips). Traditionally, it had been thought that "b" was the suppressed consonant. Subsequent studies in phonology indicated, however, that if one of the three labial consonants is lacking in a language, it is least likely to be the one sounded as "b" in English and other living European languages.

On that basis we decided to reexamine the entire system of consonants posited for the protolanguage, and as early as 1972, we proposed a new system of consonants for the language. Our proposal remains in the crucible of debate from which consensus forms in every science. The debate now focuses more strongly on features that relate the Indo-European protolanguage to other major language families and that have at last begun to bring their common ancestor into view.

According to classical theory, the "stop" consonants—those that are sounded by interruption of the outward flow of the breath that excites the vibration of the glottis, or vocal cords—are divided into three categories [*see top of illustration on this page*]. The labial stop consonant "b" appears in the first column as a voiced consonant; the parentheses enclosing it there indicate its supposed suppression. It is associated with two other voiced stop consonants: "d" (stopped by the forward part of the tongue against the palate) and "g" (stopped by the back of the tongue against the palate).

In the scheme we have developed [*see bottom of illustration on this page*], the corresponding consonants are sounded with a glottalized stop: a closure of the throat at the vocal cords that prevents the outward flow of breath. Here the voiceless labial stop ("p′") appears suppressed, followed by "t′" and "k′." As ("p′") is to ("b"), voiceless and voiced, respectively, so "t′" is to "d" and "k′" is to "g." Glottalized stops occur in many different language families, particularly those of northern Caucasian and southern Caucasian (Kartvelian) provenance. The glottalized stop—which hardens a consonant—tends to weaken and disappear in most languages of the world. So we surmised that—among the labial stops—it was the "p′" rather than the "b" that most likely had been suppressed in the Indo-European protolanguage.

VOICED	VOICED ASPIRATES	VOICELESS
(b)	bh	p
d	dh	t
g	gh	k
GLOTTALIZED	**VOICED/ASPIRATES**	**VOICELESS/ASPIRATES**
(p')	b/bh	p/ph
t'	d/dh	t/th
k'	g/gh	k/kh

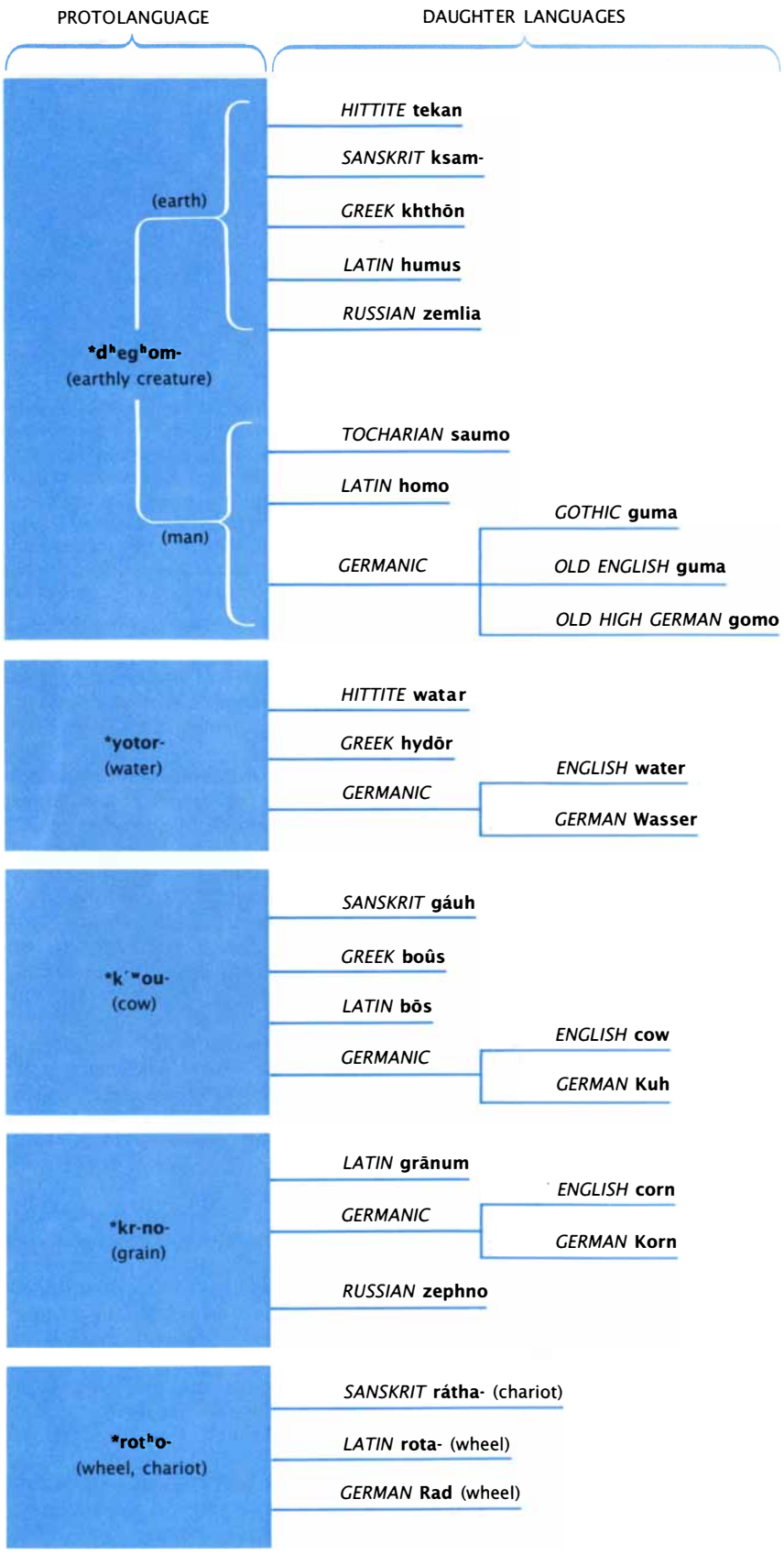
THREE SERIES OF STOPS (consonants produced when the air passage is closed) characterized proto–Indo-European. The classical model (*top*) posits that one series was voiced (followed by a vibration of the vocal cords, as in the "g" in "tiger"), another was voiced and aspirated (with an "h" sound not found in English) and the third was voiceless (like the "k" in "disk"). In the authors' model (*bottom*), the first series was glottalized (formed voicelessly by constricting the passage near the vocal cords, as in the Cockney pronunciation of the "t" in "bottle"), the second had voiced and voiced-aspirated forms and the third had voiceless and voiceless-aspirated forms. Glottalized stops bear a prime; absent stops are within parentheses.

Our so-called Indo-European glottalic system, which has been constructed by comparing the phonology of the living and the historically attested Indo-European languages, appears more probable than the classical one. The near absence of the labial phoneme ("p′") finds a natural phonological explanation in relation to the evolution of the other two glottalized stops and to the entire system of stops shown above.

In revising the consonant system of the Indo-European protolanguage, we have also called into question the paths of transformation into the historical Indo-European languages. Our reconstruction of the protolanguage's consonants shows them to be closer to those of the Germanic, Armenian and Hittite daughter languages than to Sanskrit. This neatly reverses the classical conception that the former languages had undergone a systematic sound shift, whereas Sanskrit had faithfully conserved the original sound system.

The transformation of consonants from parent to daughter languages may be illustrated by the word "cow" in English and *Kuh* in German; in Sanskrit the word for "ox" is *gáuh,* and in Greek it is *boûs.* All have long been recognized as descending from a common Indo-European word for "ox," or "cow." The word has different forms, however, in the glottalic and classical systems. In the glottalic it has the voiceless consonant **k′ʷou-* (the asterisk before a word designates it as a word in the protolanguage), which makes it phonetically closer to the corresponding words in English and German than to those in Greek and Sanskrit.

In the classical system the word is **gwou,* which is practically the same as that in Sanskrit. In accordance with Grimm's law, the transformation of **gwou* to the German would require devoicing of the first consonant from "g" to "k." And so the glottalic system seems to make the most sense: it eliminates the need for devoicing and correlates the voiceless stops in the Germanic languages (German, Dutch, Scandinavian and English) with voiceless glottalized stops in the ancestral Indo-European protolanguage. In this respect the Germanic languages are more archaic than Sanskrit and Greek. The glottalic system is seen, correspondingly, as more conservative than the classical system. It has brought the



WORD GENEALOGIES are traced as far as literary records go and then are reconstructed, for the preliterate period, on the basis of laws governing the evolution of sounds. Reconstructed words are marked with an asterisk. Many Indo-European languages derive words for "man" or "earth" from **dʰegʰom-*, a root in the protolanguage.

protolanguage closer to some of its daughter languages without resorting to such difficult phonological transformations as that from "g" to "k."

We can learn more about the earliest Indo-Europeans from other aspects of their reconstructed vocabulary. Some words, for example, describe an agricultural technology whose existence dates back to 5000 B.C. By that time the agricultural revolution had spread north from its origins in the Fertile Crescent, where the first archaeological evidence of cultivation dates back to at least 8000 B.C. From this region agriculture also spread southward to sustain the Mesopotamian civilizations and westward to Egypt. The Indo-European words for "barley," "wheat" and "flax"; for "apples," "cherries" and their trees; for "mulberries" and their bushes; for "grapes" and their vines; and for the various implements with which to cultivate and harvest them describe a way of life unknown in northern Europe until the third or second millennium B.C., when the first archaeological evidence appears.

The landscape described by the reconstructed Indo-European protolanguage is mountainous—as evidenced by the many words for high mountains, mountain lakes and rapid rivers flowing from mountain sources. Such a picture cannot be reconciled with either the plains of central Europe or the steppes north of the Black Sea, which have been advanced as an alternative homeland for the Indo-Europeans. The vocabulary does, however, fit the landscape of eastern Anatolia and Transcaucasia, backed by the splendor of the Caucasus Mountains. The language clothes its landscape in the flora of this region, having words for "mountain oak," "birch," "beech," "hornbeam," "ash," "willow" or "white willow," "yew," "pine" or "fir," "heather" and "moss." Moreover, the language has words for animals that are alien to northern Europe: "leopard," "snow leopard," "lion," "monkey" and "elephant."

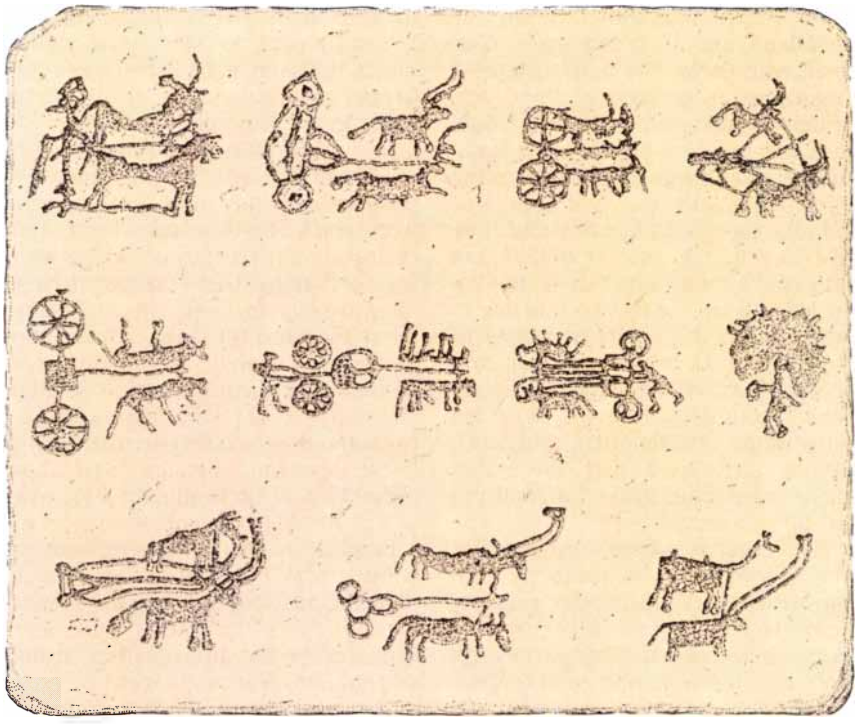
The presence of a word for "beech tree," incidentally, has been cited in favor of the European plains and against the lower Volga as the putative Indo-European homeland. Beech trees, it is true, do not grow east of a line drawn from Gdansk on the Baltic to the northwest corner of the Black Sea. Two species of beech (*Fagus orientalis* and *F. sylvatica*) flourish, however, in modern Turkey. Opposing the so-called beech argument is the oak argument: paleobotanical evidence shows that oak trees (which are

listed in the reconstructed language's lexicon) were not native to postglacial northern Europe but began to spread there from the south as late as the turn of the fourth to the third millennium B.C.

Another significant clue to the identification of the Indo-European homeland is provided by the terminology for wheeled transport. There are words for "wheel" (**rot^{h}o-*), "axle" (**hak^{h}s-*), "yoke" (**iuk′om*) and associated gear. "Horse" is **ek^{h}os* and "foal" **p^{h}olo.* The bronze parts of the chariot and the bronze tools, with which chariots were fashioned from mountain hardwoods, furnish words that embrace the smelting of metals. Petroglyphs, symbols marked on stone, found in the area from the Transcaucasus to upper Mesopotamia between the lakes Van and Urmia are the earliest pictures of horse-drawn chariots [*see illustration at right*].

The postulated homeland of the Indo-Europeans is, if not the only region, certainly one of the regions in which the horse completed its domestication and was harnessed as a draft animal in the fourth millennium B.C. From here wheeled vehicles spread with the migration of the Indo-Europeans in the third and second millennia B.C. eastward to central Asia, westward to the Balkans, and in a circular motion around the Black Sea and thence to central Europe.

The chariot provides significant evidence of cultural mixing, for chariots figured in the funerary and other religious rites of both the Indo-European peoples and the Mesopotamians. Contact with other western Asiatic cultures is also evidenced in the sharing of various mythological subjects—for example, the theft of the Hesperian apples by Hercules and similar tales in Norse and Celtic. Moreover, the Semitic and Indo-European languages each identify man with the earth. In Hebrew, *adam* means "man" and *adamah* means "earth"; both were derived from a root in the Semitic protolanguage (cf. Genesis 2:7, "... God formed man from the dust of the ground"). "Human" and "humus" came to English through Latin (*homo, humus*) from **d^{h}eg^{h}om-,* the word for "earth" and "man" (etymologically, "earthly creature") in the Indo-European protolanguage. The rooting of the Indo-European languages in eastern Anatolia is also suggested by the frequency of words borrowed from a number of languages that flourished there: Semitic, Kartvelian, Sumerian and even Egyptian. Conversely, Indo-European contributed words to each of those languages. Nickolai I. Vavilov, a prominent Soviet plant geneticist, found a vivid instance of such an exchange: the Russian *vinograd* ("grape"), the Italic *vino* and the Germanic *wein* ("wine"). These all reach back to the Indo-European **woi-no* (or **wei-no*), the proto-Semitic **wajnu,* the Egyptian **wns,* the Kartvelian **wino* and the Hittite **wijana.*



PETROGLYPHS from the Uzbek Soviet Socialist Republic (dating from the second or third millennium B.C.) provide archaeological corroboration of linguistic evidence that the Indo-Europeans had chariots. Wheeled vehicles, such as those drawn here, facilitated agriculture and the migrations that resulted from a growing hunger for land.

We concede that in the broad territory in which we have placed the homeland of the Indo-Europeans there is no archaeological evidence of a culture that can be positively linked to them. Archaeologists have identified, however, a number of sites that bear evidence of a material and spiritual culture similar to the one implied by the Indo-European lexicon. The Halafian culture of northern Mesopotamia decorated its vessels with religious symbols—bulls' horns and sometimes rams' heads, which are masculine symbols, and ritual images of leopard skins—that are shared by the somewhat later Çatal Hüyük culture of the seventh millennium B.C. in western Anatolia. Both cultures have affinities with the later Transcaucasian culture in the region embraced by the Kura and the Araks rivers, which includes southern Transcaucasia, eastern Anatolia and northern Iran.

In the 2,000 years before the Indo-Europeans who remained in the homeland began to write history, the success of the agricultural revolution brought a population explosion to the Indo-European community. The pressure of population, we may surmise, compelled the migration of successive waves of Indo-Europeans to fertile areas that were not yet cultivated.

The linguistic translocation of the Indo-European homeland from northern Europe to Asia Minor requires drastic revisions in theories about the migratory paths along which the Indo-European languages must have spread across Eurasia. Thus, the hypothetical Aryans who were said to have borne the so-called Aryan, or Indo-Iranian, language from Europe to India—and who were conscripted into service as the Nordic supermen of Nazi mythology—turn out to be the real Indo-Iranians who made the more plausible migration from Asia Minor around the northern slopes of the Himalaya Mountains and down through modern Afghanistan to settle in India. Europe is seen, therefore, as the destination, rather than the source, of Indo-European migration.

Speakers of the Hittite, Luwian and other Anatolian languages made rel-

atively small migrations within the homeland, and their languages died there with them. The more extensive migrations of speakers of the Greek-Armenian–Indo-Iranian dialects began with the breakup of the main Indo-European language community in the third millennium B.C. Two groups of Indo-Iranian speakers made their way East during the second millennium B.C. One of them, speakers of the Kafiri languages, survives to this day in Nuristan, on the southern slopes of the Hindu Kush in northeast Afghanistan. In *Five Continents,* a posthumous book recounting his many botanical expeditions between 1916 and 1933, Vavilov speculated that the Kafirs might perpetuate some "original relics" of Indo-Iranian.

The second group of Indo-Iranians, who followed a more southerly path into the Indus Valley, spoke a dialect from which the historical languages of India are descended. Their earliest literary ancestor is embodied in the *Rig-Veda* hymns, written in an ancient variant of Sanskrit. The indigenous peoples of the Indus Valley, known from the archaeological discoveries at their capital Mohenjo-Daro, were apparently displaced by the Indo-Iranians. After the separation of the Indo-Iranians and their departure for the east, the Greek-Armenian community remained for a time in the homeland. There, judging by the numbers of loan words, they had contact with speakers of Kartvelian, Tocharian and the ancient Indo-European languages that later evolved into the historical European languages. One such borrowing from the Kartvelian became the Homeric *kōas,* "fleece."

A bilingual cuneiform tablet found in the Hattusas archives records the mythological tale of a hunter in the then already dead Hurrian language along with a translation into Hittite. This remarkable discovery gave us the Hurrian word *ashi* from which Homer's *askós,* for "hide" or "fur," apparently stemmed. Before their migration to the Aegean, the Greeks borrowed the Hittite word *kursa,* which by a familiar phonological shift became *búrsa,* another synonym for "fleece." These words seem to confirm the Greeks' belief that their ancestors had come from western Asia, as recounted in the myth of Jason and the Argonauts, who sought the Golden Fleece in Colchis, on the eastern shore of the Black Sea. The evidence that the Greeks came thence to their historical homeland puts the Greek "colonies" on the northern shore of the Black Sea in a new light. The colonies may now be considered as very early settlements that were established when the Greeks began migrating to their final home in the Aegean.

The historical European languages—those that left literary remains—provide evidence that the dialects from which they descended had found their way into central Asia along with the Tocharians. These languages have many words in common. An example is the word for "salmon," once regarded as a weighty argument for a homeland in northern Europe. Salmon abounded in the Baltic rivers of Europe, and the word *lox* (German *Lachs*) in the Germanic languages is perhaps echoed by *lak-* in Hindu, for a lacquer of a pink color that evokes the color of salmon flesh. One species of salmon, *Salmo trutta,* is found in the streams of the Caucasus, and the *lak-s-* root denotes "fish" in earlier and later forms of Tocharian as well as in the ancient European languages.

The migration of the speakers of some of the early Indo-European dialects into central Asia is established by loan words from the Finno-Ugric language family, which gave rise to modern Finnish and Hungarian. Under the influence of Finno-Ugric, Tocharian underwent a complete transformation of its system of consonants. Words in the ancient European languages that are clearly borrowed from the Altaic and other languages of central Asia give further testimony to the sojourn of their speakers there.

Circling back to the west, the ancient Europeans settled for a time north of the Black Sea in a loosely federated community. Thus, it is not entirely wrong to think of this region as a second homeland for these peoples. From the end of the third through the first millennium B.C., speakers of ancient European languages spread gradually into Europe. Their coming is demonstrated archaeologically by the arrival of the seminomadic "pit grave" culture, which buried its dead in shafts, or barrows.

Anthropometry, which is the scientific measurement of the human body, has begun to chart the imposition of the Hittite physiognomy, typified in Hittite reliefs, on certain European populations. The blue-eyed, blond-haired Nordic must still be regarded as the product of interbreeding between the Indo-European invaders and their predecessors in the settlement of Europe. The culture of the indigenous populations of Europe is memorialized by the megalithic structures, such as Stonehenge, which they built near the periphery of the continent.

The languages of the previous inhabitants of Europe, with the exception of Basque—a non–Indo-European language with possible remote relatives in the Caucasus—were crowded out by the Indo-European dialects. Nonetheless, those languages made contributions to the historical European language families that account for certain differences among them. In his study of the megalithic cultures and their disappearance, as well as of the spread of farming from the ancient Near East, the British archaeologist Colin Renfrew has reached conclusions about the coming of the Indo-Europeans that agree well with ours [see "The Origins of Indo-European Languages," by Colin Renfrew; SCIENTIFIC AMERICAN, October, 1989].

Our deductions, resting so preponderantly on linguistic evidence, must find confirmation in archaeological investigations that remain to be done. Undoubtedly, the counting of base-pair substitutions in the DNA of human cells will contribute to the family tree of the speakers of the Indo-European languages and to the mapping of their migrations. Anthropometry and history also will contribute to the ultimate picture. Pending the elaboration and correction of our work, we may state with a high order of certainty that the homeland of the Indo-Europeans, the cradle of much of the world's civilization, was in the ancient Near East: "*Ex oriente lux*!"

FURTHER READING

INDO-EUROPEAN AND THE INDO-EUROPEANS: A RECONSTRUCTION AND HISTORICAL TYPOLOGICAL ANALYSIS OF A PROTOLANGUAGE AND PROTO-CULTURE. Parts I and II. Thomas V. Gamkrelidze and Vjacheslav V. Ivanov. Tbilisi State University, 1984.

ARCHAEOLOGY AND LANGUAGE: THE PUZZLE OF INDO-EUROPEAN ORIGINS. Colin Renfrew. Cambridge University Press, 1988.

RECONSTRUCTING LANGUAGES AND CULTURES: ABSTRACTS AND MATERIALS FROM THE FIRST INTERNATIONAL INTERDISCIPLINARY SYMPOSIUM ON LANGUAGE AND PREHISTORY, ANN ARBOR, NOVEMBER 8–12, 1988. Edited by Vitaly Shevoroshkin. Studienverlag Dr. Norbert Brockmeier, 1989.

IN SEARCH OF THE INDO-EUROPEANS: LANGUAGE, ARCHAEOLOGY AND MYTH. J. P. Mallory. Thames and Hudson, 1989.

WHEN WORLDS COLLIDE: INDO-EUROPEANS AND PRE-INDO-EUROPEANS. Edited by John Greppin and T. L. Markey. Karoma Publishers, Inc., 1990.

"Cambridge SoundWorks May Have The Best Value In The World. A Winner."

David Clark, *Audio* Magazine, Sept. '89

Henry Kloss, creator of the dominant speaker models of the '50s (Acoustic Research), '60s (KLH), and '70s (Advent), brings you Ensemble, a genuinely new kind of speaker system for the '90s, available only factory direct from Cambridge SoundWorks.

Cambridge SoundWorks has created Ensemble,™ a speaker system that provides the sound once reserved for large, costly speakers. It virtually disappears in your room. And because we market it directly, Ensemble costs hundreds less than it would in stores.

Your living room works *with* Ensemble, not against it.

No matter how well a speaker performs, at home your room takes over. Putting other speakers where the room helps the bass may hinder the upper ranges, or vice-versa. Ensemble consists of four units: *two* bass units and two high frequency satellites. The compact woofer units produce the bass that normally requires large speakers. Place them behind furniture, on bookshelves or under a couch. The small satellites blend into *any* decor. The result: a full range, musically accurate speaker system without big boxes.

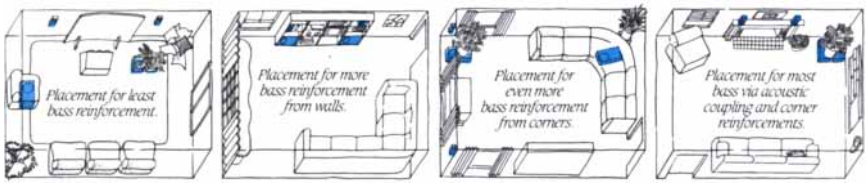
Unlike seemingly similar satellite systems which use a single large subwoofer, Ensemble uses two separate, compact bass units. They fit more gracefully into your living environment, and help minimize the effects of the listening room's standing waves.

Ensemble is a Trademark of Cambridge SoundWorks, Inc. Reprinted with the permission of Audio Magazine. © 1989 Diamandis Communications.

Try Ensemble in your home risk-free for 30 days.

Ensemble is sold factory direct...by helpful, experienced audio experts. So you can audition Ensemble the *right* way—in your home for 30 days, with no risk, no salesman hovering nearby. At $499* it is *the* value on today's speaker market. For literature, expert advice, or to order, call 1-800-252-4434 9 AM to Midnight, (ET), seven days a week.

Canada 1-800-525-4434 Fax: 617-332-9229.



You can put Ensemble's low-frequency units exactly where they should go for superb bass. You can't do this with conventional speakers because you have to be concerned about the upper frequencies coming from the same enclosures as the low ones.

CAMBRIDGE SOUNDWORKS
Suite 105M, 154 California St., Newton, MA 02158

☐ Send more information and test reports.
☐ Send Ensemble risk-free for 30 days, for $499.*
☐ Send an Ensemble Gift Certificate for $499.*
I'm paying by ☐ Check ☐ MC ☐ Visa ☐ AmEx

Acct. Number_____ Exp._____
Signature_____
Name_____
Address_____
City_____ State_____ Zip_____
Phone (Area Code)_____ Number_____

FOR IMMEDIATE SERVICE: 1-800-AKA-HIFI.
We ship worldwide, including APO & FPO.
MA residents add 5% sales tax.
*Plus freight ($7–$24). Delivery time usually 2–7 days.

Prepared as a public service by Ogilvy & Mather.

Will this become the only way to save the jaguar?

The choice is yours. Either save the photo above. Or, help World Wildlife Fund fight poachers, create nature reserves, hire guards, do research, and change things. Help us save life on earth.

World Wildlife Fund
Dept. A, 1250 24th St. NW, Washington, D.C. 20037

WWF®

MATHEMATICAL RECREATIONS

Lunar infants, lotteries and meteorites expose the dangers of math abuse

by A. K. Dewdney

"To follow foolish precedents, and wink/With both our eyes, is easier than to think."

—WILLIAM COWPER, *Tirocinium*

A nurse I know believes that more babies are born during the full moon than at any other time. She recently told me, "The moon was full last week, and we had double the normal number of births. It happens all the time." Someone else I know canceled a trip to Europe this past summer for fear of a terrorist hijacking but thinks nothing of commuting to work every day. An acquaintance who has played the stock market for two decades swears by a certain fund: "Any money manager who can consistently outperform the market eight years in a row gets my money!"

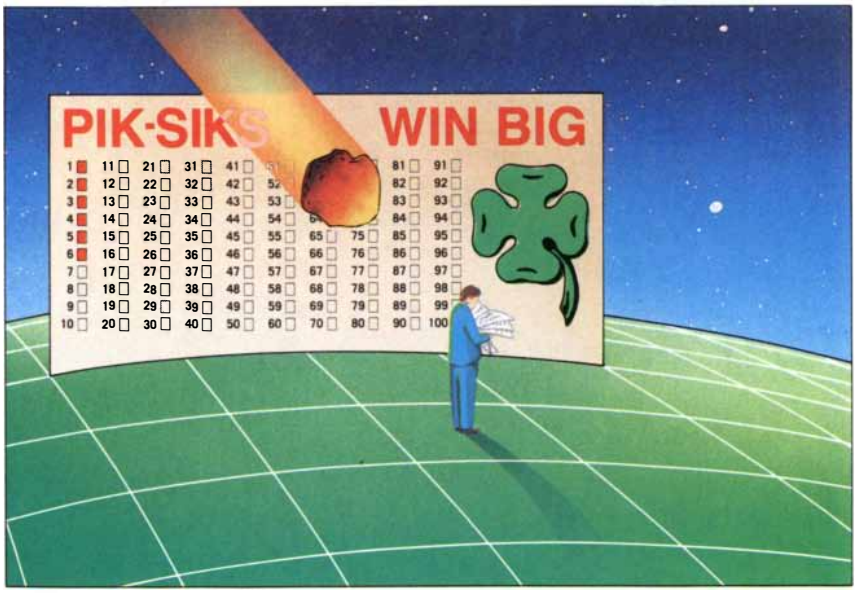
The foregoing examples illustrate various forms of math abuse: the inability or unwillingness to apply a simple logical analysis to certain situations that arise in everyday life. Many more examples can be found in *Innumeracy,* a book by John Allen Paulos. Someone who cannot deal with simple numerical ideas is innumerate, just as someone who cannot read or write is illiterate. I prefer the term math abuse because it has a wider scope. It includes errors that are not strictly numerical, and it also has a moral dimension. We abuse mathematics by failing to apply even the little we know of it to the false or questionable ideas that we encounter. We do not want to be duped, but most of us are fooled on a regular basis by politicians, the media and even friends.

The readers of this magazine, who are among the most numerate in the world, have a particular responsibility to recognize and combat math abuse wherever it may arise. What is wrong, for example, with the notion that more babies are born during the full moon? The idea certainly has charm, and I for one would not be upset in the slightest if it turned out to be true. Suppose for the moment, though, that it is not true—that as many babies are born on average during the full moon as during the new moon or the first quarter or the last quarter. If this were so, why might the nurse still claim that the peak baby-delivery period comes during the full moon?

Suppose that after watching the delivery of 15 babies in one day, the nurse looks out the window and sees a full moon. There it is again! A month later the maternity ward is relatively quiet, but the nurse does not bother to check the phase of the moon. If one watches only for the events that reinforce a belief, one is screening out all the events that falsify it. This phenomenon of belief is what Paulos calls a filter.

Filters can be found everywhere. Casinos that contain dozens of slot machines ring with the sound of winning. Every time the three little cherries line up, a machine disgorges a bunch of quarters that clatter into a tray. Losing makes no sound. Someone entering a casino may well be overwhelmed by the impression that everyone is winning. Yet even a few quarters won every 10 tries on average would produce a more or less continuous clatter from just 10 active machines.

The filtering phenomenon accounts for a good deal more than the "charitable casino" illusion. It can misguide us in our investment strategies. Was my friend right to place such confidence in a fund that beat the odds eight years in a row? Here is a simple way to judge the issue. First assume that the success of a fund depends on just plain luck—say, the flip of a coin. If the coin comes up heads one year, the fund will outperform the market index. In other words, the net value of all the stocks that make up the fund will increase faster than the index—a representative average of stocks. If the coin comes up tails, on the other hand,

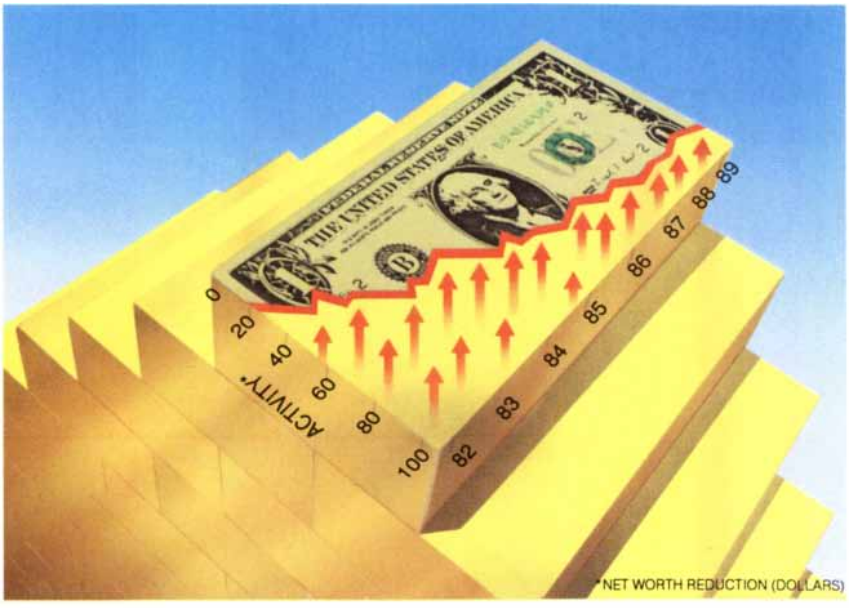


Is it more likely that I will win the lottery or that a meteorite will fall near me?

Under the new title of "Mathematical Recreations," this department will puzzle over numbers, reflect on logic and entertain computations of all sorts. But please do not shut down your thinking machine; computers will often be needed for mathematical adventures. "The Amateur Scientist" will appear in this space in alternate months.

the fund will drop below the index. Now if 1,024 funds were operating in 1982, consider their fate over eight years. By 1983 say half of the funds, 512, had outperformed the index. By 1984 half of these, 256, had again beaten the index. Each year the number of funds that continued to show superior returns was halved: 128, 64, 32, 16, 8, 4. The last number represents the "hot funds," those that produced unusually good profits every year from 1982 through 1989.

It was amusing to watch a television interview of the manager of a hot fund. To what did he attribute his unusual success? The manager carried on at some length about waves, cycles, bulls and bears. But the fact is that his success might have been just luck. A simple probabilistic model accounts for the hot-fund phenomenon quite well. In the face of this model, investors are well advised to view any and all claims with a certain degree of skepticism.



This diagram should be found guilty on at least four counts of chart abuse

An interesting stock-selling scam described by Paulos is based on a similar abuse of mathematics. A stockbroker sends out letters to 1,024 clients. In half of these letters the broker predicts that United Suspenders will increase in value over the next several days. In the other half he claims that the stock will drop in price. The broker now waits for the price of the stock to change. If the stock goes up, the broker sends 512 new letters to clients in the first group. He points out that his prediction was correct and then makes a second prediction. In 256 of the new letters he predicts the stock will continue to rise, and in the remaining 256 letters he predicts a fall in United Suspenders. The method is clear. An increasingly small group of clients is increasingly impressed by the broker's predictive powers. At some point this small group begins to bank on the stockbroker, when they should have left their money in savings bonds. Deliberate math abuse can be very difficult to detect.

Deliberate math abuse of a less insidious but more pervasive type is found in charts published publicly or privately. Two forms of chart abuse are possible. I call them compressive calming and stretch extortion. In compressive calming the vertical scale is compressed to tame the ups and downs of a variable. For example, to make a drop in sales seem far less precipitous, a company might plot its annual earnings on a graph that compresses the decline by a factor of seven. The company simply chooses a scale of $0 to $21 million even though its earnings have ranged from $1 to $4 million during the past 30 years. Conversely, a company whose sales are booming could use stretch extortion to make things look extraordinarily good; it stretches the scale so that the upward slope leads to the conviction that the company is indeed "taking off" [*see illustration on next page*].

Paulos gives a startling example of the dangers of innumeracy in the legal domain. "A man is downtown, he's mugged, and he claims the mugger was a black man. However, when the scene is reenacted many times under comparable lighting conditions by a court investigating the case, the victim correctly identifies the race of the assailant only 80 percent of the time. What is the probability his assailant was indeed black?" Most people would answer immediately, "Why, .8 of course."

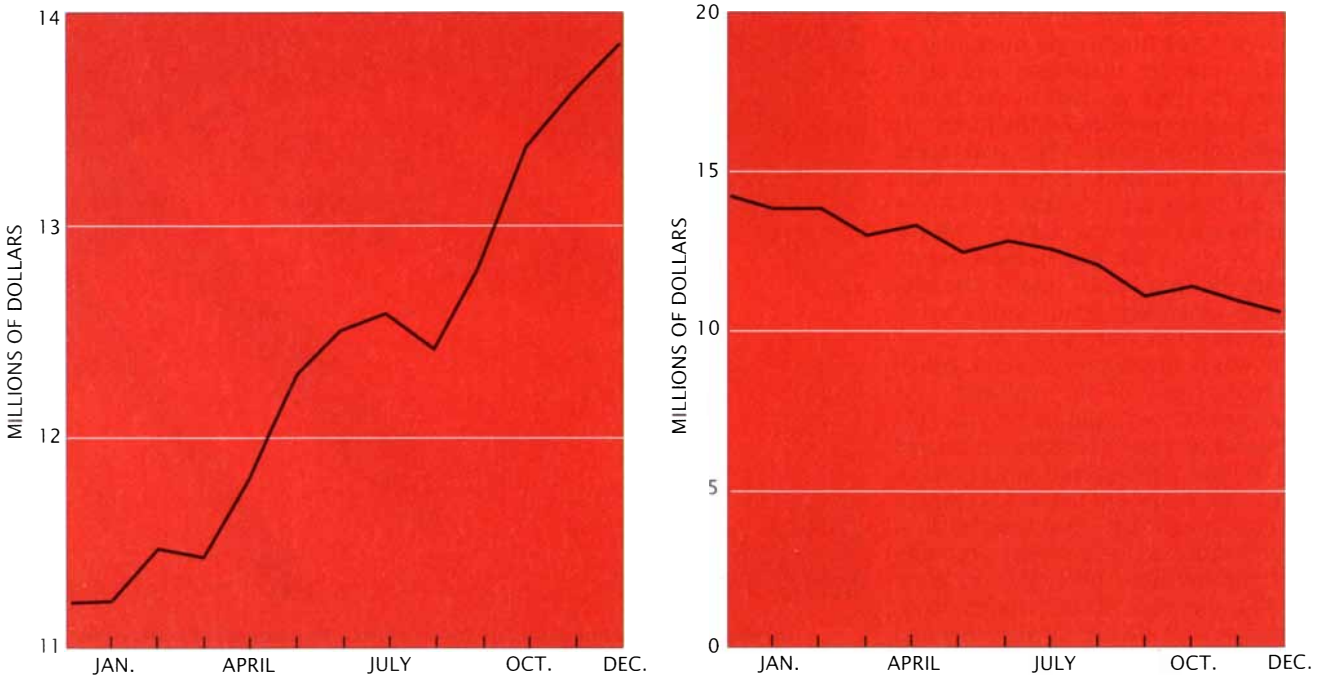
Paulos analyzes the problem, making some reasonable assumptions. First assume that the population in that area is 10 percent black and 90 percent white, next assume that members of either race are equally likely to mug people and finally assume that the victim is equally likely to misidentify race in either direction, white for black and black for white. Under these conditions, Paulos argues, "in a hundred muggings . . . the victim will on average identify twenty-six of the muggers as black—80 percent of the ten who actually were black, or eight, plus 20 percent of the ninety who were white, or eighteen, for a total of twenty-six. Thus, since only eight of the twenty-six identified as black were black, the probability that the victim actually was mugged by a black [assailant] given that he said he was is only 8/26, or approximately 31 percent!" A jury might react differently to this information than to the original crude estimate, reasonable doubt notwithstanding.

Lotteries make a wonderful introduction to probability and its abuses. What are my chances of winning a lottery? Consider the case of the six-number ticket. Players buy a ticket on which they select six numbers from 1 through 100. The point should be granted that one number series is as good as another: 1, 2, 3, 4, 5, 6 is as likely to be chosen as 6, 33, 45, 56, 69, 92.

The chance of winning follows from the simplest formula of probability:

$$\frac{\text{number of chances to win}}{\text{total number of outcomes}}.$$

If I purchase, say, five tickets and play a different sequence of numbers on each one, the upper part of this fraction becomes five. The lower part involves a bit of calculation: How many ways can one select six items from a hat containing 100 items? The first item picked can be one of 100 possibilities. The second item must be drawn from among the remaining 99 possibilities. The first two items are therefore selected from among 100 times 99 possibilities, or 9,900. In a lottery, of course, the numbers can be picked in any order. Since two num-



Charts can make good trends look better **(left)** ***and bad trends look not so bad*** **(right)**

bers can be ordered in two ways, the number of different lottery outcomes is 9,900 divided by two, or 4,950. But the calculation is not nearly finished. The third item must be picked from the 98 remaining in the hat. Yet one third of the resulting sequences will be redundant. The number of lottery outcomes now becomes 4,950 multiplied by 98 and then divided by three, giving the result 161,700. Obviously the bottom part of the fraction is growing rather quickly.

To calculate the total number of lottery outcomes, I must multiply this number by 97, then by 96 and finally by 95 and then divide the result by four, then by five and finally by six. The answer barely fits in my calculator: 1,192,052,400. If I buy one ticket, the number of chances to win is one. The chance of winning the lottery is therefore about one in one billion, and the probability is approximately .000000001!

My chance of winning the lottery is about two times smaller than the chance that the next meteorite to hit the earth will land within the square kilometer that has me at its center. In making such a statement, I am trying to combat another type of math abuse: number numbness. To appreciate a tiny (or extremely large) number, one must compare it with another number that is somehow more meaningful.

Take my example of a meteorite falling to the earth. The surface area of the earth is approximately 510,000,000 square kilometers. The chance of the meteorite falling within my particular square kilometer is therefore one in 510,000,000: the probability is about .000000002. Because this probability is about two times greater than the chance of winning the lottery with one ticket, I would have to purchase two tickets to even the odds. (Readers might have a much better chance of winning a hypothetical lottery described at the end of this article.)

Numbers large and small are the source of much confusion for most people, including me. We encounter large numbers every day, but frequently we forget our sense of scale. A favorite tactic of special-interest groups is to exploit this potential for math abuse. Consider the politician who rants, "Last year my opponent's party ran up a $3-million debt in its campaign fund. I ask you, can one expect fiscal responsibility from such a party?" Three million dollars sounds like an awful lot to most of us, especially if someone owes it. At first sight it stands right up there with the national debt, which currently stands at approximately $3 trillion.

How much is $3 trillion? One way to appreciate a figure like this, especially if it is a debt, is to distribute it among wage earners. Approximately 100 million people in the U.S. earn a salary of some kind. A debt of $3 million spread over this population would come to just over three cents per person. But a debt of $3 trillion would amount to $30,000 owed by every working soul in the country. (To take the analysis one step further, paying $30,000 at 8 percent interest would require yearly payments of about $4,470 over 10 years.)

I am equally concerned when groups espousing causes I believe in raise alarms that distort reality. This form of abuse might be called multiplier terrorism. Recently an antipollution group predicted a 100-fold increase in cancer risk in a neighborhood where someone had dumped dioxins. On the face of it, the prospects sound terrible. Is it not best to move out of such a neighborhood? The prediction may well be correct, but what is the actual risk? Assuming the probability of contracting cancer because of dioxin ingestion in a "clean" area is only .00001 to begin with, it would now be .001. How bad is this? The probability of contracting and dying from cancer from all sources is already .2 for the general population. Is the difference between .200 and .201 worth selling one's house for? The decision is obviously a personal one, but it might as well be an informed one.

Another form of math abuse is compound blindness. We willfully ignore the phenomenon known as compound growth. If a number is allowed to grow by proportional amounts in regular steps, the resulting numbers can be-

come large very quickly—far faster, in fact, than most people appreciate. For example, if someone sets aside $1,000 of hard-earned money in a special account that earns, say, 10 percent a year, and if the interest is reinvested at the same rate, the resulting sum can become quite large over a relatively short period. A sum invested at the rate of 10 percent will double in just seven years and three months.

Compound blindness applies not only to growth but also to shrinkage. Money shrinks through inflation. If the annual rate of inflation hovers around, say, 5 percent a year, money loses its purchasing power (the ultimate test of value) at a compound rate. In each succeeding year, one's money is worth 5 percent less than it was the year before. How long does it take money to lose half of its value? In the case of 5 percent inflation, it takes only 20 years.

One will frequently see advertisements from banks that exploit compound blindness: "If you had invested just $10,000 with us in 1970, you would now have $42,000!" True, but not as impressive as it first seems. At an annual inflation rate of 5 percent, the $10,000 invested in 1970 would have the same real value as $20,000 does now. Thus, in real terms the money has not quadrupled, but only doubled.

A modest ability to deal with probabilities, large numbers and other elementary mathematical ideas can go a long way toward eliminating math abuse in oneself. But it also requires the will to apply these ideas to arm oneself against the slings and arrows of sloppy thought.

Earlier I promised to describe a hypothetical lottery. The drawing is held, in effect, as soon as one buys a ticket. Along one end of the ticket, a number is printed, say 372981. Buying the ticket gives one the privilege of scratching away the silvery coating that covers the other end of the ticket. There, another number emerges, say 448372. Bad luck! Of course, the number *might* have been 372981. In such a case the prize is $1 million, tax-free!

Such a lottery could easily be engineered and, for all I know, could already exist. One end of the ticket is printed with one of a consecutive series of numbers, such as 244718, 244719, 244720. At the other end of the ticket, another six-digit number is printed, but it is chosen at random (from among those numbers not previously selected). Then the second number is painted over.

This scheme carries certain risks for both the buying public and the lottery corporation. For example, it might happen that no one wins, even if all the tickets are sold. What is the probability of this outcome? (Amazingly enough, it has almost no relation to the size of the lottery. Even if the lottery were restricted to single-digit numbers and only 10 tickets, the probability that no one would win would hardly be different!)

The risk for the lottery corporation, on the other hand, is that everyone could win—quite a serious problem! How many winners should one expect, on average? The formula can be represented as follows:

$$(0 \times P_0) + (1 \times P_1) + (2 \times P_2) + \ldots$$

What does this mean? P_0 is the probability that no one wins the lottery, P_1 is the probability that one person wins, P_2 is the probability that two people win and so on. If these probabilities are multiplied by the numbers 0, 1, 2..., respectively, and the products are all added up, the resulting sum gives the expected number of winners. Who would believe that the expected number of winners is exactly the same as if the winning number were drawn from a hat—namely, one? Two months should be plenty of time for readers to resolve the two puzzles before I publish the answers.

I would also be interested in any and all examples of math abuse noticed by our readers in their own lives. The best of these will also be published.

Computers around the world have been invaded by tur-mites, the cybernetic life-forms introduced in "Computer Recreations" this past September. Starting life in an infinite grid of black squares, each tur-mite walks about repainting squares according to its state-transition table. Each tur-mite is in reality a two-dimensional Turing machine, and the colors laid down on squares of the grid would ordinarily be regarded as the symbols that Turing machines are wont to read and write. Even the simplest one-state machines may produce elegant and complex patterns, as Greg Turk, the human subject of that column, showed.

Thomas Chrapkiewicz of Dearborn, Mich., has pointed out that watching the tur-mites move is far more exciting than the final patterns. Chrapkiewicz wrote a short program for the simplest tur-mite, the one that produces a random-looking cloud of colored squares and then heads off on a regular-looking line. The tur-mite's behavior made Chrapkiewicz curious about what it would do if it encountered a line already implanted in the grid. A tur-mite likes to follow lines and sometimes moves them in the process. This led to the idea of placing "snow" in the grid: randomly placed nonblack squares with which a tur-mite can interact. This will practically guarantee behavior that will never fall into an endless loop.

Tony Durham, a well-known English science writer, has programmed some of Turk's tur-mites, particularly the one that produces a peculiar box-like structure. I wondered if bilateral symmetry would arise in a tur-mite whose behavioral table has only cyclic symmetry. Durham has found a reversed version of this tur-mite that will undo whatever its unreversed colleague does.

From Hong Kong, Leonard Pratt writes about his tur-mite program that will accept any two-state table specified by the user. The program has enabled Pratt to follow a number of variations on the two-state theme. Changing just one table entry more often than not results in a completely different creature.

O. A. Olsen of Trondheim, Norway, has set two tur-mites in motion on the same grid to watch his charges mill about in a random cloud of colored pixels. Then one sets off toward infinity, tracing out a filigreed channel as it goes. Eventually its mate discovers the channel and "sets out along the path to tell him about his bad conduct."

Finally, I must mention the pretty pictures sent in by Homer C. Waits of Columbia, Va. Trying variations on the four-color tur-mite, Waits discovered one that produces an elongated blob from which it occasionally sends out finite "spiral arms." But the tur-mite always returns to the blob to make it even more complicated, elongated and random-looking before going off on yet another arm. With tur-mites, getting there is nine tenths of the fun.

FURTHER READING

LIES, DAMN LIES, AND STATISTICS: THE MANIPULATION OF PUBLIC OPINION IN AMERICA. Michael Wheeler. Liveright, 1976.

THE VISUAL DISPLAY OF QUANTITATIVE INFORMATION. Edward R. Tufte. Graphics Press, 1983.

INNUMERACY: MATHEMATICAL ILLITERACY AND ITS CONSEQUENCES. John Allen Paulos. Hill and Wang, 1988.

BOOKS

Health and wealth, electronics as art, oceans of air and of sage, elegant hulls

by Philip Morrison

HEALTH AND THE RISE OF CIVILIZATION, by Mark Nathan Cohen. Yale University Press, 1989 ($29.95).

One look at any current international list of vital statistics makes the case: the citizens of developed states, from Norway to Japan, enjoy an overall life expectancy at birth of well over threescore years and 10 and need mourn the death of only one infant in 100. Most human beings do not yet share those numbers, which measure the life-and-death value of a modern economy—at least while at nuclear peace—for its individual members (even though the averages conceal exclusions, as these United States grievously document). Everyone knows there is another, complementary parameter: the annual birth rate runs at about 15 births per 1,000 of population in those well-off lands.

An anthropologist at SUNY in Plattsburgh, the author has brewed up out of a varied and copious literature an absorbing study of the course of these measures of the human condition since earliest times. No trend in the entire historical and archaeological record is clearer than the dizzying growth in human numbers, from millions to billions, and in the size of human societies, with the foraging bands of a couple of dozen people replaced by superpowers that extend their writ across many time zones. That is what the vital numbers describe. But Professor Cohen puts a searching question. How far has what we are pleased to call the progress of civilization advanced the well-being of individual human beings themselves, weighed in this crudest of scales of life and death?

Three lines of evidence allow long-term assessment of human health, even though. The underlying adaptive strategies are complicated, political, filled with compromise and in constant change. First, we know enough now about the biological mechanisms of health and disease, epidemiology in the broadest sense, to construct limiting patterns for very different circumstances of life. Second, in the past few decades we have gained systematic vital statistics for isolated modern hunter-gatherers such as the !Kung San, the Hadza and the Pygmies, who are deemed to function at least approximately as people did within the long-gone world of hunters. The assumption of the pristine nature of such survivors has recently been put in doubt, but at a minimum they are expert simulations of ancient times. Third, the dead bones can now speak. Osteology and paleopathology make it possible to evaluate age at death and to diagnose disease from ancient human remains—witness more intimate than any other testimony of the spade. These difficult and fallible inferences gain credibility by the consistency of their patterns; together they "paint a surprising picture of the past . . . more complex than we commonly realize."

Infectious disease is an important problem for most plants and animals; we are all too clearly no exception. Larger and less mobile populations are more subject to epidemic infection. Malaria, for example, requires settlement because mosquitoes do not travel far and must bite and re-bite to spread the illness. Measles so surely confers lifetime immunity that it can persist only within a population large enough to nourish the virus on newborn susceptibility. Many diseases have probably arisen as their agent changed its host from animal to human after domestication had led to intimate contact: smallpox from cattle, influenza from pigs, the common cold from horses. There is an ironic twist. Once the large group has adapted, once ill children can count on long-immune parents to care for them, the fatal epidemics dwindle; but that very trait gives bigger civilizations a kind of unseen weapon against small groups newly in contact. The conquest of the New World—where there were no cat-

tle, pigs or horses—may have been largely a variant of that phenomenon.

Meat becomes scarcer and scarcer in this world; only the wealthy can now afford its energy inefficiencies on this crowded planet, along with the few who long ago hunted in grasslands crowded with bison, antelope or caribou. In Mesolithic times there was a shift from a narrow spectrum of foraging for the most desirable game to a broader spectrum of taking what one could get: Eden lost. The bones in the old cemeteries suggest a drop in stature at the same time. Tooth-surface patterns record episodes of nutritional stress; the data are not all concordant, but it appears that the early agriculturalists had found no marked relief from the foragers' episodes of seasonal hunger.

From somewhat uncertain data, the San and the Hadza, modern African hunter-gatherers, reckon a life expectancy of less than 30 years and an infant mortality of around 20 percent; their birth rates are moderate enough to lead to a near-stationary population. Ancient city civilizations did not greatly increase life expectancy, even though the number of people they supported grew by two orders of magnitude. Probably an increase in birth rates was matched by the agriculturalists' ability to feed more people.

The senate and citizens of Rome saw an average life expectancy of about 30; the study of medieval and later European cities leads to something much the same, falling to a minimum perhaps in the centuries just before the rise of modernity. That may be one reason for our complacent assessments of steady improvement. The simplest human societies a couple of hundred centuries back seem to have faced life and death with results that would "not have embarrassed most European countries" until well into the 19th century. The poor have not shared in the developments we celebrate—they do not yet share—and imagination has "assumed too close a fit between technological advances and progress for individual lives."

Professor Cohen has made a good case against the view that larger societies always improved living for their people. To be sure, the case is tendentious; his shrewd arguments must often go beyond what has been tested quantitatively. Was the introduction of cooking pots healthy or not? Pots can preserve bacteria; roasted meat is safer. Such logic can raise a hypothesis, but can it fix a rate?

The past 100 years or so, with all their tragedies, stand out as excep-

tional. The statistics today do not extend the long reign of stagnation. There is something new in the world, a manifest major improvement. The author addresses this change with the caution, entirely appropriate, that it has been seen over only a rather short run. Yet it does appear that technology, the skills of specialists, the general spread of better food, clothing, shelter, water, immunizations and know-how—all have at last brought to a billion humans or more a life span well beyond the prospects of those who knew the woolly mammoth. Those changes have not ended. Shall we not see some measure of gain spread among billions in less developed lands? The potential appears real; but in an unjust world the challenge of 20,000 years is still unmet.

THE ART OF ELECTRONICS, by Paul Horowitz and Winfield Hill. Second edition. Cambridge University Press, 1989 ($49.50).

The headlong development of electronics technology is a distinction of our epoch. This bulky, cheerful and hugely successful text and reference (now circulating in eight or nine languages) offers in its new second edition an unusual measure of change, one Nilometer in the digital flood. It is aimed not at the symbol-sorting computers of our banks, offices and homes but at a target that may be more congenial to this department's readers: the domain of the scientist and engineer in almost any discipline—one who seeks data and their meaning, and neither words nor dollar accounts.

The first edition was praised here in 1981 for its genuine depth, its good-humored, wry and casual style ("the very speech of the laboratory") and its match of the conceptual and practical at every level from testing to draftsmanship and safety. The pedagogy remains the same: making the most of approximate insights, it progresses to quantitative results only on the basis of such understanding. It is no book of theory but one of the concrete, although it expounds clearly the theory it needs. "By 'art' we meant the kind of mastery that comes from an intimate familiarity with real circuits, actual devices, and the like."

The chapters do span the art. The first three treat the fundamentals of circuitry in a low-power context; then discrete transistors, with models that enable practical design; and then a striking account of field-effect transistors, delicate and ubiquitous devices that respond to input currents in the picoampere range. The fourth chapter opens the sovereign idea of feedback, again not abstractly but in the context of operational amplifiers, an arsenal of clever analogue devices (the list of those available is a dozen pages long) that will amplify, oscillate, regulate or compare in one or another context. Every op-amp is a single silicon block, an integrated chip that in effect realizes a score or more of interconnected and well-fed transistors.

Digital electronics is the big topic. It earns nearly 400 pages; most of the 300 pages added in this edition are within this large section. Instructive overall but too detailed to master merely at a reading, these pages will be of high value to those who work at digital electronics for a term or two of a lab course, oscilloscope at the ready, or who consult them to augment the art they know.

The treatment begins with logic and its coding and ends in a complete detailed design example based on a central processing unit, a computer-on-a-chip. (The specimen chip is the Motorola 68000 series, viewed in the architectural style of IBM PC.) The result—in the end the well-instructed CPU requires nine more integrated chips in its support, one of them "with a zillion modes of operation"—is pretty much a small general-purpose computer, its task determined by its software. The task chosen here, a scheme to catch and average many repeated analogue signals, demands a program about three times as long as what was required for the six-channel event counter treated way back in 1980.

You might try to build an instrument like that on a circuit board. Wires will hardly do—too many; they are replaced by printed and plated conducting paths of thin metal. Perhaps the 10-chip averager could be built by handwork techniques. But if you wanted a board "paved wall to wall with fifty or a hundred IC's," you would need half a dozen distinct layers of circuitry. You could not be accurate enough, and moreover, "you'll need a month's vacation to recover from the concentration required to work out the routing."

The solution is computer-aided design. The logical rules of circuitry are set into a program; the output of the program is a layout of all the layers of foil and all the mounting holes with precise locations, printed legends and stencils that locate the dots of solder. A printed-circuit house will, for a condign fee, make up custom boards (by the million, if you wish).

Engineering is a culture; it is defined by the state of the art. Joseph Henry could not buy even insulated wire and had to wrap his magnet wire with dressmaker's scraps. To make a microcomputer, you need more ample support. Here are about 100 pages of up-to-date tables—doubled since 1980—that outline conventions and codes and list and annotate the wide range of component parts to be bought, from batteries and fans to the silicon arcana of buffers, latches and decoders—all the springs and gears of this new culture of functional intricacy, now reported at two moments in its precipitate growth.

SPACIOUS SKIES, by Richard Scorer and Arjen Verkaik. David & Charles Publishers, 1989. Distributed by Sterling Publishing Co., Inc. ($35).

Spectacular eye-filling images of a special sort of matter crowd this up-to-date field study. Just as invisible atoms make up all visible solids, so do barely visible but familiar droplets of water (or their icy progeny) make up a curiously protean substance: the clouds that throng the sky. Droplets in clouds, unlike real atoms, do not only move in and out of view; they freely appear and vanish in place to generate the shifting sky.

The cloud photographs that fill these large pages are of two kinds. In dramatic black-and-white we peer through the eyes of a dozen NOAA satellites, to encompass continents and oceans on a heroic scale. And in crimson, gold and blue we view clouds and sky as if through our own eyes, the horizon bounding the ground observer's view just tens of miles away.

The authors divide their dozen chapters according to scale and instrument. The two met first five years ago; they were already at one in a shared passionate pleasure and wonder at the skies of this planet. Arjen Verkaik had, with his wife, Jerrine, spent a decade and more chasing sky events all over North America to fix both beauty and meaning in brilliant sequences of sky photography. Richard Scorer, although a senior British meteorological theorist, had never been bound to armchair and computer. He began as a forecaster and became a glider-borne researcher amid the clouds and finally a specialist in atmospheric waves and clouds.

They have jointly compiled a book whose subject is visible change. It opens with brief, engaging introductions to air mass and front, stability and mixing, the taxonomy of clouds, satellite imaging and droplet formation. Most of both the text and the

series of exciting pictures to which it is keyed outline physical cause and form as clouds disclose change. Attentive to many scales at once, this candid treatment is no reader's royal road to quick knowledge, even though it is never hedged about either with algebra or with evasions. The authors hope to help our intuition grow jointly with experience, beginning with their striking samples and ample guidance. Readers already well versed in atmospheric flow are apt to see this volume as a classic display, not elsewhere matched, of what it is they know.

Some points are simple and compelling. The cloud that has a hard, bright edge is newly formed, for in it there are plenty of droplets, large and small, to make it opaque and reflecting. As it ages with drift outward into drier air, its smaller droplets evaporate, so that the less dense cloud appears thinner and grayer against the light—dimmed, as it were, in reflection. A variety of waves make remarkable isolated clouds smoother than any that arise from convective motion.

The contrails of aircraft, and the inverse "distrails" (engine-warmed downwash that evaporates holes in clouds just below the wake), are transient cloud forms, neatly shown and described. There are surprisingly long trails from ships, too—never in air mass newly come from land but out there in the eastern Atlantic, where the clean air is starved of droplet nuclei. Flow shadows, cast by islands across the stable airflows of cold climates, are nothing less than amazing: the orbital shots collected here of downwind cloud patterns—vortex streets and wavy wakes—from Jan Mayen Island and Crozet Island, where perhaps they provide skymarks for wandering albatross, rival any showpiece of the wind tunnels.

Here is a wonderful sequence from Ontario. A long line of cottony cumulus stretched across the sky on that cold October day, to deposit a pattern of oddly localized snow. The clouds could be traced to a single visible source: a lake, still warm, sent its vapors into the newly cold air mass to generate an icy snow plume and a long, drifting line of clouds. The parade was followed downwind, to yield photographs of cumulus "glid[ing] silently . . . like an endless procession of vapoury jewels" rising higher and higher. Again, the Verkaiks collected for us a big multicellular storm in north Texas; it is shown here both from the ground and from polar orbit. From above, the big anvil top was seen to spread into a giant egg that finally broke away to evaporate overnight above the warm waters of the Gulf.

Rotation marks the sky at all its scales. The tornado is well treated in the up-to-date context of the complex supercell storm, where adjoining updrafts and downdrafts make an organism of spins within which the thin tornado funnel forms, far from the simple model of a tornado as nothing more than one tightly spun vortex. That ominous tube was daringly followed by automobile in a couple of cases; a cabin-size dust devil offers a simple contrast. Ocean-size cyclones and anticyclones, swift jet streams flowing around them, evince the grandest patterns of the air ocean, laid out here in the four dimensions wherein this blue marble spins.

THE SAGEBRUSH OCEAN: A NATURAL HISTORY OF THE GREAT BASIN. Text and photographs by Stephen Trimble. University of Nevada Press, 1989 ($34.95).

The dryland ocean of sagebrush laps at two mountain walls. The Rocky Mountains mark its eastern shore in the ski-slope suburbs of Salt Lake City; to the west the sagebrush reaches the steep Sierra cliff front below Lake Tahoe. The wide sagebrush plains drained by the Snake River border it on the north, and its southern boundary is the warmer land of the creosote bush. (The southern edge is placed by Trimble well to the north of the tunneled and cratered mesas that swell and quiver to the buried explosions of nuclear weapons.) Many travelers cross the sagebrush sea on Interstate 80, often speeding by night to avoid 12 hours of sunstruck loneliness. This is among the "least novelized, least painted, least eulogized" of American landscapes, even though its flatlands, as level as if paved, were (in spite of daunting waterlessness) the wagonmasters' choice on the way west. They still hold deep-worn wagon ruts.

From salt flats to alpine tundra, writer-naturalist Trimble has camped there, has endured the harshest of seasonal changes, searched out the ancient bristlecone pines and photographed the tumbling tumbleweed. (The oldest living organism was a bristlecone pine cut in error in 1964 in the Snake Range of the Great Basin, more than 4,900 years of age by ring count.) His introduction to the ecology and biogeography of a strange ocean and its archipelago of mountains benefits from an unusual breadth of approach. This natural history is divided by living communities. It is no guidebook; only in an occasional allusion does the author mention specific landmarks of human use. A reader is drawn in by the subtle richness of the austere land, whose geology, trees, shrubs, birds and fishes are each the topic of another volume in this same well-illustrated series. This one takes the wider view of a celebrant.

The arbiter is geology, its decisions less than 10 million years old. The land rose then and tautened until its surface cracked. The fault blocks that formed are now mountain islands, sediments raised high above the sagebrush sea. The fire-born Sierras and Cascades at the basin's edge rose much higher, to cast their formidable rain shadow across the westerlies that bring Pacific moisture. The ocean winds lift, cool and rain themselves dry. Then they slide down to the basin, to rewarm into winds that make this desert of sunny skies and winter cold. Its scant waters do not reach the world ocean; they drain only toward the basin's interior.

The rainfall is close to the upper extreme for desert lands. Sunny Nevada averages less rain than any other state, nine inches a year, but that is far more than deserts like the barren Sahara or the nearly rainless Chilean coast. Evaporation is fierce, and life here must often adapt to a substrate turned too salty and alkaline for most forms. When the glaciers were last at climax, some 25,000 years ago, the evaporation was low, the summers cool. Two wide lakes comparable in depth and area to modern Lake Michigan then filled the basin 1,000 feet deep. The little mountain ranges were snow-covered and glacial: "This scene of ice white and lake blue is the unlikely stage for the desertification of North America."

A warmer time followed worldwide, and soon evaporation dominated. A residue of a few very salty lakes remains still, to nourish rafts of brine shrimp and clouds of brine flies. The wide flats between the relic lakes are the old lake bottoms gone dry, some clay-covered, some salt-crusted. Such a playa is even today a lake in the wet season, two inches deep and tens of miles wide, a shallow that bursts with strangely transient life adapted to the recurrent magic of spring.

A century ago the rancher and his cattle came at last to the sagebrush ocean (sagebrush now covers about 45 percent of the area of the Great Basin). Within a few decades half a million cattle grazed on its wide range. The perennial native grasses were adapted to make all their growth in the spring; they had no way to recuperate

from grazing. Only the sagebrush kept some winter leaves and could tap water deep below ground; also, it tasted bitter to livestock. In 40 years the grasses were no more, and cattle-raising vanished with the grass.

Big sagebrush (no kin to kitchen sage) then grew almost as a monoculture in a dry pool of dust all across Nevada. Plants introduced by chance thrived in that ecological vacuum; most of the tumbleweed of today's West is an imported Russian thistle. An interventionary drama of many acts has ensued, until now the whole intermountain range is again in use to raise livestock. The region (the Great Basin and much beyond it) is a "wilderness with cows," all closely managed under a technology based on Russian wheatgrasses resistant to grazing, on herbicides and on controlled fire, "juggling a billion acres worth of . . . realities . . . . We should wish ourselves luck."

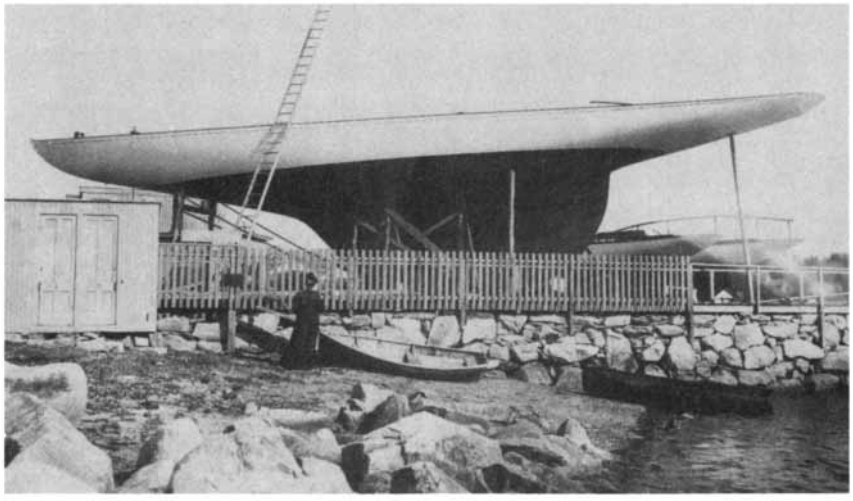
HERRESHOFF OF BRISTOL: A PHOTOGRAPHIC HISTORY OF AMERICA'S GREATEST YACHT AND BOAT BUILDERS, by Maynard Bray and Carlton Pinheiro. WoodenBoat Publications, 1989. Brooklin, Me. 04616 ($45).

In the years after the Civil War John Brown Herreshoff set up a small boatyard on his family's dockside property at Bristol Harbor off Narragansett Bay. JBH, although blind by the age of 15, was a designer and builder of stature, determined to race the fastest boats on the Bay; Nathanael Greene Herreshoff was his gifted younger brother, a young mechanical engineer from still younger M.I.T.

In 1878 they joined in an enterprise, the Herreshoff Mfg. Co., to build steamers of modest size on their own shoreland. Their novel entry was the Safety Coil Boiler just patented by their brother James, a chemist. It was wound from a single long, tapered iron tube and was costly to hand-forge and maintain but was more compact and able to raise steam much faster than any other boiler. A year or two before that they had, under contract from Cuban insurgents, designed and farmed out for construction the fast steamer *Estelle*. The U.S. revenue cutter set to guard the *Estelle* had to keep up steam even when at anchor for fear that the new ship would suddenly steam away. The U.S. government indeed seized the ship after her sea trials, but the brothers had charted a new course.

This fondly compiled and captioned collection of period photographs of boats, shops and people is the chronicle of the Herreshoff Mfg. Co. from the start until the razing of its shops after an Indian-summer revival during World War II to build naval patrol craft. It authenticates the powerful legend that grew around the brothers until the sudden death in 1915 of JBH, always the acute manager. Then the focus narrowed to independent Captain Nat, artist and wizard of design, until his retirement in the 1930's. Even under new owners during the lean years, the Herreshoff Mfg. Co. had its talented resident family designer, A. Sidney Herreshoff—"Mr. Sid."

The forms rivet even the casual glance and catch the knowledgeable eye even more surely. Here the elegant boats are seen at sea, at dockside, ready on the ways for launch or in embryo within sunlit shops. There is a wallful of the half-models Captain Nat favored, with his beautiful little dayboat hull designs in numbers. You see



The Herreshoff sloop* Altair *in winter storage, probably in 1902

a heap of the splendid yacht anchors he made on a new design. You can see the sunny machine shop, for NGH designed and the yard built steam engines as well as hulls, although never internal-combustion engines; those shafts and belts overhead are all powered by NGH's own steam engine.

Many little rowed craft (one lovely wooden tender went with each yacht, of course) speak quietly of the yard's sense of spare elegance and unstinted workmanship, manifest at small scale as at large. They built devotedly in cedar, mahogany, steel, bronze and aluminum at many sizes.

The pages record the sinister low black naval torpedo boats of 1895, the steam yacht for William Randolph Hearst, the steam-powered commuter boat for J. P. Morgan and the wonderful 46-foot sloop *Gloriana* of 1891, which took her every race that year to bring the yard the glory she was named for. The largest single-masted vessel ever sailed is seen under way: the America's Cup 1903 defender *Reliance,* too broad-beamed by rule to please Captain Nat completely. Her mast was a steel tube 189 feet high, and her racing crew was 66 men. Bronze plating over a steel frame all open within, her hull resembled an airplane fuselage. The last lofty craft were the beautiful aluminum alloy and steel Peacocks, as the local people called the four 76-foot J-boats, Cup defenders in the depression 1930's.

This is the record of a family institution, an industry and an art that served sport and war as equals, and most of all enjoyed a warlike sport. (Just one workaday boat was built, for the menhaden trade.) It is a tale of wealthy patrons, gifted architects and little-known master artisans. It is warming to read that the yard was paying the highest wages in New England before the turn of the century. A photograph documents the waterside spirit: workmen left their toolboxes open right where they ended their day's work. No one would disturb those tempting and costly kits.

The Mfg. Co. site in Bristol is now the site of newly built shore houses and of the Herreshoff Marine Museum; it is Museum scholarly energy that compiled this rewarding book from a wealth of family notes and photographs. The book is a pictorial narration that evokes in quieter key the architects and gardeners of the Sun King. In some ways this spirit survives; Herreshoff designs, and their successors rendered in fiberglass, still leave many a sporting wake in the waters of New England.

ESSAY

"Holographic" science to meet energy needs

by Richard F. Tucker

Some of us who manage scientific investigation are talking these days about a new approach. We call it holographic. It entails viewing any problem or area of investigation in depth (in three dimensions and at many different levels of detail) and from every possible angle, exploiting and uniting the varied perspectives of a host of disciplines.

This will surprise people who mistakenly think of the energy industry, and in particular the oil industry, as technologically mature and far removed from the cutting edge of scientific research. But the new approach is in fact a growing reality, driven by steady growth in society's energy demands, shifts in the energy-resource base and mounting concern about the integrity of the environment.

The oil industry is searching for breakthroughs at three frontiers. They are "georeaction engineering"—using the earth as a reactor for in situ processes to unlock more hydrocarbon resources; "catalytic templating"—a new way to manufacture precisely shaped molecules; and the technological fulfillment of the "environmental covenant"—the obligation we all share to future generations. These are holographic frontiers. Each one calls for sophisticated analytical and imaging techniques, massively parallel supercomputing and dynamic, multidimensional simulation at scales from the molecular to the global.

Let us look first at georeaction engineering. One of the petroleum industry's prime objectives is to secure an energy-resource base for the 21st century—to reverse the decline in economically recoverable oil reserves that is once again making America ever more dependent on foreign oil, with the political risks that go with such dependence.

The problem is not that the U.S. has run out of oil resources. North American reservoirs still hold a third of a trillion barrels of oil—enough, if we could recover it all, to take us into the 22nd century at today's consumption rates. The real problem is that even with current advanced reservoir-enhancement capabilities, the easier-to-produce oil in North America is almost gone. Most of what remains is oil that by its very nature resists extraction.

The challenge is to overcome the chemical and physical forces immobilizing the oil—to change the nature of the oil while it is in the earth and thus make it easier to extract. How? By treating the reservoir not as a mere underground storage tank but as a georeactor. Doing that will require quantum advances in our understanding of geothermal, geochemical and even biological reactions.

It will also require a holographic melding of geoscience and reaction engineering. Geoscientists need to think of the reservoir as a vessel of enormous scale and complexity, filled with beds of porous rock wrenched out of shape by immense tectonic forces, containing water, gases and the mixture of hydrocarbons that is called crude oil. Reaction engineers need to understand the complex changes that take place during in situ processing and how these affect the fluid dynamics of the entire reservoir.

We in the oil industry have been using imaging tools such as X-ray CAT scanners to observe complex fluids as they move through the pores of reservoir rock in the laboratory; we use acoustical tomography to make 3-D seismic snapshots of the reservoir itself and to track the chemical and physical changes taking place underground. Today, as advanced reservoir simulators push the limits of the emerging data-management and parallel-supercomputing technologies, dynamic holographic imaging of the georeactor is beginning to come within our grasp.

With supercomputers and 3-D imaging technology, the industry is also beginning to learn how to engineer molecules better. To engineer a molecule whose precise shape gives it a precise function (in a new fuel or lubricant, say, for the high-temperature, high-efficiency engines of the future), we must first design the precise shape-selective industrial catalyst for manufacturing that molecule. And this brings me to the second frontier: catalytic templating, the goal of which is to synthesize solid-state catalysts with an enzymelike ability to discriminate between molecules and to produce molecules that have highly specific functions.

Imagine such precision-designed catalysts working like chemical assembly lines, manufacturing molecular structures to order, carrying out multiple consecutive reactions at different sites—each site approaching the incredible selectivity of a biological enzyme designed by eons of evolution. To create such catalysts, we will have to understand more of the fundamentals: how the catalytic structures snap into place to form the template, how reactant molecules maneuver through a catalyst's microscopic pore system and how the product molecules can be extracted from the sites where they have been generated. To achieve catalytic templating, teams of surface chemists, organic chemists, crystallographers, applied mathematicians, materials engineers and chemical engineers will be required.

The most fundamental understanding of all will be needed for progress at the third frontier, the fulfillment of what I call the environmental covenant. The covenant recognizes that humankind, creator of the industrial system that has transformed civilization, is at the same time responsible for unintended environmental consequences of that industrialization. Maintaining the health of the planet's life-support system must be given the very highest priority. We all want a safe and healthy world; we must all become environmentalists.

Here the multidisciplinary approach is imperative if pollution is to be prevented at every possible source. This implies devising microscale solutions that have macroscale consequences: catalysts and processes that fine-tune manufacturing to minimize or eliminate unwanted by-products, computer modeling to optimize every manufacturing or production system, new molecular structures that pass harmlessly through the environment, and recycling of wastes. Moreover, prevention will have to evolve on a global scale. Today's energy industry is committed to ensuring that tomorrow's generations have not only an abundance of energy but also a healthy world in which to use that energy.

This environmental challenge—together with the world's changing resource base, society's needs and advances in science itself—will keep driving the industry to bring about dramatic change and perhaps even to conjure up a radically new slate of technologies within the next 25 years. Organizing science to meet the challenge will require a holographic approach—and a closer interdisciplinary partnership of scientists in industry, academia and government.

RICHARD F. TUCKER is president and chief operating officer of Mobil Oil Corporation and a director and vice-chairman of Mobil Corporation. He serves on several U.S. corporate boards and on Cornell University's Board of Trustees.

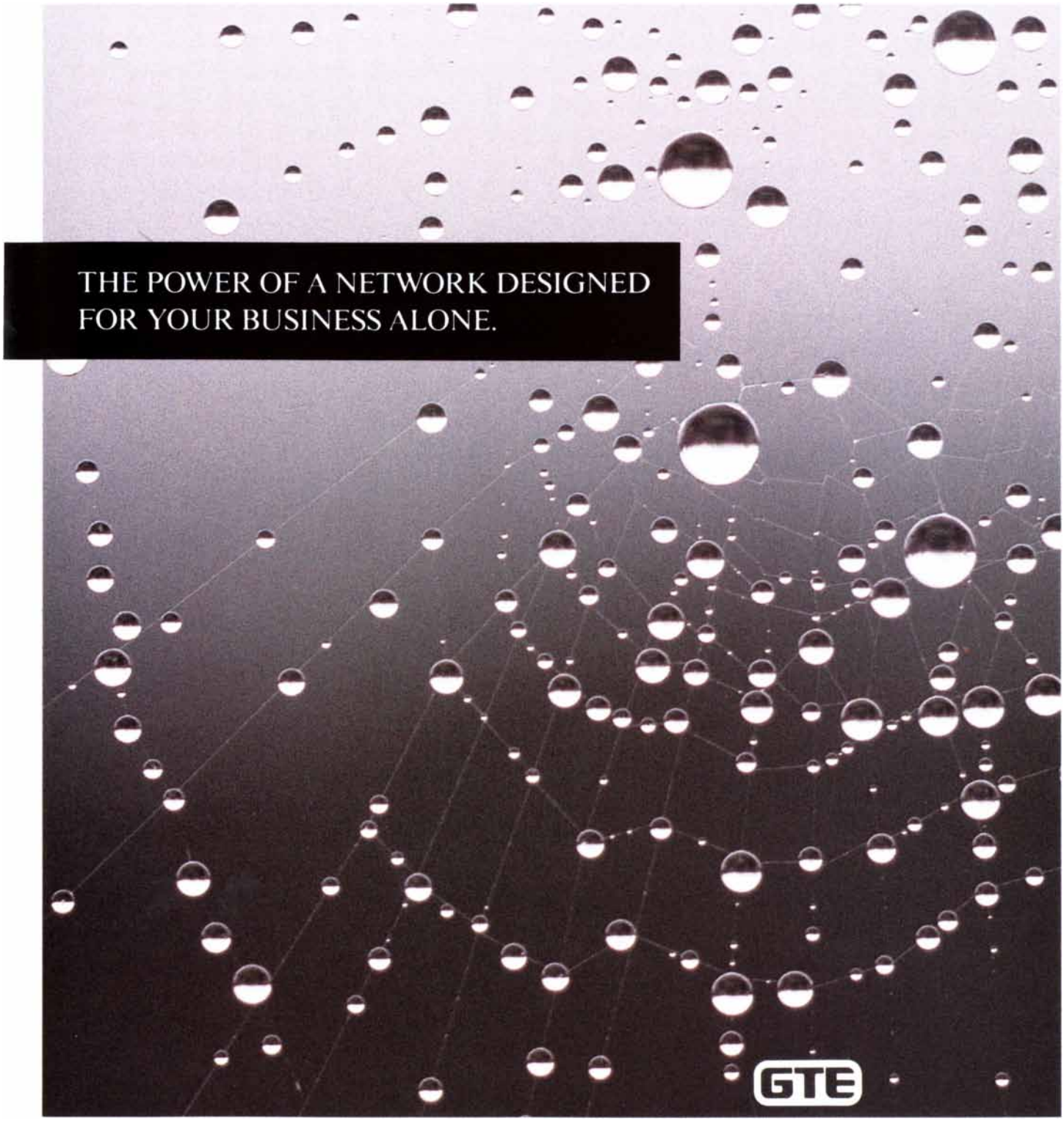
For the growth of your business, the prosperity of your business, there is no more powerful tool than the telecommunications network.

And at GTE, there is nothing vague about the power we put into your hands when we design that network for you.

For one of the nation's leading retailers, we are now implementing one of the largest satellite-based private data and video networks in the world. A network like no other because it was created to power the success strategies of that company alone.

The power of a personalized telecommunications network. We've given it to others. We can give it to you.

Because at GTE, the power is on.



THE POWER IS ON

ABSOLUT INTELLIGENCE.

FOR GIFT DELIVERY OF ABSOLUT® VODKA (EXCEPT WHERE PROHIBITED BY LAW) CALL 1-800-243-3787.
PRODUCT OF SWEDEN. 40 AND 50% ALC/VOL (80 AND 100 PROOF). 100% GRAIN NEUTRAL SPIRITS. ©1989 CARILLON IMPORTERS, LTD., TEANECK, NJ.