SCIENTIFIC AMERICAN

NOVEMBER 1991
$3.95

Microlasers—the tiniest points of light.

Clues to a cause of Alzheimer's disease.

More evidence for an African genesis.

Cosmic window: *the Keck Telescope's giant mirror reflects a revolution in optical technology.*

JOHN HOLTKAMP was spending more on gas than on film, which posed a problem for him.

John's a photographer. And he used to drive a truck. That big box he's leaning on is his bread and butter, and he takes it (and more) wherever the work happens to be. Which, on the average, means 750 miles from Monday morning to Friday night.

Hoping to pay less for those miles, John shopped the small imports. But they just wouldn't do. Then he heard about us. So he came by, with all his stuff, and looked at a Saturn sedan.

(Our 60/40 split fold-down rear seatbacks, while a real mouthful to say, make it easy when you have more than a trunkful to haul.)

On the weekends, John hauls animal feed. Because the Holtkamps have a llama, eight miniature donkeys, and an emu.

Since we didn't know what an emu was either, we asked, before meeting, if he wouldn't mind holding his emu while we took a picture of him and his Saturn. "Well," John said, "it's okay with me, but emus don't really like that sort of thing."

SATURN

A DIFFERENT KIND *of* COMPANY. A DIFFERENT KIND *of* CAR.

If you'd like to know more about Saturn, and our new sedans and coupe, please call us at 1-800-522-5000.

Attila t

So much power in so little space. The Quadra 900 is just 18.6" high and fits comfortably next to your desk. The Quadra 700 fits comfortably on top of it.

While the engineering and architectural disciplines have always prized the elegant solution, there are times when brute strength is imperative.

Introducing the Apple® Macintosh® Quadra™ 700 and Quadra 900 computers.

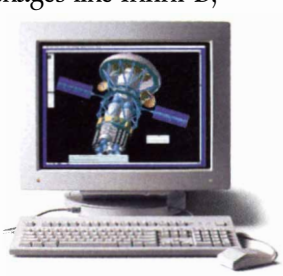
Awesomely powerful. Ferociously fast. But each is still very much a Macintosh.

Up to twice as swift as any of their forebears, they're the first Macintosh computers to be built around the Motorola 68040, rated at 20 MIPS and running at 25 MHz. A highly integrated design, the 040 combines the processor, math coprocessor, memory controller, and cache memory all onto one chip.

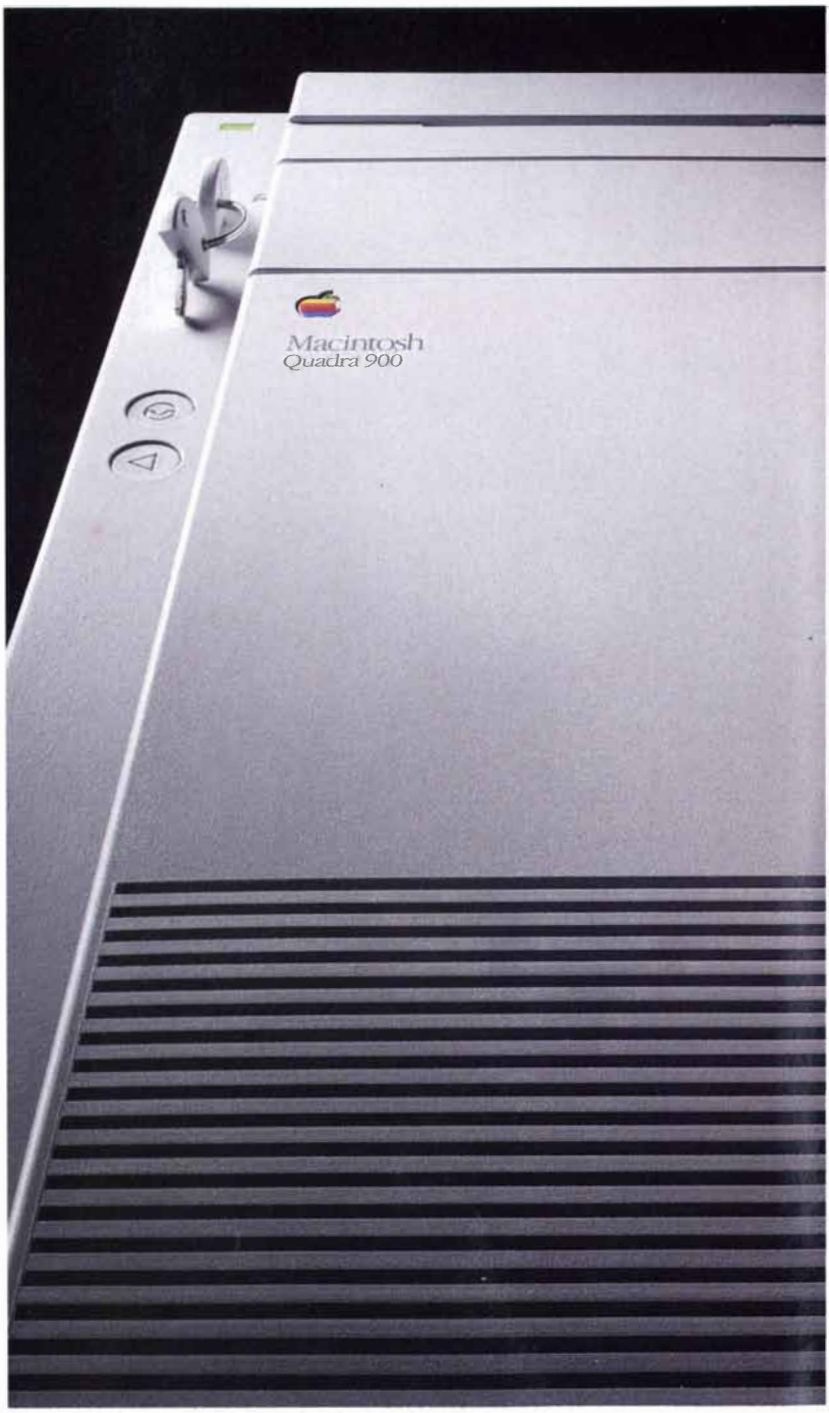
More important than merely technical measurements, the Macintosh Quadra computers are totally harmonized systems. The hardware architecture, operating system, interface, peripherals, and networking were all designed from the start to optimize the 040's power and work together smoothly as a single integrated system.

Anyone using compute-intensive applications — like 3-D modeling and stress analysis — will immediately appreciate the difference.

Popular software packages like Infini-D, MicroStation Mac, and Virtus WalkThrough perform more nimbly and



Big just got bigger. The new Macintosh 21" Color Display gives you more drawing board to work on. Colors are vivid, focus is crisp, brightness and contrast are high.



*24-bit video support for up to 16" monitors; 8-bit video support for up to 21" monitors. **There are two easy ways to do it. Simply add an application called SoftPC, or one of two cards from Orange Micro, the Mac286 or Orange386. © 1991 Apple Computer, Inc. Apple, the Virtus Corp. Camera 1 image was created in Infini-D; Infini-D is a trademark of Specular International, Ltd. Orbiter image was created with MicroStation Mac. MicroStation is a registered trademark of Bentley Systems Inc., an Intergraph affiliate. ORACLE is a registered Instruments. Lotus and 1-2-3 are registered trademarks of Lotus Development Corporation. SoftPC is a registered trademark of Insignia Solutions Inc. WordPerfect is a registered trademark of WordPerfect Corp. By the way, this ad was designed, typeset, and otherwise*

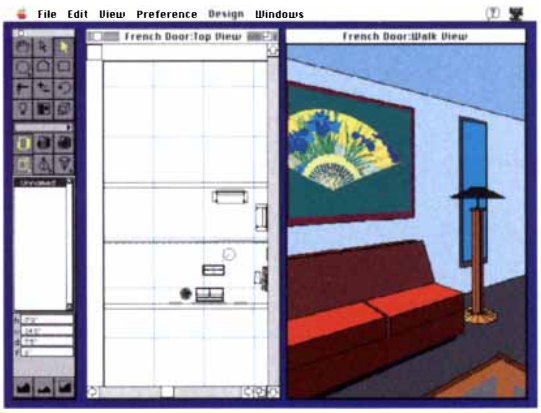
he Mac.

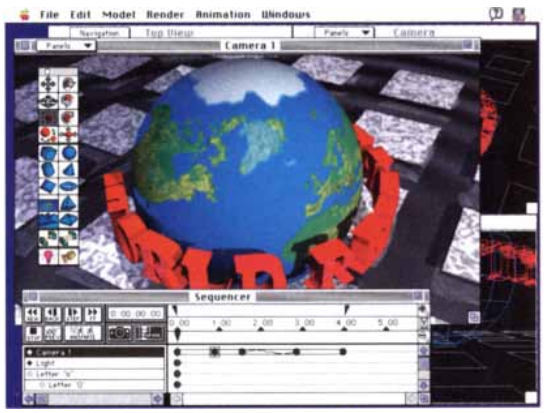
responsively than they ever did before.

And because you do more than design and engineering, these computers also run thousands of Macintosh productivity programs like Lotus 1-2-3 and WordPerfect. Accounting programs like Great Plains. Database programs like ORACLE and FoxBASE+/Mac. And presentation programs like PowerPoint.

In addition, RAM is expandable up to 20MB.

The Quadra 900 is a standing tower of immense capacity with five NuBus expansion slots, SuperDrive, plus three additional half-height expansion bays for CD-ROM drives, magneto-optical disk drives, tape backups, or hard disk storage of over 1 gigabyte. RAM can be added up to 64MB. It also features a key lock, not only





Because it's a Macintosh, extremely sophisticated programs for interior spatial emulation, 3-D modeling, and CAD/CAM are easy to use. Because it's a Macintosh Quadra, they've got the muscle to run nimbly and quickly. Pictured in action, Virtus WalkThrough and Infini-D.

Both Mac® Quadra models offer a generous array of expansion slots. Which you may never need since so much is already on board.

That includes sound input and output ports. And high-performance 24-bit color video controllers built onto the logic boards which will run any Apple and many third-party monitors.* Saving both a slot and the cost of a video card. And both come with Ethernet. So you can move large CAD files around the office at warp speed.

The Macintosh Quadra 700 is the same compact size as the popular Macintosh IIci.

It also comes with a SuperDrive™ disk drive, two NuBus™ slots, and a hard drive of up to 400MB.

for security, but to protect against interruption of your long, compute-intensive jobs.

Despite their unprecedented abilities, the Quadra computers are as easy to set up and use as every Macintosh and are capable of running not only thousands of Macintosh applications, but MS-DOS** programs as well.

For the name of your nearest authorized Apple reseller, call 800-538-9696, extension 320.

They'll be pleased to demonstrate a power that has vaulted once again to new heights. The power to be your best.®

Introducing Macintosh Quadra.

Apple logo, Mac, Macintosh, and "The power to be your best" are registered trademarks and Macintosh Quadra and SuperDrive are trademarks of Apple Computer, Inc. French Door screen was created with Virtus WalkThrough. Virtus WalkThrough is a trademark of trademark of Oracle Corporation. FoxBASE +/Mac is a registered trademark of Fox Software, Inc. Great Plains is a registered trademark of Great Plains Software, Inc. MS-DOS and PowerPoint are registered trademarks of Microsoft Corp. NuBus is a trademark of Texas created quickly and affordably using Macintosh personal computers.

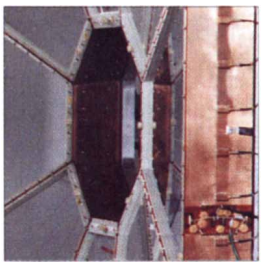
SCIENTIFIC AMERICAN

November 1991 Volume 265 Number 5

48 

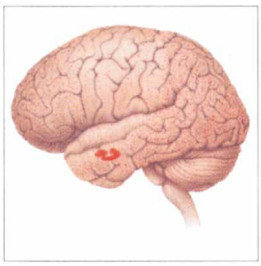
Firearms, Violence and Public Policy
Franklin E. Zimring

Guns are the weapon of choice in more than 60 percent of the homicides committed annually in the U.S. Handguns in circulation now number more than 35 million. But many state and federal gun-control laws, such as the 1991 Brady bill, may be misdirected. They do little to restrict access to those weapons that are most often implicated in violent crime and fatal accidents.

58 

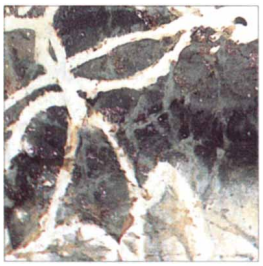
The Nuclear Equation of State
Hans Gutbrod and Horst Stöcker

The fiery death of an exploding star cannot be duplicated in any laboratory. Yet physicists know much about what happens to matter under such extreme conditions. Just as an equation can describe the states of water (solid, liquid, vapor) at various temperatures and pressures, so it can predict similar phases as matter breaks up into its ultimate constituents, quarks and gluons.

68 

Amyloid Protein and Alzheimer's Disease
Dennis J. Selkoe

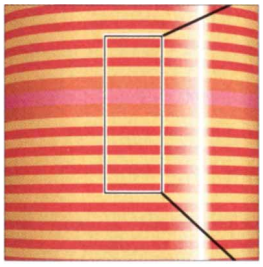
In 1907 Alois Alzheimer observed deposits of "a peculiar substance" in brains from patients who had suffered from senile dementia. Whether these amyloid plaques are a cause or an effect of Alzheimer's disease has long been debated. The author cites evidence that in at least some forms of the disease the protein is a causative agent. Understanding how the plaques form may lead to treatments.

80 

SCIENCE IN PICTURES

Chromoskedasic Painting
Dominic Man-Kit Lam and Bryant W. Rossiter

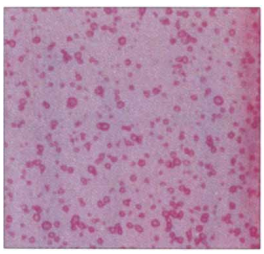
The well-known chemistry of black-and-white photography has an unexpected dimension. An artist has found a way to produce full-color paintings by controlling the size of the silver particles so that they scatter light in particular wavelengths.

86 

Microlasers
Jack L. Jewell, James P. Harbison and Axel Scherer

Compared with the transistors on a computer chip, the tiny red lasers in compact-disc players are clumsy behemoths. Researchers are closing the gap by making micron-scale lasers that promise eventually to be as minuscule as their electronic cousins. These efficient, low-power lasers may pave the way to optical computers and find applications from remote sensing to machine vision.

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. Canadian GST No. R 127387652. Subscription rates: one year $36 (outside U.S. and possessions add $11 per year for postage). Subscription inquiries: U.S. and Canada 800-333-1199; other 515-247-7631. Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537.

96 

Cultured Cells for the Treatment of Disease
Howard Green

Replacing severely damaged skin with grafts obtained by growing a patient's own cells in the laboratory is now easing the painful and protracted healing of burns. The cell culture technique is also curing intractable ulcerations and may one day be applied to other tissues and organs, including blood vessels and liver cells.

104 

Genes, Peoples and Languages
Luigi Luca Cavalli-Sforza

Evidence for the African genesis of humanity now correlates on three major fronts. Family trees based on an exhaustive analysis of human genetics trace the divergence of languages during successive waves of migration by ancient peoples. Both are supported by the archaeological record.

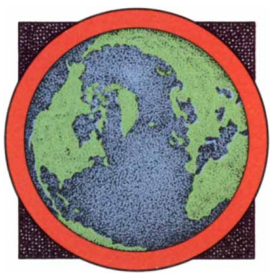
112 

TRENDS IN ASTRONOMY

Mirroring the Cosmos
Corey S. Powell, staff writer

When the Hale Telescope was completed in 1947, many astronomers believed it would be the largest ever built. But thick, rigid mirrors are being replaced by multifaceted reflecting surfaces, lightweight honeycombs and flexible sheets of glass. Combined with optical technologies that cancel out atmospheric distortion, the new telescopes promise the clearest and brightest view yet of the heavens.

DEPARTMENTS

17 

Science and the Citizen

The worrisome threats to Soviet nuclear safeguards.... Scanning the skies for maverick asteroids.... Checkmate in 223 moves.... The single-atom switch.... Are amphibians disappearing?... PROFILE: Superstring guru Edward Witten.

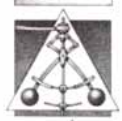
124 

Science and Business

Combating obesity with brain chemicals.... Computer-aided ergonomics .... Transdermal diagnosis.... The first new plastic in a decade.... Greening employment.... THE ANALYTICAL ECONOMIST: Insuring the unthinkable.

10 

Letters
The right way to draw a bow.... Pilot *vs.* computer.

14 

50 and 100 Years Ago
1941: Surgery through a tiny abdominal incision.

136 

The Amateur Scientist
Coaxing colors from black-and-white photographic emulsions.

139 

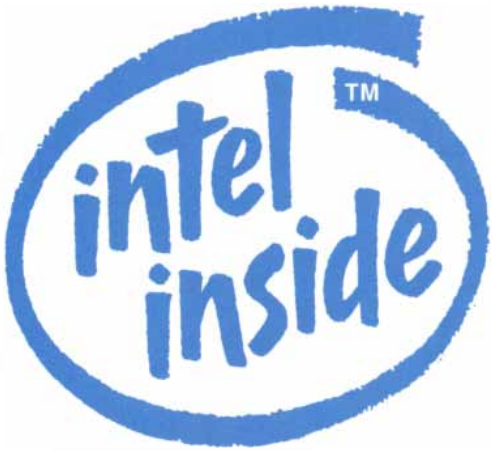
Books
London's docks.... Prehistoric Americans.... Illusions of India.

144 

Essay: *Donald R. Griffin*
Animals can tell us their thoughts—if we stop to listen.

How to spot the

very best PCs.



It's really quite easy. Just look for PCs that have a genuine Intel microprocessor inside. Either the Intel386,™ Intel386 SX, Intel386 SL, Intel486™ or Intel486 SX microprocessor.

Intel is the world's leader in microprocessor design and development. In fact, Intel introduced the very first microprocessor. So with Intel inside, you know you've got unquestioned compatibility and unparalleled quality. And you'll know you're getting the very best PC technology.

So look for the Intel Inside symbol on ads for leading PCs. Or call 800-548-4725 for more information. It'll show you've got an eye for spotting the best.

intel®
The Computer Inside.™

THE COVER photograph shows the elegant simplicity of the dome housing the Keck Telescope on Mauna Kea in Hawaii. When it is completed early next year, Keck will be the largest optical telescope in the world (see "Mirroring the Cosmos," by Corey S. Powell, page 112). Keck is part of a new generation of sophisticated telescopes that will incorporate giant mirrors and advanced electronics to provide remarkably sharp, bright views of the heavens. Early results from these devices hint at an astronomical bonanza soon to come.

THE ILLUSTRATIONS

Cover photograph © 1991 Roger Ressmeyer/Starlight

Page	Source
49	Kevin Horan/Picture Group, Inc.
50–51	Jason Küffer
53	Currier & Ives
54	Jason Küffer
58–59	Achim Zschau
60–64	Ian Worpole
66	NA35 Collaboration, CERN
69–70	Tomo Narashima
71	Dennis J. Selkoe
74–75	Tomo Narashima
76	George G. Glenner
80–85	Dominic Man-Kit Lam
87	Axel Scherer
88	Michael Goodman
92–93	Michael Goodman
94	Axel Scherer
97	Stanley Rowin
99–101	Carol Donner
102	William Hickerson, BioSurface Technology, Inc. (*left*), Carolyn C. Compton, Shriners Burns Institute (*right*)
104–105	Photo Researchers, Inc.
106–109	Laurie Grace
110	L. Tommaso Cavalli-Sforza
112–114	© 1991 Roger Ressmeyer/Starlight
115	© 1991 Roger Ressmeyer/Starlight; California Association for Research in Astronomy (*far right*)
116	Andrew Christie
118	© 1991 Roger Ressmeyer/Starlight
120	Andrew Christie
121	Tod R. Lauer, Kitt Peak National Observatory
122	© 1991 Roger Ressmeyer/Starlight (*top left*), European Southern Observatory (*top right*), Robert Q. Fugate, Phillips Laboratory (*bottom*)
123	© 1991 Roger Ressmeyer/Starlight (*left*), Robert D. McClure, Dominion Astrophysical Observatory (*right*)
136–137	Dominic Man-Kit Lam

SCIENTIFIC AMERICAN®

Established 1845

EDITOR: Jonathan Piel

BOARD OF EDITORS: Alan Hall, *Executive Editor;* Michelle Press, *Managing Editor;* Timothy M. Beardsley; Elizabeth Corcoran; Deborah Erickson; Marguerite Holloway; John Horgan, *Senior Writer;* Philip Morrison, *Book Editor;* Corey S. Powell; John Rennie; Philip E. Ross; Ricki L. Rusting; Russell Ruthen; Gary Stix; Paul Wallich; Philip M. Yam

ART: Joan Starwood, *Art Director;* Edward Bell, *Art Director, Graphics Systems;* Nisa Geller, *Photos;* Johnny Johnson

COPY: Maria-Christina Keller, *Copy Chief;* Nancy L. Freireich; Jonathan Goodman; Daniel C. Schlenoff

PRODUCTION: Richard Sasso, *Vice President Production & Distribution; Managers:* Carol Albert, *Prepress;* Tanya DeSilva, *Projects;* Carol Eisler, *Manufacturing & Distribution;* Carol Hansen, *Composition;* Madelyn Keyes, *Systems;* Leo J. Petruzzi, *Manufacturing & Makeup;* William Sherman, *Quality Assurance;* Carl Cherebin

CIRCULATION: Lorraine Leib Terlecki, *Circulation Director;* Cary Zel, *Circulation Manager;* Rosa Davis, *Fulfillment Manager;* Katherine Robold, *Assistant Business Manager*

ADVERTISING: Robert F. Gregory, *Advertising Director.* OFFICES: NEW YORK: Michelle Larsen, *Manager, New Business Development;* William Buchanan; Peter Fisch; Meryle Lowenthal; Elizabeth Ryan. CHICAGO: 333 N. Michigan Avenue, Chicago, IL 60601; Patrick Bachler, *Advertising Manager.* DETROIT: 3000 Town Center, Suite 1435, Southfield, MI 48075; Edward A. Bartley, *Detroit Manager;* William F. Moore. WEST COAST: 1554 S. Sepulveda Blvd., Suite 212, Los Angeles, CA 90025; Kate Dobson, *Advertising Manager;* Joan Berend, Lianne Bloomer, San Francisco. CANADA: Fenn Company, Inc. DALLAS: Griffith Group

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

INTERNATIONAL: EUROPE: GWP International, Düsseldorf; Roy Edwards, GWP International, London; Kurt Bühlmann, GWP International, Zürich. SEOUL: Biscom, Inc. TOKYO: Nikkei International Ltd. SPECIAL PROJECTS: Barth David Schwartz, *Director*

ADMINISTRATION: John J. Moeling, Jr., *Publisher;* Marie M. Beaumonte, *Business Manager*

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

PRESIDENT AND CHIEF EXECUTIVE OFFICER: Claus-Gerhard Firchow

CORPORATE OFFICERS: *Executive Vice President and Chief Financial Officer,* R. Vincent Barger; *Senior Vice President,* Linda Chaput; *Vice Presidents:* Jonathan Piel, John J. Moeling, Jr.

CHAIRMAN OF THE BOARD: Dr. Pierre Gerckens

CHAIRMAN EMERITUS: Gerard Piel

SOMETIMES WHAT YOU WEAR TO WORK MAKES ALL THE DIFFERENCE.

At 1,500 degrees Fahrenheit, dressing for success is no cliche. It's a matter of life and death.

Which is why Du Pont engineers worked to create Nomex® and Kevlar® fire resistant fibers. The remarkable performance of these fibers has made them the standard for state-of-the-art firefighting clothing all over the world.

And everyday, they not only save the lives of thousands of firefighters rushing into fires, but those they carry out as well.

At Du Pont, our dedication to quality makes the things that make a difference.

DU PONT

BETTER THINGS FOR BETTER LIVING.

LETTERS TO THE EDITORS

On Target

"Early Bow Design and Construction," by Edward McEwen, Robert L. Miller and Christopher A. Bergman [SCIENTIFIC AMERICAN, June], was enjoyable and educational, but it ignored a major component of the archery system of which the full recurve composite bow was one part: the thumb ring.

In the Western system exemplified by the English longbow, the bowstring is drawn with the curled index, middle and ring fingers. That technique is called the Mediterranean release. In contrast, the Oriental archery system uses the Mongolian thumb lock. The thumb wraps around the string, and the index finger locks over the top of the thumb to produce a very strong hold. A ring worn on the thumb protects the flesh and provides a low-friction surface over which the string can slide.

The omission of the Mongolian release technique led to a mistake in the description of the cover painting. The side of the bow from which the arrow shoots is dictated by the release, not by the engineering of the bow. The arrow must be fired from the side toward which the curling fingers or thumb points. If you curl your right thumb in the Mongolian thumb lock, it points right. [*The cover painting shows a Persian bow drawn by a right-handed archer; the arrow is on the right side of the bow, in keeping with the Mongolian release.—the Editors*]

In my experience, both release systems work well when used with either recurve or self-bows. Neither one works at all, however, with the arrow placed on the side opposite the release: the arrows spin end over end and fly off erratically.

LARRY PENLAND
Redding, Calif.

In Truth, Beauty

In his essay "The Poetry of Science" [SCIENTIFIC AMERICAN, July], John Timpane makes the kind of connections between science and poetry that were also made by great minds of our century. Albert Einstein has been quoted as saying, "The most beautiful and most profound emotion we can experience is the sensation of the mystical." In *The Immense Journey,* biologist Loren Eiseley wrote, "I have come to suspect that this long descent down the ladder of life, beautiful and instructive though it may be, will not lead us to the final secret.... It is only that somewhere among these seeds and beetle shells and abandoned grasshopper legs I find something that is not accounted for very clearly in the dissections to the ultimate virus or crystal or protein particle."

BOB MCCANN
Battle Ground, Wash.

Science and Abortion

Your readers deserve at least barebones evidence for John Horgan's article "Exporting Misery" ["Science and the Citizen," SCIENTIFIC AMERICAN, August]. To link the U.S. refusal to fund abortion consulting and rising Third World abortion-related death rates, Horgan leans on conjectures from Planned Parenthood, the World Watch Institute, lobbyists and politicians. Yet he doesn't quite make the case that the problem is getting worse. He does not, for example, show or even claim that deaths are rising faster than population. He does not discuss what was happening to the relevant statistics before the Mexico City policy or what other factors might affect them. He does not even specify what it is that is rising. Is it the number of deaths or perhaps the ratio of abortion-related maternal deaths to all maternal deaths?

Horgan also gratuitously slams conservative groups for discouraging researchers from contraceptive work. Let him weigh conservative pressure against the impact of liberal consumer group litigation on contraceptive research in this country. Is the French-made drug RU 486 unavailable in the U.S. because of conservative pressure or fear of U.S. liability laws?

BRIAN C. ZIV
Chicago, Ill.

It is appalling to see the U.S. government's hypocritical stance on human life and the downright irresponsible double standards that are the norm in U.S. policy-making. The policy of denying foreign aid to any organization that deals with abortion issues points to the frightening and abhorrent level of gullibility and ignorance that prevails thanks to self-serving groups such as religious fundamentalists and antiabortionists. Do these self-proclaimed guardians of "morality" know that the making of policies based on religious dogmas is a direct violation of the First Amendment?

Thanks to this "morality," thousands of children will continue to be born every day to live in poverty and misery and to suffer hunger, malnutrition, starvation and ignorance. Ignorance is what allows the symbiotic relationship between self-serving governments and the manipulative religious establishment to exist and to perpetuate each other.

LUIS A. PUYOL
San Francisco, Calif.

Plane Truth

Speaking as an airline passenger, I believe there is one overwhelming reason why I would prefer to have a human pilot rather than a computer in ultimate control of an airliner ["Along for the Ride?" by Gary Stix; SCIENTIFIC AMERICAN, July]. No matter how sophisticated the system, no matter how reliable, no matter how much quicker its reactions, a computer lacks an incentive to do the job right. A human pilot, on the other hand, has just as much riding on his level of performance as I do.

JEFF STERLING
Rockledge, Fla.

ERRATA

In "Antichaos and Adaptation," by Stuart A. Kauffman [SCIENTIFIC AMERICAN, August], the basis for the estimate of the number of human genes is misstated on pages 83–84. The 100,000 figure is based on the diversity of RNA sequences that are transcribed, not on the DNA content of the cells. Also, on the accompanying graph, the zero on the axis labeled "Number of Attractors and Cell Types" should be a one.

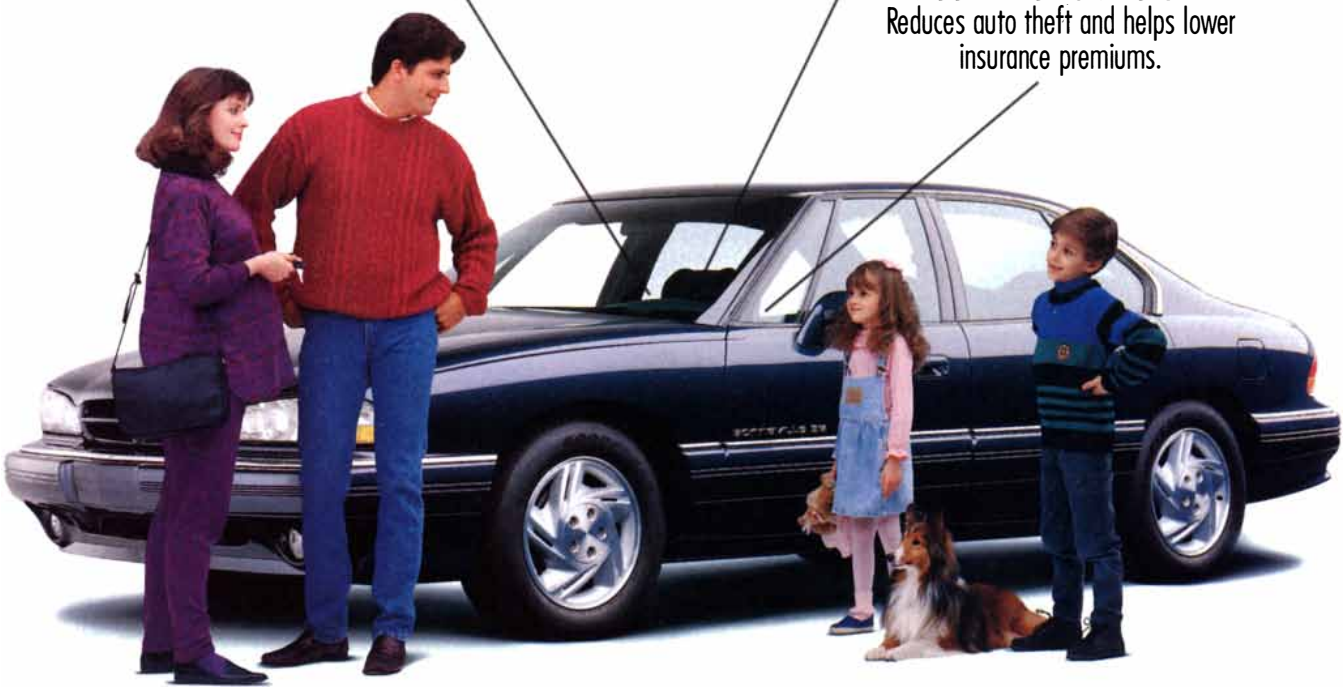
In "Mind Reader," by Timothy Beardsley ["Science and the Citizen," SCIENTIFIC AMERICAN, April], the primary affiliation of Rodney L. Lowman is incorrect. He is the director of Career & Personal Development Laboratories in Houston.

DELCO ELECTRONICS GIVES YOU A SAFER, MORE SECURE PLACE TO KEEP YOUR VALUABLES.

AIR BAGS
Of all the features we help provide, this is one we hope you'll never need.

HEAD-UP DISPLAY INSTRUMENTATION
Lets you see how fast you're going, while keeping your eyes on the road.

PASS-KEY® SECURITY SYSTEM
Reduces auto theft and helps lower insurance premiums.



At Delco Electronics, we don't just design electronic products for Geos, Chevrolets, Pontiacs, Oldsmobiles, Buicks, and Cadillacs.

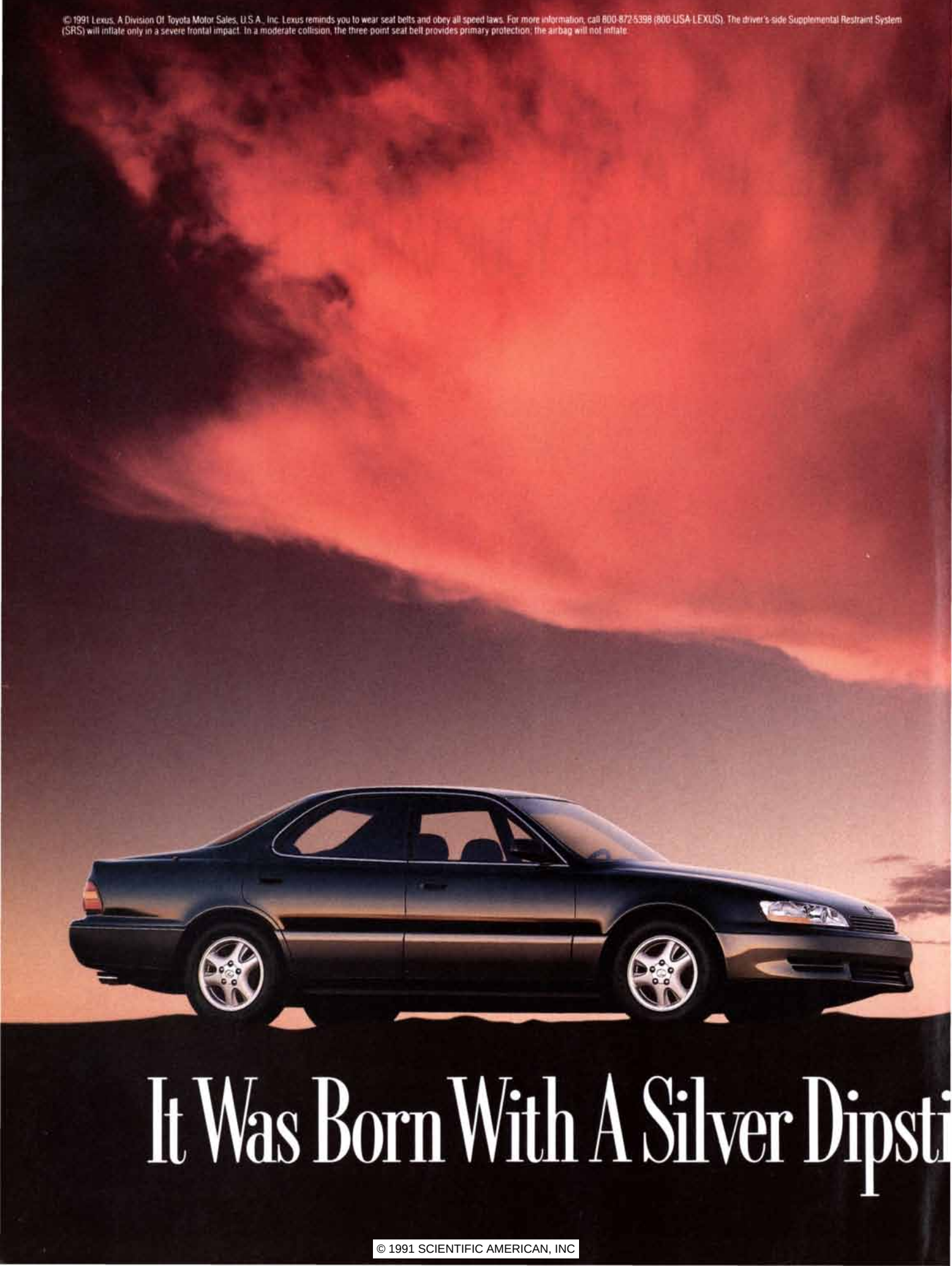
We like to think we're also helping to bring you peace of mind.

By making the family car a safer, more secure place for families. After all, we know just what kind of valuables you put into it.

Delco Electronics–driving the world in automotive technology.

Delco Electronics
WORLD LEADER IN AUTOMOTIVE ELECTRONICS™
Subsidiary of GM Hughes Electronics

It's who we are.



It Was Born With A Silver Dipsti

The ES300 Sports Sedan

Driver's-side airbag SRS

8-speaker, 180-watt audio

California Walnut trim

Anti-lock braking system

Automatic climate control

Remote-controlled entry

Sound-dampened cabin

LEXUS
The Relentless Pursuit Of Perfection.

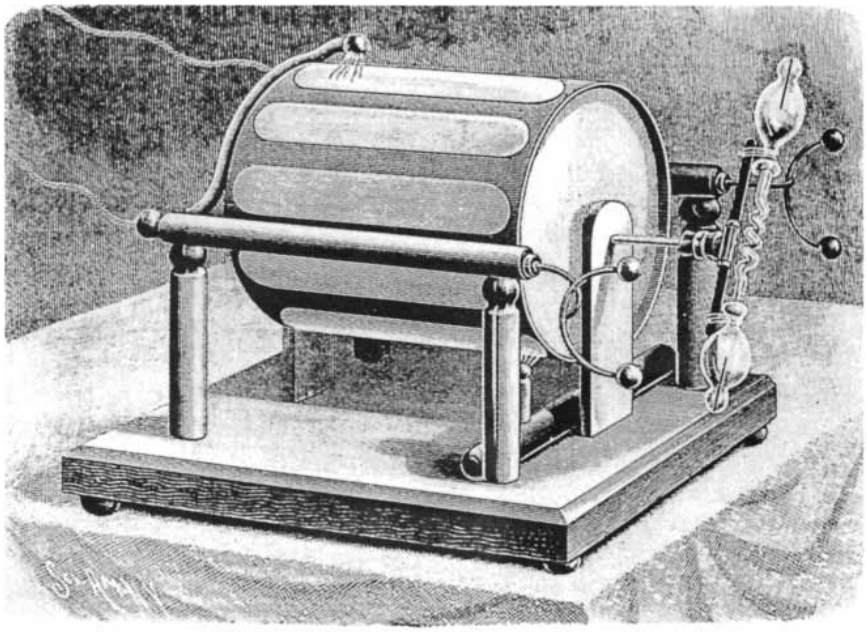
NOVEMBER 1941

"The pilots who fly the big bombers over the Atlantic to Britain sometimes stay at 15,000 feet, or higher, the whole way over. It's 50 degrees below zero up there but the heated planes are comfortable. Insidious, though, is the effect of altitude: at first you don't recognize the dreamy, don't-care feeling as the higher centers of the brain gradually cease functioning and you may wait too long before attaching the oxygen tube. One pilot, flying in winter at 20,000 feet to avoid icing, had to detach his tube and go back to help a passenger. When he returned to his seat he had to readjust his tube, a simple operation. But he couldn't do it. The tube in his hand approached the socket—wavered away. That went on for five minutes while they slipped down toward the dangerous icing level. Then the navigator realized what was wrong and came to the rescue."

"Dr. John G. Ruddock, Los Angeles surgeon, has perfected a device known as the Ruddock Peritoneoscope to study the organs within the abdominal cavity. First, a dull-pointed needle five inches long enables the surgeon to penetrate the abdomen and distend it with air. After withdrawal of the needle, a metal sheath containing a removable tip is inserted and a telescope is slid into the sheath. A tiny light is placed near the tip, ahead of the optical system. The surgeon simply scans his field through the tube and mirrors. Should tissues suspected of being diseased be encountered, the surgeon sends down a pair of tiny forceps, eases the open forceps into the tissue, closes them with thumb scissors, and removes the specimen. If bleeding takes place, he touches a switch, sending high-frequency electricity down through the lower end of the forceps to coagulate the blood."

"We have strong reasons to believe that the nuclear transformations by which the Sun's heat is maintained run so much faster with increased temperature that practically the whole energy-liberation occurs in the hottest and densest part of the Sun, near its center; the rest of the Sun acts only as a nearly opaque envelope, which keeps the heat from escaping to the surface faster than it does."

"Important among the developments which have made possible consistent broadcasting to far-distant countries is the beam or directional antenna. With this system it is possible to direct a fan-shaped beam of radio energy instead of broadcasting it to the four winds. The steerable antenna is of the greatest value in serving the areas centered around Rio de Janeiro and Buenos Aires. These two areas, as viewed from New York, are 20 degrees apart and are the most important language areas of South America."



An electrostatic motor

SCIENTIFIC AMERICAN

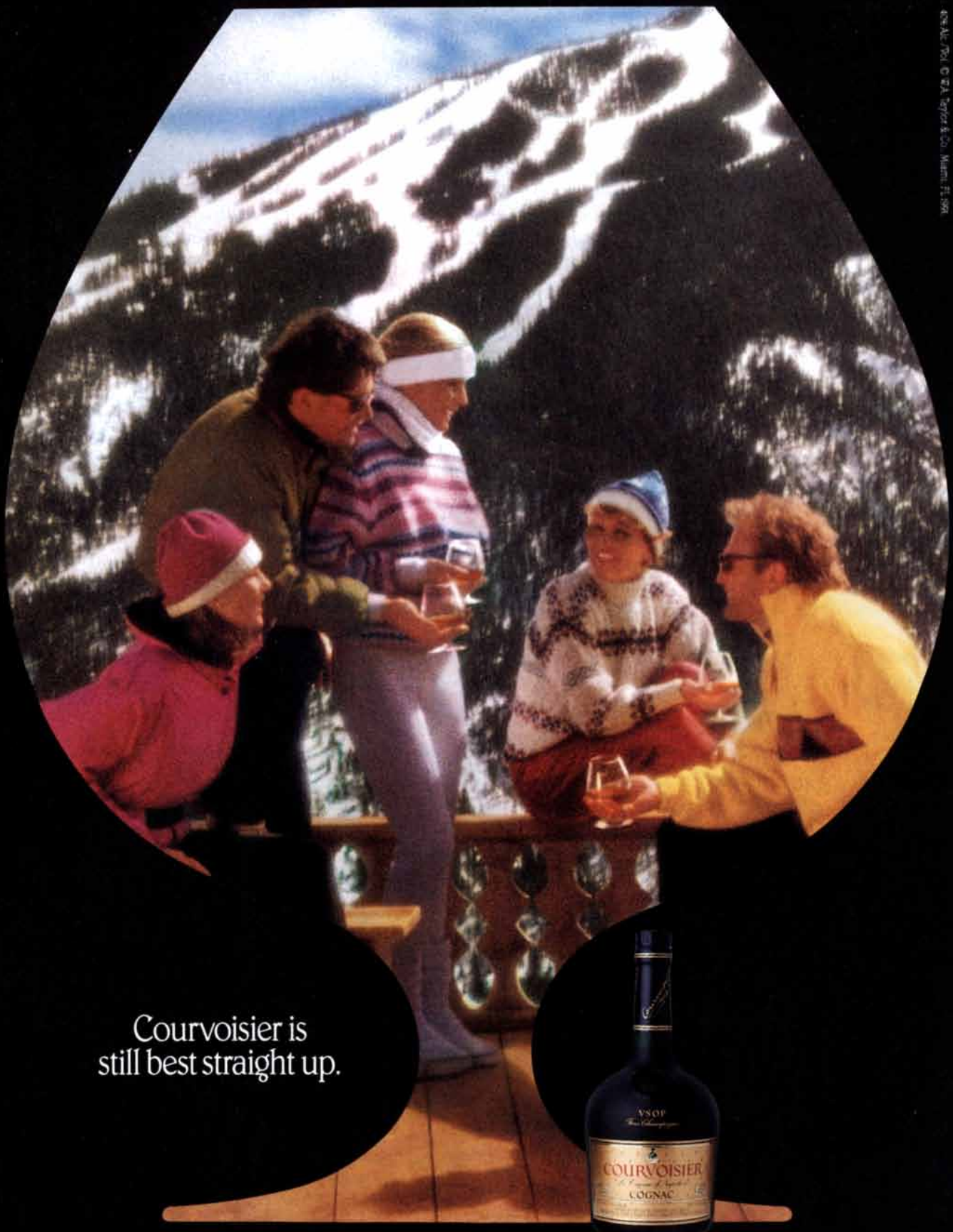
NOVEMBER 1891

"The Pittsburg *Leader* reports that the Hon. Henry Edwards, ex-member of the Pennsylvania Legislature, was seen running down the road minus hat, coat, and vest, and loudly calling for help. He was pursued by a mammoth hoop snake, which was running, or rather rolling, after him. The reptile had its tail in its mouth, and was rolling along hoop fashion.

"In the SCIENTIFIC AMERICAN for November 30, 1889, we gave an engraving and an interesting description of the hoop or milk snake. It will be seen from the information there given that the alleged rolling of the hoop snake is an optical illusion. The snake gathers itself up into large loops, and pushes itself forward, all with such amazing rapidity as to appear, to a frightened beholder, as if it actually rolled. The mind of man is very easily deceived by false impressions made through the eye."

"A curious instance of one poison killing another is reported from Yackandandah, Victoria, where Dr. Mueller has recently administered strychnine in cases of snake bite. The two poisons are antagonistic, and the characteristic effects of the strychnine only show themselves after the venom has been neutralized. The first independent action of the drug is evinced by slight muscular spasms and the injections must then be discontinued, unless after a time the snake poison reasserts itself. So long as the latter is active the strychnine can be applied in quantities which would be fatal in the absence of the virus."

"In the figure we have the static electromotor in its most efficient form. It does actual work, and is here shown rotating a Geissler tube, which is lighted by the same machine that furnished the power for the motor."



COURVOISIER®

Le Cognac de Napoleon

Precision-guided "Smart Weapons" played a vital role in Operation Desert Storm, many of which were designed and developed by Hughes Aircraft Company. These systems helped coalition forces identify enemy targets and defenses, and execute missions with pinpoint accuracy and minimum losses to troops. These weapons included airborne radar systems, tactical missiles, ground and naval systems, and infrared/laser systems for night vision and targeting. The lessons learned from Operation Desert Storm will help Hughes feed important information back into their development and production cycle to implement future improvements.

Dual Spectrum fire sensors will be installed on operational aircraft for the first time, as a Hughes-built system has qualified for use aboard the Air Force's C-5 Galaxy military transport. The proven Dual Spectrum fire sensor can detect an explosion or fire in an enclosed space so quickly, a fire suppressant can be released before any damage occurs. The units, located in the pylon and nacelle of each engine, can be connected to warning panels on the flight deck and nitrogen purging systems on the aircraft. The sensors, which offer tremendous safety benefits to aircraft personnel, have long been used on tanks and other military vehicles, and more recently aboard civilian passenger buses.

PC boards and hybrids can now be electronically trimmed and configured, thanks to a new family of nonvolatile, serially programmable (NSP) integrated circuits developed by Hughes. These NSP circuits enable designers to electronically calibrate PC boards and hybrids with test stations and computers. This automated procedure is a tremendous advantage over mechanical methods, which are less reliable and often difficult to perform. Presently, the new Hughes NSP family consists of nine types of devices. They all feature low-power consumption and redundant circuit techniques to ensure reliable operation and long life.

Corporations in Great Britain now will be able to transmit data via rooftop satellite dishes instead of using their current terrestrial-based telecommunications systems. It's the result of British Telecom's new satellite business network using Hughes' Very Small Aperture Terminal technology. The satellite service provides business customers with high-quality data circuits to potentially thousands of remote sites throughout the U.K. and Europe. Supported applications include database access and downloading, reservations, stock transactions, credit card verifications, electronic mail, data broadcast, and business television.

Hughes Environmental Systems, Inc. (HESI), needs environmental engineers, environmental planners, QA engineers, industrial hygienists, and managers at all levels. A subsidiary of Hughes Aircraft Company, HESI is rapidly becoming a full-service environmental management and remediation firm. We design and construct systems and implement programs in air quality/toxics, hazardous waste, wastewater treatment, and water supply. With a range of positions and facilities nationwide, we provide many challenging assignments and professional opportunities. If you're creative, ambitious, and have a BS/MS/Ph.D. in Environmental Engineering or a related field, send your resume to: HESI, Employment Department., Bldg. A-20, MS 2E264, P.O. Box 10034, Manhattan Beach, CA 90266-8534. U.S. Citizenship may be required. Equal opportunity employer.

For more information write to: P.O. Box 45068, Los Angeles, CA 90045-0068

HUGHES

Subsidiary of GM Hughes Electronics

Red Alert

Who's controlling the world's largest nuclear arsenal?

STILL A THREAT: Soviet missiles in Red Square. Photo: S. Salgado/Magnum.

The three days in August that Mikhail S. Gorbachev spent as a prisoner in the Crimea were the first time in history that nominal control of nuclear weapons had been seized by force. That possibility, long envisaged in war games, still came as a surprise when it actually happened. Now, with the plotters either dead or awaiting trial, U.S. nuclear strategists are trying to weigh the long-term implications of the events.

Strategists in the U.S. want to know how many separate nuclear powers from the former Soviet Union they will have to deal with (the hoped for answer: one). A related question is how many republics would have the power to veto a nuclear attack by a new union (the hoped for answer: as many as possible). Could Soviet nuclear weapons find their way into the international arms trade? And will the proposed new Union of Sovereign Republics maintain the strict controls that so far have prevented the unauthorized use of such weapons? Finally, could a new hard-line junta seize control?

In retrospect, U.S. nuclear weapon experts have downgraded the initial crisis to a nuclear nonevent. Although the coup plotters took charge of the briefcase containing codes that Gorbachev would use to launch a nuclear strike, Paul Bracken, a political scientist at Yale University, likens the "black bag" to the mace of medieval kings—a symbol of legitimacy that would have been of little use to the putschists.

High-level officers in the military and the KGB—both of which have key roles in authorizing the use of strategic nuclear weapons—remained loyal to Gorbachev. General Yuriy P. Maksimov, the commander of the Strategic Rocket Forces, went so far as to order mobile SS-25 missiles back to their bases during the coup, lowering attack readiness in order to reassure the U.S. And although U.S. sources disagree on the details of the command and control system, all concur that the plotters probably never had effective control.

To outsiders, the most obviously threatening weapons are the offensive strategic forces—the nuclear submarines, bombers and intercontinental ballistic missiles that can attack distant targets directly. The Soviet Union was believed to have about 11,000 offensive strategic nuclear warheads.

Although the great majority of the weapons are based in Russia, other republics have been host to a substantial proportion of the force. The Ukraine has been a base for SS-19s, older missiles now being eliminated under the Strategic Arms Reduction Treaty (START). It has also been a base for multiple-warhead SS-24s. Single-warhead mobile SS-25s have been based in Byelorussia and two fields of SS-18s, heavy missiles with 10 warheads apiece, are based in Kazakhstan.

Yet even before the collapse of the Soviet Union, antinuclear sentiment was growing in Kazakhstan and the Ukraine, according to Bruce G. Blair, a senior fellow at the Brookings Institution in Washington, D.C. Ukrainians have been resentful of all things nuclear since the Chernobyl disaster, and the Ukraine's declaration of sovereignty, issued in the summer of 1990, states its intention to become a nuclear-free zone. Nevertheless, Ukrainian officials have started to backpedal, apparently angling for a larger say in future nuclear decisions.

Kazakhstan's opposition to Soviet nuclear forces was as strong. The republic announced its intention to close the nuclear test site at Semipalatinsk two years ago and is now demanding reparations for health problems the site supposedly caused. Kazakhstan's antinuclear stance was tacitly recognized in START, which includes a provision allowing SS-18s from Kazakhstan to be rehoused in silos in Russia.

Yet Kazakhstan may also be having second thoughts about forsaking its weapons. The republic has a long-running border dispute with Russia and controls access to some SS-18 sites in Russia, which could tempt Kazakh president Nursultan A. Nazarbayev to use them for political leverage. Indeed, Nazarbayev has reportedly declared that Kazakhstan will not allow "any other entity" to have control of weapons on Kazakh territory. Even so, Blair thinks the Kazakhs as well as the Ukrainians will eventually realize that their political interests are not well served by striving to become independent nuclear republics. Byelorussia has also issued a nuclear-free declaration.

U.S. analysts seem confident there is

Strategic Missile Sites in the Soviet Union

SOURCE: Defense Intelligence Agency

almost no chance of an independence-minded republic seizing control of strategic weapons, whether they are returned to Russia or not. The weapons have sophisticated "permissive-action links" that prevent their use unless the correct series of codes are inserted. Nuclear bombers and submarines have similar checks.

The codes have to come through special channels from both political and military central commands, and further physical safeguards (such as sensors that trigger disabling devices) protect the integrity of the weapons. By Blair's account, given in testimony to Congress, the safeguards in the Soviet strategic control system are more impressive than their U.S. equivalents [see "Accidental Nuclear War," by Bruce G. Blair and Henry W. Kendall; SCIENTIFIC AMERICAN, December 1990]. "The Strategic Rocket Forces can dud their entire forces regionally or nationally," Bracken asserts.

Yet even if—as seems likely—the existing command and control infrastructure remains intact under a new central authority, it is unclear whose finger will be on the proverbial button. Russian president Boris N. Yeltsin has said a committee controls nuclear weapons but has also said Russia should control those on its own territory.

Russia will clearly dominate the new Union, but Ashton B. Carter of Harvard University points out that several other republics produce crucial components for nuclear weapons, such as arming and fusing mechanisms and special materials. They may be able to make a strong argument for having a voice in nuclear decisions—if they continue to cooperate in the weapon business at all.

More alarming than the strategic weapons, according to U.S. analysts, are the estimated 15,000 tactical and other short-range nuclear weapons. Until a few years ago, tactical weapons could probably be found in virtually every Soviet republic, according to Matthew Bunn, associate director of the Arms Control Association in Washington, D.C. In the past few years, ethnic strife in the Baltic States and the Transcaucasus has so alarmed the Soviet authorities that tactical weapons have been withdrawn from those regions.

Although many of the tactical weapons have permissive-action devices of some kind, they are thought to be less effective than the controls on strategic weapons. One reason is that, in contrast to the launch codes for strategic weapons, the codes to unlock tactical weapons are probably more widely distributed, says Edward L. Warner III, a nuclear weapon analyst at the Rand Corporation. But some tactical weapons have no physical controls, and some are merely "in containers with padlocks," according to Blair.

All of which points to a worrying scenario. If famine and ethnic conflict lead to civil strife, a mutinous military force or political grouping could conceivably commandeer a tactical weapon, Blair says. "No safeguard is foolproof forever—a fairly ingenious technical person could in time circumvent it." Some observers have even speculated that tactical nuclear weapons from the former Red Army might appear on the international arms market. "No Pentagon scenario of the past 30 years has anticipated this situation," Bracken states.

The best the U.S. can hope for to minimize the chances of such a diversion, Carter says, is for Soviet nuclear weapons to be confined to a small number of republics—probably those that now have strategic weapons. But he recognizes that whatever central authority emerges in the weeks to come might not want to risk the move, which would temporarily make the weapons more vulnerable to capture.

If the U.S. wants to prevent nuclear proliferation, Carter adds, it should be thinking about diversions from all the components of the Soviet nuclear production system, not just the finished weapons. That includes technical personnel. "There are going to be a whole bunch of unemployed bomb builders running round the Middle East if we are not careful," he says.

The changes of recent months could offer opportunities as well as threats. The new geopolitical landscape has convinced many weapon experts that the time is right to improve controls on all nuclear weapons. Rose Gottemoeller, a former State Department Soviet specialist now working for the Rand Corporation, points out that the Soviet troops in charge of nuclear weapons were an elite force. There is no guarantee that things would go so smoothly in the event of political turmoil in other nuclear states, she notes.

The National Academy of Sciences has recently added its voice to the chorus calling for change. The academy argues that the U.S. and the Soviets should agree in the near term to reduce the number of strategic warheads in their arsenals to between 3,000 and 4,000—a factor of three below the levels that will remain after START is implemented. Further reductions should follow. The academy also urges the U.S. to install permissive-action links on sea-based weapons and to press other nuclear powers to use similar controls on their weapons.

President Bush seized the initiative in late September with a bold set of proposals to eliminate U.S. tactical weapons, reduce alert levels and accelerate planned reductions in strategic weapons. The proposals suggest that despite its concerns about controls on nuclear weapons in the former U.S.S.R., the administration is confident that some form of strong central authority will persist. Among the topics Bush wants to discuss with Soviet leaders is the eventual elimination of missiles with multiple warheads—the most destabilizing of strategic weapons. In the end, Moscow's failed coup may have produced not a strategist's nightmare but a major step toward world security. —*Tim Beardsley*

"Nothing great was ever achieved without enthusiasm."

Ralph Waldo Emerson

The desire to break through boundaries. To make things better. It's the engine that continues to drive technology forward.

And in **Lebanon Tennessee** that translates to the production of over a million award winning color television sets every year.

Built by people who know that their opinion means more to the company than their quota.

And that the world's greatest achievements always begin with the belief in a better tomorrow.

Photograph by Pete Turner

Color TVs & VCRs • Computers & Printers • Copiers • CT Scanners • Digital Instrumentation • Disk Drives • Facsimile Systems • Industrial Motors & Controls • Liquid Crystal Displays • Medical Tubes • Mobile Electronics • MR Imaging • Nuclear Medicine • OPTO Electronic Devices • Picture Tubes • Professional Video Systems • Room Air Conditioners Satellite Receivers • Semiconductors • Telecommunication Systems • Turbines & Generators • Ultrasound • Uninterruptible Power Supply • Variable Speed Drives • X-Ray Systems

In Touch with Tomorrow

TOSHIBA

Atomic Turn-on

Researchers demonstrate the first atom switch

Although there are all manner of handy devices to control the flow of electric current, from the standard toggle to the Clapper, there is one model the local hardware store does not stock. This switch is flipped on or off by the motion of a single xenon atom. Such an atom switch may help pave the way for dramatic increases in computer speed and in the development of atomic-scale electronics.

The atom switch is yet another application for the scanning tunneling microscope (STM), an invention that won a share of the 1986 Nobel Prize for its inventors, Gerd K. Binnig and Heinrich Rohrer of the IBM Research Division in Zurich. Besides being capable of producing three-dimensional images of atoms, the device can pick up and move individual atoms.

The STM relies on a very fine tungsten needle. When the tip of the needle is brought close to the surface of a sample, electrons may "tunnel" between the tip and the sample, inducing a current that varies with the distance to the sample. To produce images, a feedback mechanism moves the stylus to maintain a constant current, thereby keeping the tip the same distance above any part of the surface. Sweeping the tip across the sample and using a computer to process its vertical motion creates a three-dimensional representation of the surface.

Researchers soon discovered that applying a voltage across the tunnel junction can cause an atom on the surface to stick to the tip. Donald M. Eigler and his colleagues at the IBM Almaden Research Division in San Jose, Calif., demonstrated they could drag atoms across a nickel surface. In famous stunts, they used xenon atoms to spell out their employer's initials and carbon monoxide molecules to draw a stick figure. Other investigators have used the STM tip to pick up single atoms and clumps of atoms.

Now Eigler and his co-workers Christopher P. Lutz and William E. Rudge have created an atom switch. By moving a single xenon atom between the microscope tip or a nickel surface, they were able to alter the amount of tunneling current between the tip and sample. The off position corresponds to the xenon resting on the surface. The researchers flipped the switch on by applying a 64-millisecond, 0.8-volt pulse to the tip, which made the xenon jump to the tip. The tunneling current then increased by about a factor of seven. Reversing the polarity of the applied voltage pulse made the xenon atom jump back to the surface, toggling the switch off.

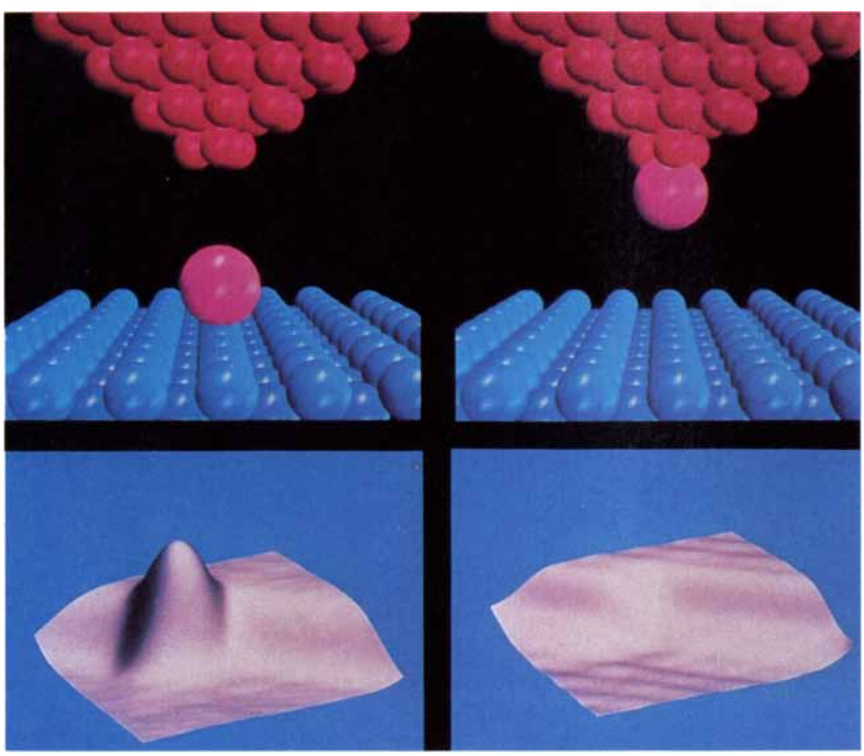
Why the pulse of current makes a xenon atom jump remains somewhat of a mystery. "The physics behind the transfer is not well understood," says Calvin F. Quate, a physicist at Stanford University. Eigler thinks the forces acting on the xenon may be similar to those that cause impurity atoms in certain solids to migrate. In that process, called electromigration, the flow of electrons causes impurity atoms to move through a conducting solid. The same effect may well cause the xenon atom to jump back and forth across the tunnel junction. The applied voltage scatters some electrons in the sample, and the scattering may help kick the xenon off the surface.

Immediate applications of the switch seem unlikely. The massive microscope and low-temperature, high-vacuum conditions make the switch in its present form "the epitome of impracticality," Eigler remarks. Still, there's hope. Although the STM is crucial in building the switch, it acts only as an electrical contact. In principle, then, a contact other than the tip of an STM needle is possible. In addition, other researchers have moved atoms around under less exacting conditions.

All this atomic manipulation "suggests we can build structures of nanometer dimensions," Quate says. Judging from the pace of development, he thinks some practical application from the control of atoms may emerge within five years.

The initial application would probably involve information storage. A computer bit, after all, simply indicates whether something is on or off. According to Quate, storing a bit in a cluster of 1,000 atoms—"which I think is very doable"—means that the entire contents of the Library of Congress could go on a 12-inch-wide silicon disk. Current technologies would require 250,000 such disks. Other possibilities that may stem from the manipulation of atoms include quantum "wires" and "dots," structures with promising electronic properties that confine electrons in one or zero dimensions. Investigators may also be able to assemble molecules atom by atom.

What other minuscule devices the switch may lead to depends on "being able to read the future," Eigler observes. He adds, "I haven't been trained in that area." —*Philip Yam*

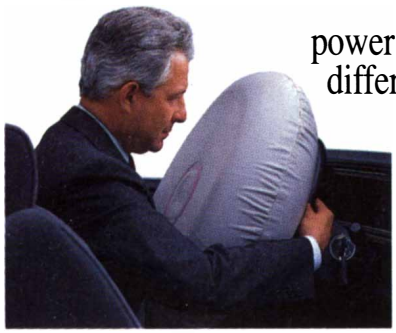


ATOM SWITCH components **(top diagrams)** ***include a scanning tunnel microscope (STM) tip*** **(red),** ***a nickel surface*** **(blue)** ***and a xenon atom*** **(pink).** ***The switch is off when the xenon sits on the nickel*** **(left)** ***and on when the atom is on the tip*** **(right).** ***The corresponding STM images*** **(bottom)** ***show the presence and absence of the xenon on the nickel surface. Magnification is about 15 million times.***

It can make driving easier, smarter, smoother and safer.

Relax. It's all in there. All the amenities you'd expect to find in a fine luxury car. Plus one feature you'd never expect. A very affordable price.

A driver's airbag is standard equipment on a Dodge Dynasty, for instance. A luxury no car should be without. Available anti-lock brakes will pay off down the road, too. With improved stopping power on wet and slippery surfaces.



A driver's airbag could be a real life saver.

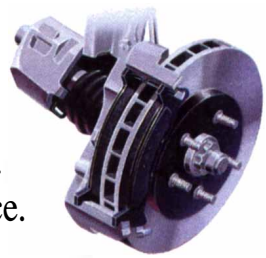
A 3.3L V-6 offers plenty of power for uphill roads and freeway ramps.

Under the hood, there's another powerful measure of security. An available 3.3L V-6, for more merging and passing power.

Position sensitive front suspension struts are another standard luxury. You'll love how quietly your Dynasty rides. And how well it smoothes out the bumps.

Then get a load of these smarts*: Doors that lock automatically when you reach 15 mph. A self-dimming rearview mirror. There's even a power driver's seat that'll remember two different positions.

Even smarter is our new Owner's Choice Protection Plan: 7 year/70,000 powertrain warranty or 3 year/36,000 mile bumper-to-bumper coverage, your choice. And with no deductible.†



Anti-lock brakes help you brake and steer when the going gets wet and slippery.

Dodge Dynasty gives you all the conveniences. Plus a price that makes it an outstanding value, as well. And that's the easy part.

All in all, it's quite a package.

Rediscover American Value.

Call 1-800-4-A-DODGE for a free product brochure.

Advantage: Dodge

OWNER'S CHOICE PROTECTION PLAN 7/70 POWERTRAIN OR 3/36 BUMPER-TO-BUMPER

BUY OR LEASE AT YOUR DODGE DEALER. Buckle up for safety.

*Dynasty LE options. †See limited warranties, restrictions & details at dealer. Excludes normal maintenance, adjustments & wear items.

OFFICIAL SPONSOR OF THE 1992 U.S. OLYMPIC TEAM USA

36 USC 380

Dodge Dynasty

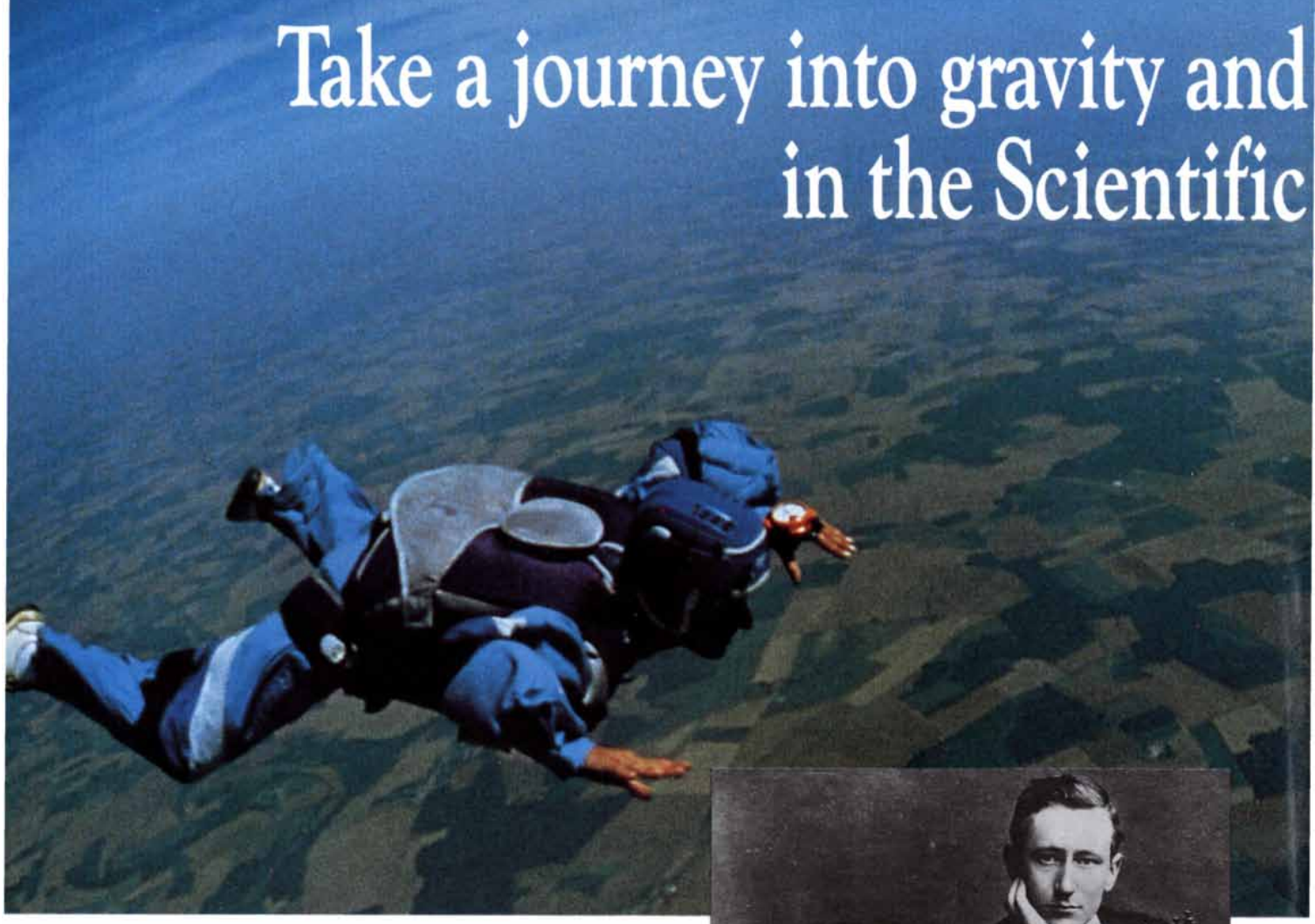
A substance abuse program at work can save employers money on insurance.

It costs more to replace an employee than it does to rehabilitate one. A substance abuse program at work may also reduce accidents and insurance premiums. Best of all, it can help you keep those people you wouldn't give up for anything. If you'd like us to send some helpful guidelines for setting up a substance abuse program at your company, write: Fighting Substance Abuse at Work, P.O. Box 303, Hartford, Connecticut 06141.

Aetna. A policy to do more.

Aetna

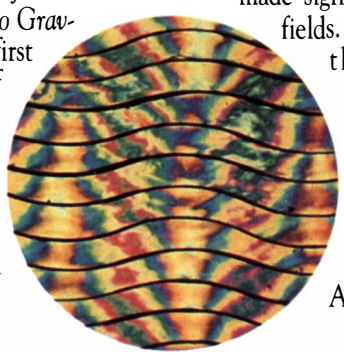
Take a journey into gravity and in the Scientific



Meet the giants who make science happen. (Guglielmo Marconi in 1896. Courtesy Marconi Company Ltd.)

Now you can preview *A Journey into Gravity and Spacetime* for 15 days risk-free when you join the Scientific American Library.

Gain fascinating insights into one of nature's fundamental forces. Discover the latest thinking about Einstein's general theory of relativity. Enjoy the special viewpoint of one of our most imaginative physicists. All when you join the SCIENTIFIC AMERICAN LIBRARY and receive John Archibald Wheeler's *A Journey into Gravity and Spacetime* – the first in a remarkable series of books that will take you on a wondrous journey to the frontiers of scientific discovery.



Enter the world of force and matter. (Polarized light view of stressed graphite fiber-epoxy resin composite. Courtesy Tsu-Wei Chou.)

Your invitation to share the ideas of today's leading scientists

SCIENTIFIC AMERICAN LIBRARY books are created in a unique collaboration between our editors and renowned scientists – men and women who have made significant contributions to their fields. The result is a series of books that are readable and *authoritative*.

Beautiful, informative volumes you'll treasure for years to come

Entertaining and informative, the SCIENTIFIC AMERICAN LIBRARY provides you with an excellent resource for study and research. Each volume features abundant photos, paintings, and illustrations, most in full color. Helpful diagrams, charts, and tables clarify points made in the text.

And SCIENTIFIC AMERICAN LIBRARY books are produced to exacting standards. Printed on high-quality paper and bound in cloth, they make a handsome addition to your library, providing you with years of pleasure.

Preview a new title every other month at no risk

Your premier SCIENTIFIC AMERICAN

spacetime as your first adventure
American Library.

Take pleasure in new ideas, delivered clearly, crisply, and in full-color illustrations.

Join the thousands of readers who have already discovered just how exciting science writing can be. Send for your risk-free premier volume today.

LIBRARY volume, *A Journey into Gravity and Spacetime* is just the beginning. Other volumes you'll receive include *Discovering Enzymes, From Quarks to the Cosmos,* and *Animal Navigation.*

As a member, you'll enjoy:

1. A new title every 2 months or so that you can preview for 15 days without obligation.

2. A low member's price of just $24.95 for each book you keep. You save 25% off the bookstore price.

3. No prepayment or minimum purchase. You pay only for the books you choose to keep.

4. *Our guarantee that you'll be completely satisfied or you may cancel at any time.*

To order, call toll free:

1-800-345-8112

SCIENTIFIC AMERICAN LIBRARY

Mail to: **SCIENTIFIC AMERICAN LIBRARY**
P.O. Box 646, Holmes, PA 19043

☐ Please enroll me in the Scientific American Library and send my first volume, *A Journey into Gravity and Spacetime.* Thereafter, send me a volume about every 2 months for 15 days' free examination.

Name_____

Address_____

City/State/Zip_____

1191S

Are Species Specious?

Biologists still argue about what a species is

For many biologists, a species is a bit like pornography: hard to define exactly, but they know it when they see it. Although Aristotle began categorizing animals into species more than 2,300 years ago, biologists are still far from unanimous about what the criteria should be. "Not only between disciplines but even within them, there are differences in how people try to define species," comments K. Elaine Hoagland, the executive director of the Association of Systematics Collections in Washington, D.C.

At the heart of the dispute is the most commonly taught definition, the biological species concept (BSC). It holds that a species is a population of organisms that can at least potentially breed with one another but that do not breed with other populations. That definition reflects the evolutionarily significant breeding patterns found in nature, explains Ernst Mayr of Harvard University, who has been the most ardent and persuasive advocate of the concept for more than 50 years.

Many other researchers are less enthusiastic about the BSC. "The vast majority of plant systematists don't use it. Entomologists don't use it. Most invertebrate zoologists don't use it," argues one of its critics, Joel Cracraft of the Medical School of the University of Illinois at Chicago. He and some other systematists are unsatisfied with the BSC because, they say, it disregards and obscures the phylogenetic (evolutionary) relationships between different species. Most biologists today, he says, "are basically using a phylogenetic approach without really calling it that."

Many taxonomists and conservationists find the BSC hard to apply in the wild. Subpopulations within a species are often unable to interbreed because they are geographically isolated from one another. Conversely, crossbreeding between organisms classified as different species is common among plants and far from rare in animals. "There's more genetic exchange than we once thought," Hoagland remarks. "I believe the biological species concept is the best we can deal with, but we have to realize that there are other things going on."

Mayr admits the BSC is not perfect but insists it is firmly based in biological reality. In populations of organisms that cannot interbreed freely, he says, biologists can recognize species by inferring prudently from their physical, behavioral and phylogenetic characteristics. "There are some people who consider themselves too stupid to ever make an inference," he says drily.

Arguments against the BSC based on natural hybrids, Mayr believes, show a misunderstanding of what a species is. Under the BSC, the species label describes a relationship between populations—not an intrinsic characteristic of individuals. "A population does not lose its species status when an individual belonging to it makes a mistake," Mayr writes in a forthcoming article for the *American Journal of Botany*.

Alternatives to the BSC have been devised. One controversial proposal is the phylogenetic species concept, which Cracraft put forward almost a decade ago. The phylogenetic concept defines a species as the smallest recognizable cluster of individuals that share a common trait and have a common pattern of ancestry. "You're applying the species name to those things that have differentiated, to those populations that are diagnosably distinct," Cracraft says. Critics contend, however, the idea would meaninglessly designate too many groups as species.

John C. Avise, a geneticist at the University of Georgia, is sympathetic to Cracraft's goals, but he believes the phylogenetic species concept goes overboard. In 1990 he proposed a new "genealogical concordance" approach: by comparing large numbers of gene sequences in various organisms, biologists can recognize the genetic drift that occurs only between distinct species. That definition maintains the BSC's emphasis on reproductively isolated groups, Avise explains, yet it also acknowledges the phylogenetic history of the organisms written in their genes.

The varied concerns of biologists seem to guarantee that, for the foreseeable future, no one will define species in a way that pleases everyone—although people will continue to try. "I'm very open-minded," Mayr says with a trace of irritation. "Let them come up with a definition, and then we'll see how good it is." —*John Rennie*

Dyslexia: A Problem of Timing

The first indication that a child has dyslexia is usually difficulty in learning to read. For most dyslexics, letters and words on a printed page seem to move or reverse themselves chaotically. A team of researchers at Harvard Medical School and Beth Israel Hospital in Boston believe they have found the problem: a sluggish brain pathway that cannot handle visual information fast enough.

Although dyslexia is most often thought of as a reading disorder, it is not limited to language—it is actually a broad category that can include problems with depth perception, spatial location of objects and other faculties. Some dyslexics, for example, have trouble differentiating similar sounds in speech, such as those of the letters "b" and "d."

Several studies have suggested that the underlying problem in dyslexia may be an inability to perceive or distinguish very rapidly changing stimuli. Margaret S. Livingstone of Harvard and her colleagues decided to test that hypothesis in the vision system. Within the human brain, some of the nerve fibers from the eye reach into two parallel systems: the magnocellular pathway and the parvocellular pathway. Whereas the magnocellular system responds primarily to fast-moving, low-contrast images, the parvocellular system reacts to slower, higher-contrast ones.

If dyslexia did involve a misperception of quickly changing stimuli, Livingstone's group hypothesized, perhaps the problem was in the magnocellular system. By monitoring the brain activity of dyslexic subjects, the investigators confirmed that their magnocellular pathway responded poorly to rapidly oscillating black-and-white patterns. Autopsies on the brains of dyslexics showed that although their parvocellular system was normal, cells in the magnocellular pathway were unusually small. Thinner neurons conduct signals more slowly, Livingstone points out.

Those differences, she says, may cause information from the magnocellular and parvocellular systems to arrive at higher brain centers out of sequence. For most activities, the mistiming would be unimportant, but for reading, which depends on extremely rapid eye movements, it may be disastrous. The brain areas associated with the sense of hearing may also be divided into fast and slow systems, which could explain how dyslexia affects speech perception, Livingstone adds. —*John Rennie*

Put the best seat in the house in your garage.

The Buick Park Avenue Ultra.

This is what you get when you tell your interior designers to spare no expense.

You get 6-way power front seats that place you instantly at ease. An optional memory feature even recalls your favorite seating position at the push of a button.

You get deeply cushioned rear seats, designed for maximum relaxation. Rear-seat passengers even have their own lighted vanity mirrors.

And of course, the Ultra's seating areas are covered with soft, supple leather.

So if you're looking to redecorate, we suggest you start by putting a 1992 Park Avenue Ultra in your garage.

BUICK®
The New Symbol For Quality In America.

GM® Buckle up, America!

Park Avenue is a registered trademark of GM Corp.

"It expanded my horizons, without exceeding my budget."

"With PRODIGY,® there's a panorama of information, services, advice, and entertainment at my command. Hundreds of exciting features, and all for just $12.95 a month, with no online time charges. That means I can use PRODIGY without worrying that the meter is running.

"The great graphics and friendly interface are real attention-getters. And the menus and cross-referencing features make it so easy, my whole family's involved.

"There's so much to tell you about PRODIGY, I hardly know where to start.

"We talk to other members. We all love the PRODIGY clubs where we exchange information about things we enjoy. The computer club is great because it keeps me up on the latest PC technology. My wife's favorite is the food club where she shares recipes and tips.

"I do more shopping and banking in a lot less time. I go window-shopping for ideas on PRODIGY and make purchases by keyboard instead of by car. And when the bills arrive, no problem. With PRODIGY I can pay them right from my PC. It would take all day to tell you about all the exciting things I do on PRODIGY.

"You gotta get this thing."

The PRODIGY service, connected to your home computer and phone line, has games, sports information, an encyclopedia, and much, much more. And, it's just $12.95 a month, including 30 personal messages.*

PRODIGY is available for IBM® or compatible, and Macintosh® computers. Call **1-800-776-3693, ext. 219,** to find out where you can purchase a PRODIGY Service Start-up Kit now. Or look for PRODIGY already included with many popular computers.

*Additional messages will be billed at only 25 cents each. Bank fees may apply. PRODIGY is a registered service mark and trademark of Prodigy Services Company. IBM is a registered trademark of IBM Corporation. Macintosh is a registered trademark of Apple Computer, Inc. (c) Copyright 1991 Prodigy Services Company.

Murder Mystery

Ecologists seek better data on extinction

The widely reported announcement in 1990 seemed dire: populations of amphibians around the world, some biologists said, had been declining ominously since the 1970s. Indeed, two high-profile species, the gastric-brooding frog of Australia and the golden toad of Costa Rica, now seem to be on the verge of extinction. As many as 1,500 golden toads used to gather at the species' principal breeding site, according to Alan Pounds, a researcher at the Monteverde Cloud Forest Preserve, but only individuals have been seen since 1987.

The reported declines have prompted biologists to suspect that amphibians, like canaries in coal mines, might be particularly sensitive to some subtle worldwide environmental change. The greenhouse effect, pesticides, acid precipitation and a new virus have all been proposed as possible explanations. But a 12-year study of a pond in South Carolina called Rainbow Bay, which has generated the most detailed data ever assembled on amphibian populations, points to a more prosaic effect: random fluctuations caused by drought.

The Rainbow Bay study tracked populations of three species of salamanders and one species of frog. Although the populations fluctuated dramatically, the species were more likely to breed successfully in years when the pond held water for longer before drying up. The fluctuations in population size were surprisingly large, but there were no unexplained trends once allowance had been made for rainfall.

The results, published in *Science,* constitute what the study's principal author, Joseph H. K. Pechmann of Duke University, calls "a cautionary tale." He says he has no doubt that destruction of natural habitats is driving some amphibians to the brink of extinction. But his research shows that natural fluctuations can be large and may be mistaken for a trend if there is no good long-term record. Although the species he observed were affected by drought, there is no evidence that droughts are more common in South Carolina now than they ever were, Pechmann says. "At least for this site, our study shows what sort of fluctuations to expect naturally."

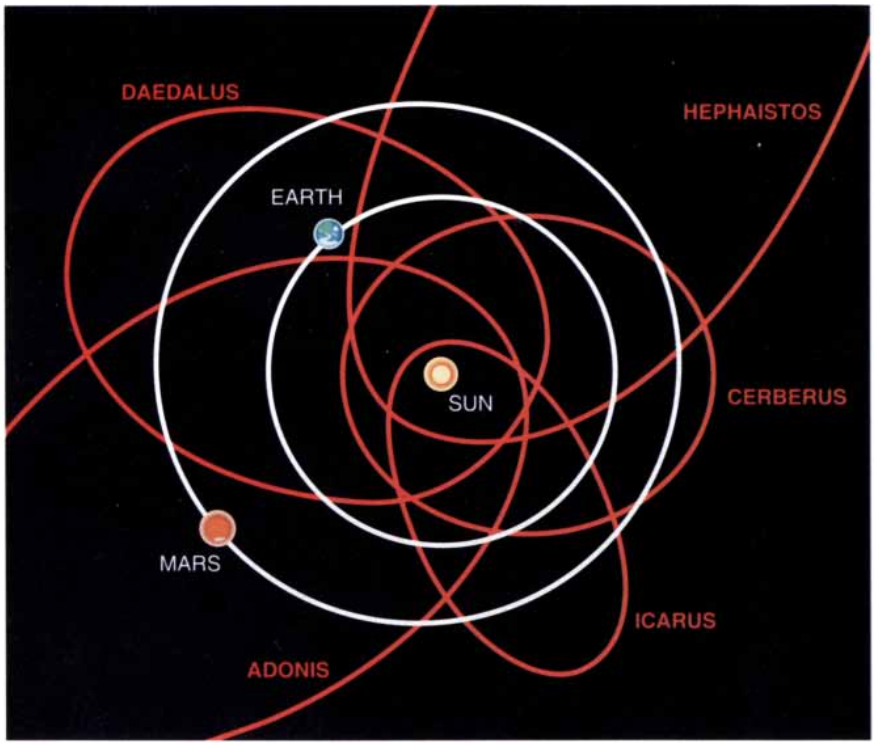
Other researchers, such as Peter J. Morin, a herpetologist at Rutgers University, go further. Morin doubts whether there is any global decline in amphibian populations. "Populations fluctuate, and occasionally local populations go extinct," he says. "What we need to know is whether they are going extinct more than they used to, and we don't have the evidence for that."

The uncertainty adds urgency to efforts to discover the causes of population fluctuations. As Pounds points out—and Pechmann agrees—habitat destruction cannot explain the disappearance of the golden toad: the species is known almost entirely from a small area within the Monteverde Cloud Forest Preserve that has remained undisturbed since the toad was discovered in 1964. The decline of the gastric-brooding frog is likewise a mystery. Pounds is now investigating a possible link between the decline of the golden toad and the unusually warm temperatures in Costa Rica during the late 1980s.

Still, ecologists recognize that the argument for habitat conservation would be a lot stronger if they had better hard data on biological diversity. Some efforts in that direction are getting under way. Earlier this year the Ecological Society of America outlined a "sustainable biosphere initiative" to improve understanding of biological diversity, global change and sustainable ecological communities. "What we point out is that we don't know what the magnitude of the biodiversity problem is and that we need more information," says Jane Lubchenco of Oregon State University, president-elect of the Ecological Society.

So far the exercise exists largely on paper, but already some federal agencies have used ideas from the initiative in planning their budgets, Lubchenco says. Proponents hope the initiative will eventually become a $500-million-a-year effort. For now, though, extinction for many species remains as plain as the business end of a bulldozer. —*Tim Beardsley*

GOLDEN TOAD, found only in Costa Rica, faces extinction. Photo: Michael Fogen/Animals Animals.



APOLLO ASTEROIDS, like the five indicated here, have orbits larger than that of the earth and cross its path at their closest approach to the sun. More than 100 have now been tracked. Source: Lucy McFadden of the California Space Institute.

Big Bang

NASA *wants to fend off doomsday asteroids*

Now that all-out nuclear war seems to be receding as an imminent threat to life as we know it, the National Aeronautics and Space Administration has come up with something else to keep us worried: doomsday asteroids. The agency will soon present to Congress its proposal for a survey to track the orbits of asteroids that could strike the earth and wreak global havoc. A separate report will describe how an impact might be averted by nudging the threatening object into a new orbit.

Fears that a wandering lump of solar system debris might put an end to humanity were confined to science fiction a couple of decades ago. But searches by enthusiasts such as Eleanor F. Helin of the Jet Propulsion Laboratory in Pasadena, Calif., have found that asteroids in orbits that bring them close to the earth are more numerous than once thought. Almost 200 such asteroids have now been tracked, and according to one estimate there may be as many as 10,000 that measure between 0.5 kilometer (about three tenths of a mile) and five kilometers across. The stock is continually replenished by comet remnants, Helin believes. Several new ones are now discovered in an average month.

Proponents of a survey point out that collisions and near-misses are far from rare. Several mass extinctions may have been caused by impacts, and this century has seen at least one direct hit—the Tunguska Event. That 10-megaton detonation, which occurred over Siberia in 1908, is believed to have been caused by an exploding meteor. In 1972 an object more than 80 feet in diameter skipped across the atmosphere over North America, and two other near-misses have been recorded recently. The most celebrated was in March 1989, when an asteroid half a mile across came within 400,000 miles, or six hours, of the earth.

"It's the 99 percent that have not yet been discovered that we are concerned about," says David Morrison of the NASA Ames Research Center, who chaired the report on asteroid detection near the earth. If an object only 0.5 kilometer across struck the earth, the blast would be equivalent to more than 1,000 megatons of TNT and would probably produce a global "nuclear winter."

NASA's proposed survey would map the orbits of objects more than 0.5 kilometer across that come close to the earth. Knowing the orbits might allow several years' warning of a collision—long enough to think about launching an intercept mission. Morrison's committee is expected to outline a network of about five dedicated wide-field telescopes at strategic locations around the world. The telescopes would be equipped with charge-coupled devices, which can search for moving objects much faster than humans can scan film plates.

Money, as always, is likely to be the problem. The telescopes would cost several million dollars apiece, and many astronomers think that kind of funding is pie in the sky. "As soon as people realize how much it's going to cost, they will not want to spend that much money," says Joseph Veverka, an astronomer at Cornell University. And the cost of a mission to intercept an asteroid would be vastly greater. Among the options being discussed are mounting a rocket motor on the asteroid, exploding a nuclear bomb alongside it or simply arranging a passive collision, according to John Rather, the official who chairs NASA's committee on the subject. "It is not in principle a difficult problem," Rather says.

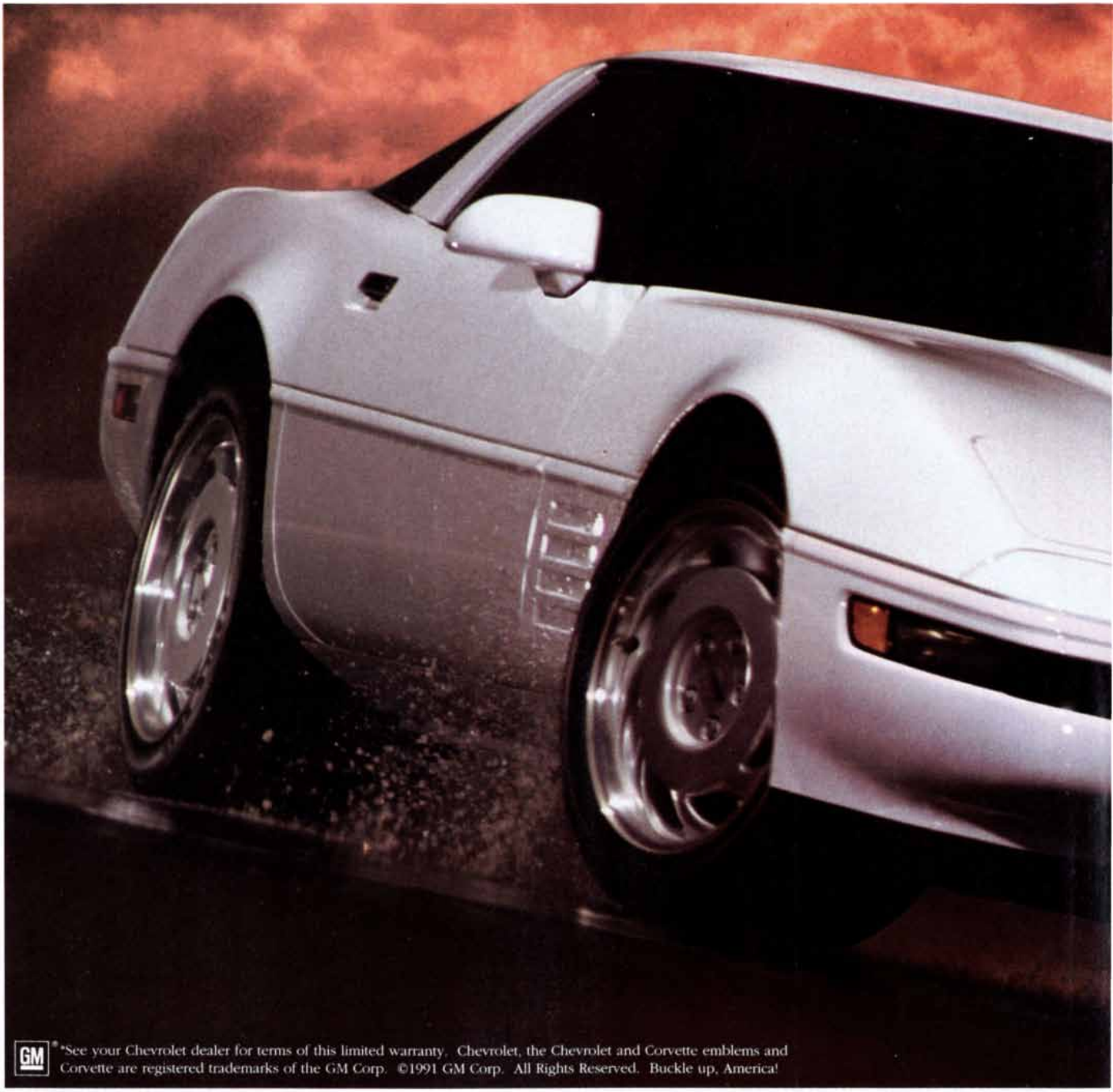
There is also room for doubt about the feasibility of finding all threatening objects more than 0.5 kilometer across. Even a vigorous survey could find only between two thirds to three quarters of them within 25 years, says Edward L. G. Bowell of the Lowell Observatory in Flagstaff, Ariz. "A space-based survey would be useful," he says.

Nevertheless, many astronomers are hoping that the NASA reports will lead to greater recognition for the asteroid searches already under way. Helin, who coordinates an international search that uses about a dozen telescopes part-time, complains that a lack of funds is holding back work. "Right now in Eastern Europe they have no film plates," she says. "I think a lot can be done with existing instruments and within the framework of existing programs," says Brian G. Marsden of the Harvard-Smithsonian Center for Astrophysics. Helin argues that arrays of charged-coupled devices, which are only now becoming competitive with film plates for wide-area searches, should first be retrofitted to existing telescopes.

That is not enough for people like Tom Gehrels of the University of Arizona, who runs an automated asteroid search. "The danger is real—if we don't do something, human society will be obliterated," he insists. Although acknowledging that objects more than 500 meters across probably strike the earth only once in 10,000 years, Gehrels points out that the unlikely event "can happen any time." —*Tim Beardsley*

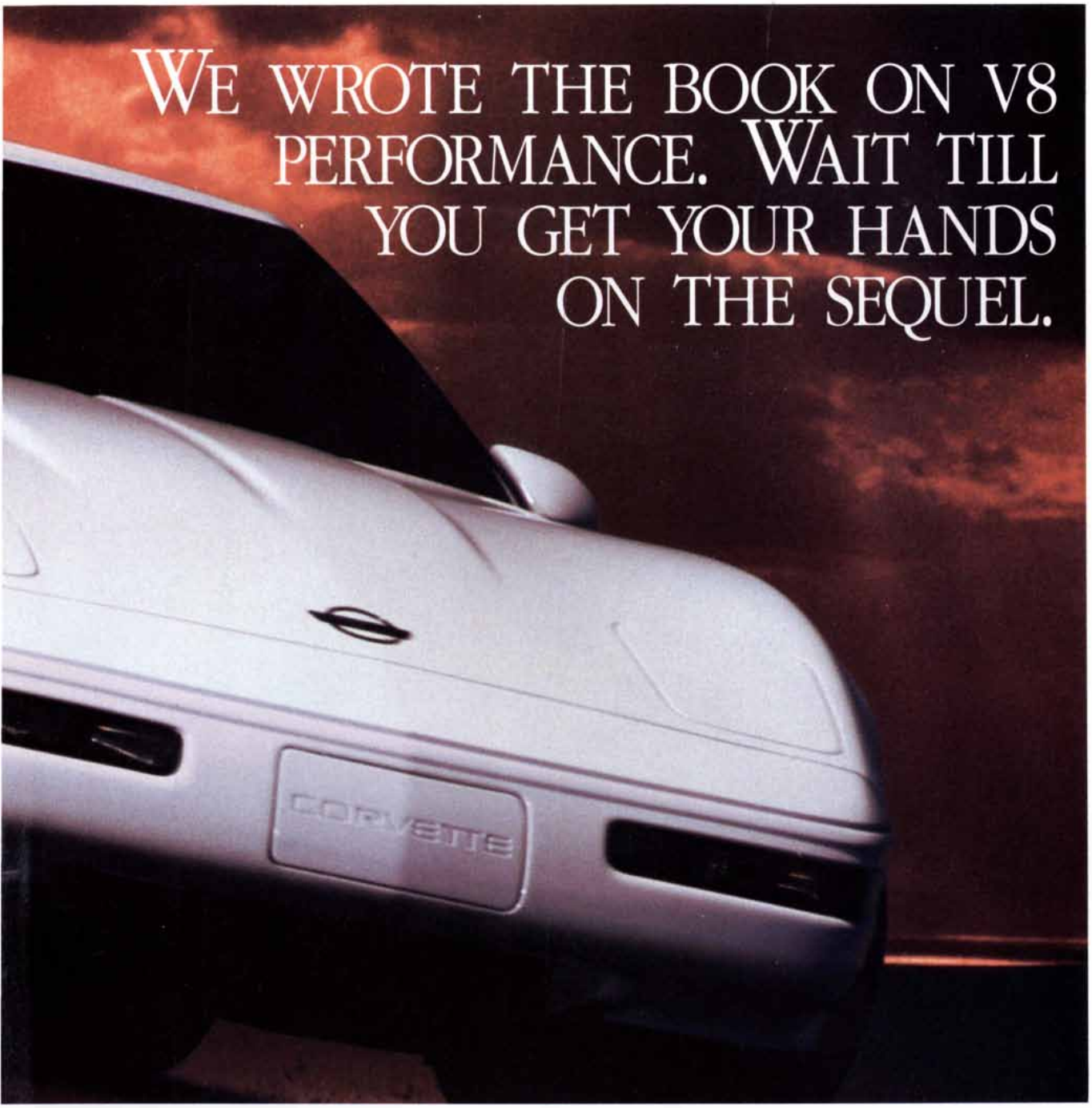
A GRAND REIGN.

To send a gift of Grand Marnier® Liqueur (except where prohibited by law) call 1-800-243-3787.
Product of France. Made with fine cognac brandy 40% alc/vol (80 proof). ©1989 Carillon Importers, Ltd., Teaneck, NJ.



THE 300-HP CHEVROLET CORVETTE LT1

The new LT1 V8 is the most technically advanced small block in Chevrolet history. And only one machine has it: The new 1992 Corvette. ♦ Flaunting such technical wizardry as reverse-flow cooling, a front-mounted Opti-Spark distributor, aluminum cylinder heads, high-compression (10.5:1) pistons and an aggressive hydraulic roller camshaft, the LT1 generates seamless, effortless power. The kind of power that makes an enthusiast's mouth water. And to help get it to the ground, Corvette features Acceleration Slip Regulation(ASR), a highly sophisticated traction control system that regulates the



application of power to the rear wheels, providing optimum grip in all traction conditions. ♦ Test-drive the 1992 Corvette LT1. Cover to cover, it's the best 'Vette yet.

Corvette LT1 Features:

• 5.7 Liter fuel-injected LT1 V8 with 300-HP @ 5000 RPM. • Acceleration Slip Regulation (ASR). • New directional and asymmetric high-performance Goodyear GS-C Eagles. • Power 4-wheel Bosch ABS IIS anti-lock disc brake system. • Choice of 4-speed automatic or 6-speed manual. • No-deductible, 3-year/36,000-mile Bumper to Bumper Plus Warranty.*

THE HEARTBEAT OF AMERICA IS WINNING.™

Light. It is magical. Powerful. Essential.

At GTE, we put that power at the touch of your hand. It is the power to illuminate the facts of science. To brighten the arenas of sports and business. And create the warm glow that means "home."

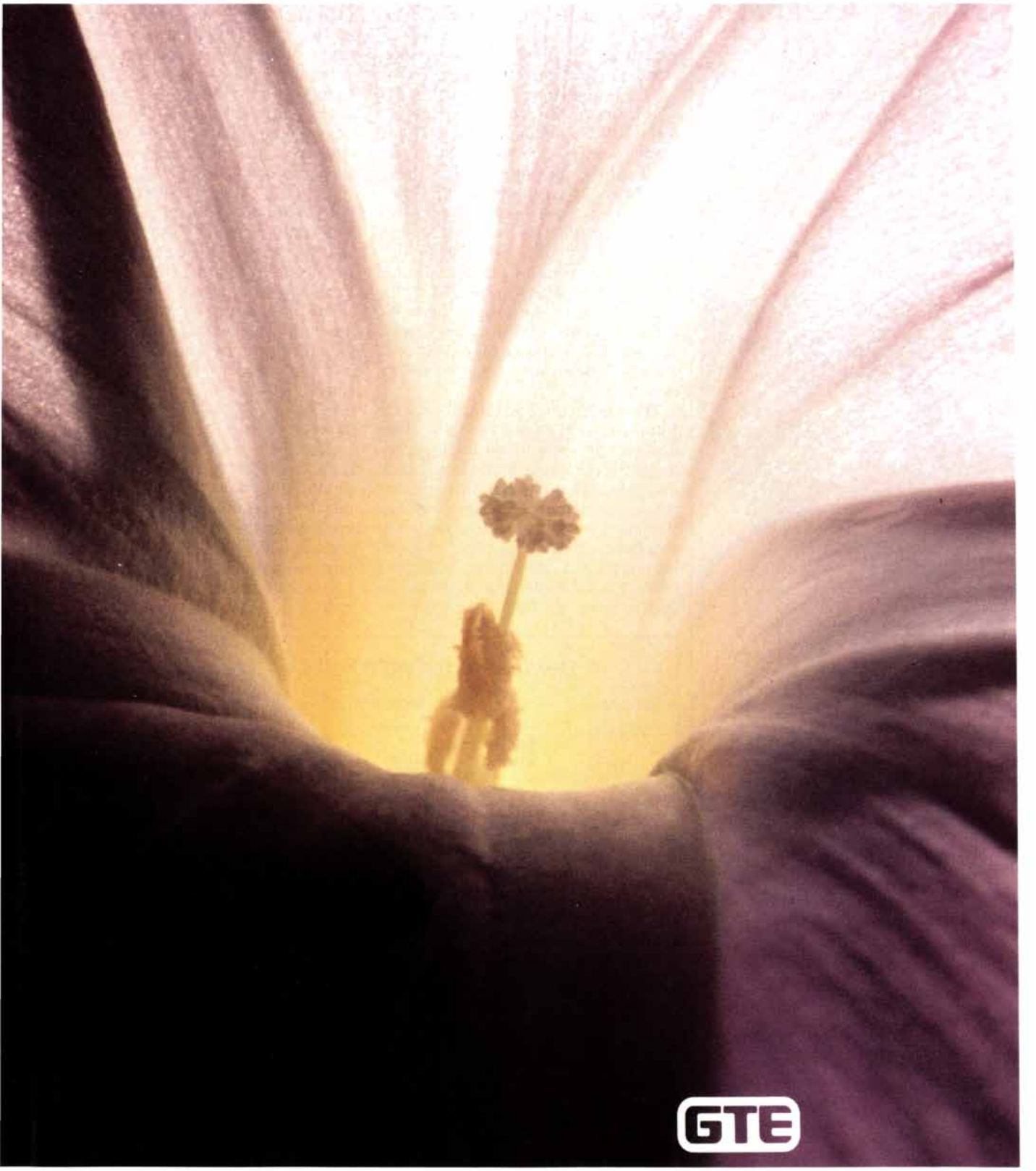
Our Sylvania engineers have introduced more lighting innovations to the U.S. than anyone in the past ten years. For example, our pixel fluorescent tube

echnology is changing the face of video scoreboards worldwide. And we were the first to bring halogen lighting technology to the automotive industry.

We can also make things grow, make businesses run more efficiently, ease a small child's fear of the dark.

At GTE, we can give you the power of light in more than six thousand ways. None of them are small.

Because at GTE, the power is on.



THE POWER IS ON

Quantum Pinball

A quantum system can be observed without an observer

Imagine a pinball machine that propels photons of light instead of steel marbles and that uses mirrors as bumpers. If you watch the path of a photon as it is fired into play, it ricochets off the mirrors as if it were a solid object. But if you stop observing, things begin to get strange. Unlike a steel ball, the photon shatters into many wavelets, each taking a different route through the machine.

Actually, it should make no difference whether you look or not. What counts, according to the theory of quantum mechanics, is whether the photon is shielded from all possible devices that could monitor its path. If it is, the photon behaves like a wave. If not, the photon acts like a particle. But how do physicists know how a bit of matter or energy is behaving when no one observes it?

They didn't until investigators at the University of Rochester built a working quantum pinball machine. In the experiment, the photons are shot from a laser and deflected at random toward either the left or right side of the apparatus. If a photon takes the path to the left, it enters a crystal and splits in two, creating one photon called the signal and another called the idler. If a photon veers to the right, its fate is virtually the same. The photon hits a second crystal, which breaks the photon into a "right" signal and a "right" idler. Both right and left signal photons are directed by mirrors to a detector, whereas the idler photons end up in a counter.

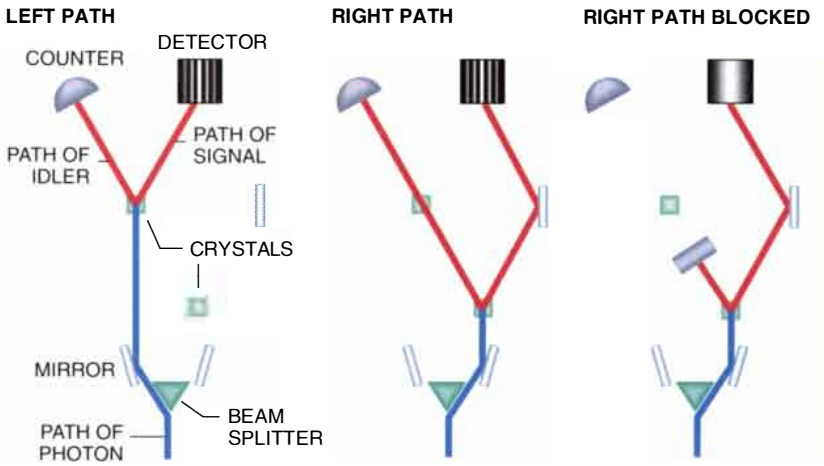
After the counter has recorded many photons, the detector reveals a pattern of light and dark bands. Such patterns are the familiar signature of wave behavior. If the photons behaved like particles, that is, like steel marbles, the detector would record a single bright band.

The pattern of light and dark bands arises because something is interfering with the signal photons. But what? If the apparatus is producing signal photons, one at a time—a perfectly reasonable assumption—then the signal photon should emerge from either the right or the left. Consequently, the signals from the left cannot influence the signals from the right.

But that's not quite right. To understand the interference pattern properly, one must delve into quantum theory. From that viewpoint, the crucial feature is that the left path of the photon is indistinguishable from the right path. In other words, regardless whether the photon takes the left path or the right path, the result is the same: a signal photon hits the detector, and an idler hits the detector. Because the two paths are indistinguishable, the signal photon actually takes both routes and interferes with itself.

To the Rochester group, the interference of the signal photons was no sur-

A Device That Can't Tell Left from Right



The beam splitter deflects the photon to either the left or right. If the photon veers to the left, it strikes a crystal, which splits the photon into an idler and a signal. The idler enters a counter; the signal hits a detector. If the photon goes to the right, its path is different, but its fate is the same. The detector and the counter cannot be used to tell left from right. Blocking the "right" idler photon allows the two paths to be distinguished, and surprisingly such action changes the pattern recorded by the detector.

WIZARD®

ORGANIZATIONAL POWER: PURE AND SIMPLE.

The Wizard® OZ-8000 has all the power you need to stay on time, on track and in control. Eleven built-in functions and a raised typewriter keyboard make it easy to store, retrieve and sort phone numbers, important dates, your schedule, memos and more. You don't need to enter DOS commands. Just touch any function key and the Wizard displays the information you want.

EASY TO EXPAND.

Insert any optional Wizard software card for easy access to more power. The Wizard's patented touch-screen turns each card into a customized keypad. Choose from reference, business, entertainment, telecommunications and memory expansion cards that are as easy to use as the Wizard itself.

EASY TO EXCHANGE DATA.

With the optional Organizer-Link and cable, you can load files from your PC into your Wizard, or download work done on your Wizard to your PC.

EASY TO COMMUNICATE.

An Organizer Fax/Modem gives you the power to send a fax directly from your Wizard. Or, connect with on-line services to send and receive data over direct phone lines and even from a cellular phone. To learn how easy it is to organize your business and personal life, call 1-800-321-8877. In Canada, call 1-416-890-2100. And ask for the dealer nearest you.

SHARP®
FROM SHARP MINDS COME SHARP PRODUCTS™

prise. They concentrated on the behavior of the idler photons to help determine the role of the observer. They blocked the path of the right idler photon so that it could not reach the counter. Now if the counter recorded a photon, they could deduce that the signal photon took the route to the left. If the counter did not record a photon, the signal must have gone to the right. Because they were able to figure out the path of the signal, the photon acted like a particle, and the detector recorded a bright band instead of an interference pattern.

They next blocked the path of both left and right idler photons so that neither could reach the counter. By doing so, they could no longer distinguish paths, thus effectively removing themselves as observers. Nevertheless, the interference pattern was not restored. "The mere possibility that the paths can be distinguished is enough to wipe out the interference pattern," says Leonard Mandel, who conducted the experiments with Li-Jun Wang and Xing Yu Zou.

In the world of quantum mechanics, therefore, observations can be made without observers. That leaves one important question unanswered: If a tree falls in a forest and no one hears it, does it make a sound?—*Russell Ruthen*

Endless Endgame?

Chess masters down the centuries have held that in pawnless endgames, such as the position shown below at the left, the advantage of rook for knight normally cannot force a win. Best play, they intoned, inevitably leads to a draw—that frustrating outcome to the locking of mind and ego over the chessboard.

The masters were wrong. A program written by Lewis B. Stiller, a graduate student at Johns Hopkins University, and run for more than four hours on a Thinking Machines Corporation massively parallel computer, has found a way to win. In doing so, the program produced a winning line 223 moves long, by far the longest in the 1,000-year history of the world's most deeply studied game.

Like certain computer-solved mathematical proofs, the analysis is too complicated to be checked, except by another computer. "We ran the first 173 moves twice and did other kinds of checks, so it's very unlikely an error crept in," Stiller says. "Still, we're dealing with roughly 32 trillion operations, and a stray neutron might conceivably have bounced into the computer chips."

Stiller designed his program to analyze all possible endgames involving two kings, four pieces and no pawns. The task, which requires the generation and classification of about 100 billion positions, begins at the end: The computer produces the handful of possible winning positions and works backward from them, generating a branching and rebranching tree of analysis. The machine then determines whether the balance of forces leads to a win in the general case (that is, where the superior side is not immediately compelled to lose a piece).

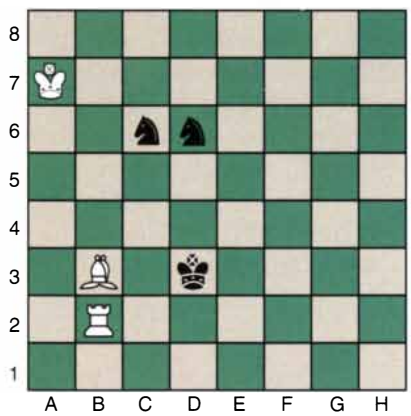
For the first 200 moves or so, the pieces seem to dance about aimlessly, conforming to no rules a human master might recognize and follow. Matters become clear only near the end, when the Black King's back is against the wall and its attendant knights can no longer protect one another. After 222 moves, the White King is about to occupy the square f5, forcing the win of a knight. The resulting struggle of king, rook and bishop against king and knight leads quickly and simply to checkmate.

Kenneth L. Thompson of Bell Laboratories was the first to apply retrograde analysis to computer chess. His program proved, for example, that king and queen win against king and two bishops—an endgame that chess manuals had concluded would end in a draw. Thompson's program exhausted the possibilities of five-piece problems, in one case finding a winning line of 71 moves.

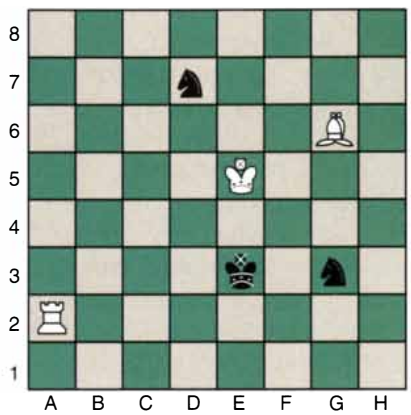
This unprecedented feat led the International Chess Federation to amend its rule on how long a game can proceed before it must be declared drawn. Previously a player had been required to force checkmate within 50 moves after the last capture of a piece or move of a pawn, on the all too human assumption that no winning line could possibly be longer.

Stiller refined Thompson's program to take advantage of the Connection Machine CM-2, which has 65,536 processors and eight billion bytes of memory. Thompson's one-processor algorithm had to incorporate an extra step to reduce the number of positions under review, whereas Stiller's parallel algorithm is simpler. It considers more positions but spends less time on each one.

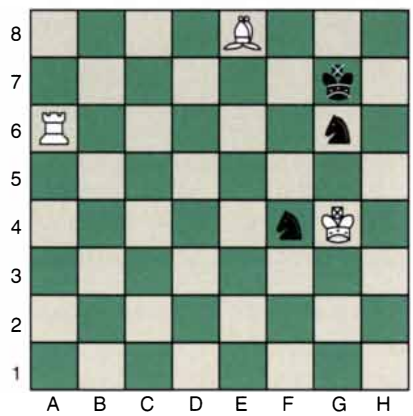
When playing from the initial position, however, computers still lose to the very best humans. Gary K. Kasparov, the chess champion of the world, has vowed that no machine shall ever defeat him. But although he has beaten Deep Thought, IBM's chess machine, its programmers expect to unveil a vastly improved version within a year or so. They say it will end the human domination of the chessboard forever. —*Philip E. Ross*



POSITION BEFORE WHITE'S FIRST MOVE



POSITION BEFORE WHITE'S 120th MOVE



POSITION AFTER WHITE'S 222nd MOVE

IBM presents the that brings

If you're a glutton for productivity, chances are you're considering open systems solutions for your business. But much of the menu is so unfamiliar or limited, it's enough to make you lose your appetite.

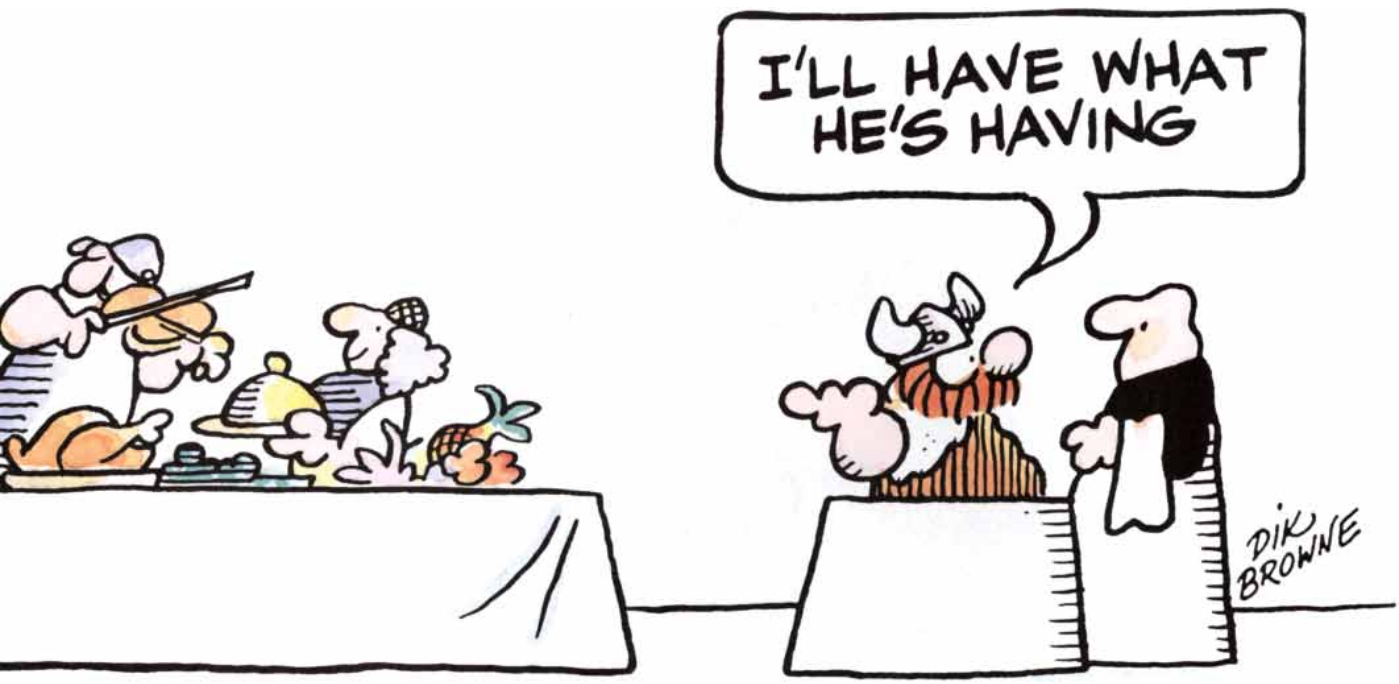
The IBM RISC System/6000™ family of POWERstations and POWERservers is for you. It's as open an open system as you'll find anywhere, to help you take a big bite out of communication barriers. With the RISC System/6000, you offer your suppliers and customers broad connectivity and management for IBM and non-IBM networks. And you also satisfy your hunger for the latest leading-edge technology, without network or equipment upset.

Naturally, the RISC System/6000 family supports all major industry standards. And it also runs a smorgasbord of over 3,500 key applications. But that's only the appetizer. The family provides the widest range of binary compatible systems, so you're free to put the right level of machine performance and price at every location.

The RISC System/6000 also serves up remarkable performance, capacity and value. The POWERserver 550 runs at speeds you'd expect from a supercomputer, but runs up a tab starting at only $58,000. And

*Prices subject to change. Dealers' prices may vary. IBM and AIX are registered trademarks and RISC System/6000 is a trademark of International Business Machines Corporation. UNIX is a registered trademark of UNIX Systems Laboratories. HAGAR THE HORRIBLE Character(s) © 1991 King Features Syndicate, Inc. © 1991 IBM Corp.

first open system more to the table.



other models' prices start below $14,000.*

AIX,® the pièce de resistance. To make UNIX® even easier to swallow, there's AIX, IBM's UNIX operating system. AIX makes it a piece of cake to develop company-wide networked solutions. And it gives companies robust operating software for enterprise-critical applications.

Service with a smile. The RISC System/6000 comes with something you won't find anywhere else—the unparalleled support of IBM. Service 24 hours a day, 365 days a year. But then, at IBM, customer service and support aren't just fillers. They're your just desserts.

To find out more about the IBM RISC System/6000, contact your IBM marketing representative or Business Partner. For literature, call 1 800 IBM-6676, ext. 824.

For the Power Seeker.

IBM®

The Pied Piper of Superstrings

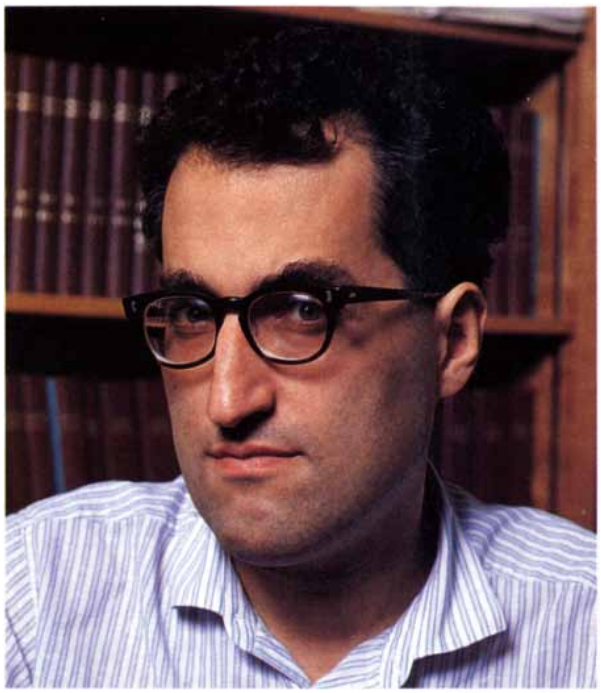
Edward Witten is not one to shy away from bending a dimension—time, say—if it suits his purpose. Witten is an ardent advocate of superstring theory, which holds that all physical phenomena—from quarks to quasars—arise from infinitesimal strings wriggling in no fewer than 10 dimensions. Defending the highly mathematical theory from the charge that it has made no substantive contributions to physics, Witten argues that, on the contrary, "all the really great ideas in physics" are "spinoffs" of superstrings.

General relativity, for instance. Of course, Witten knows that Einstein devised the theory, which explains gravity as a function of curves in space and time, more than half a century before superstrings came along. But Witten calls that "a mere accident of the development on planet Earth." It is quite likely, he says, that "other civilizations in the universe" discovered superstring theory first and then derived general relativity later.

This is not the attention-seeking hyperbole of some fringe scientist. Witten, a 40-year-old theoretical physicist at the Institute for Advanced Study in Princeton, N. J., is widely viewed as the most brilliant physicist of his and quite a few other generations. The stock comparison is Einstein, but John H. Schwarz of the California Institute of Technology, a pioneer of superstrings, reaches even further back for comparisons: he thinks that in sheer mathematical mind power, Witten may have no rival since Newton.

Witten's reputation was already firmly established some seven years ago, when he declared that superstrings represented the ultimate goal of physics: a theory that could explain all the forces of nature. Such a theory is sometimes called the "theory of everything," although Witten disdains the term. "Witten did not found string theory," says Michio Kaku of the City University of New York, "but he created a sensation by saying, 'This is it,' and then making profound contributions to the field." In no small part because of Witten's efforts, many of the world's "brightest" young theoretical physicists, Kaku says, are now studying superstrings. (Kaku, naturally, is a string theorist.)

Nonbelievers worry that Witten may be a Pied Piper, leading his followers away from reality and into a phantasmagoria of pure mathematics. They point out that superstrings, which are thought to be as small in comparison to a proton as the proton is in comparison to the solar system, are not likely to be detected in this or any millennium. Nor are the surplus dimensions the strings supposedly inhabit.



PHYSICIST EDWARD WITTEN hopes to unravel the Gordian knot at reality's core. Photo: John Pinderhughes.

One of the few skeptics willing (indeed, eager) to speak on the record is Sheldon L. Glashow of Harvard University. He contends that the theory is not physics at all but merely mathematical "smoke and mirrors." Like many critics, he also complains about the theory's impenetrability. "I don't understand Ed's recent papers," says Glashow, a Nobel Prize–winning physicist, "so I can't tell you how brilliant he is."

Witten shrugs off the criticism. He acknowledges, even emphasizes, that no one has truly fathomed the theory and that it may be decades before it yields a precise description of nature. But he is serenely confident that it is the way to truth. "Good wrong ideas are extremely scarce," he says, "and good wrong ideas that even remotely rival the majesty of string theory have never been seen."

I had read before interviewing Witten that some of his students at Princeton University, where he taught until 1987, have called him (with "great reverence and respect," the article stressed) "the Martian." Meeting him, I can understand why. His tall, slim body seems almost an afterthought in comparison to his looming head. His lofty forehead is bounded by thick black hair on the top and thick black glasses on the bottom. He speaks in a high-pitched voice so soft that I worry my tape recorder won't pick him up over the air conditioner. He pauses frequently—for 51 seconds at one point—casting his eyes down and squeezing his lips together like a bashful teenager. He seems to be striving for mathematical precision.

Yet Witten is hardly reticent. I have barely sat down in his office when he chides me for having written in these pages about the philosopher of science Thomas S. Kuhn, who believes that science is "arational." Witten insists that, as penance for wasting print on Kuhn's "silly" ideas, I should profile five mathematicians in a row. He isn't kidding. Witten also declares that he will not answer "personal" questions, such as what he majored in at college. I must consult other sources to learn that although Witten is the son of a theoretical physicist and always enjoyed physics, he entered the field by means of a curved route. Indeed, he graduated from Brandeis College in 1971 with a degree in history and designs of becoming a political journalist. After publishing articles in the *Nation* and the *New Republic* and working on George McGovern's 1972 presidential campaign (yes, a man who is arguably the

INTRODUCING DIAMANTE. PROOF THAT A GRACIOUS HOST CAN ALSO BE A STIMULATING ONE.

The proof is presented in a series of revelations.

First you notice the richness of the interior. The many small courtesies, such as the ignition key light and automatic door lock. And the seat, which supports you with almost no sense of pressure against your legs, hips and back.

You might conclude that this is a modern, well-crafted luxury car. But then it reveals a gratifying exuberance as well. A dramatic rush of power. Athletic steering response. It feels not just competent, but eager to perform.

Welcome to the new Mitsubishi Diamante. Conventional "luxury" and "performance" labels do not apply here. It is the best of both, without the limitations of either.

A panel of 49 international judges recently evaluated the Diamante in detail. And then named it Japan's "Car of the Year." To make your own evaluation, please call 1-800-447-4700 for your nearest Mitsubishi Motors dealer.

MITSUBISHI
The word is getting around.™

smartest person in the world is a liberal Democrat), he decided he lacked the "common sense" for journalism (or so he told one reporter). He entered Princeton to study physics and obtained his doctorate in 1976.

Witten also followed a roundabout path to superstrings. Much of his early work was tied relatively closely to mainstream, experimentally accessible physics. He wrote numerous influential papers on quantum chromodynamics, which explains the strong force that binds atomic nuclei together. He also addressed problems in electroweak theory, which showed the underlying unity of the weak nuclear force (responsible for certain types of nuclear decay) and electromagnetism.

But increasingly he gravitated toward more speculative unification schemes involving higher dimensions. The key to understanding such theories—and modern physics in general—is the concept of symmetry. There are various types of symmetry, from the simple rotational symmetry possessed by a sphere to more complicated varieties involving the interchangeability of particles. One way physicists seek to glimpse the hidden symmetries of a system is by imagining it from a higher-dimensional viewpoint. In a similar way, an astronaut who rises above the two-dimensional surface of the earth is able to perceive its rotational symmetry. Witten, echoing the title of a surrealistic Victorian-era novel about two-dimensional beings, has called this process "stepping out of *Flatland*."

Witten explored two higher-dimensional approaches—called supersymmetry and Kaluza-Klein theory—that were to play a role in superstrings. Supersymmetry posits an underlying symmetry (and thus interchangeability) between particles that constitute matter and those that transmit force. One predicted consequence of supersymmetry is that all known elementary particles have unseen supersymmetric twins, sometimes called sparticles. Kaluza-Klein theory, named for two European scientists who developed it early in this century, posits the existence of additional dimensions that are somehow rolled up, or "compactified," in such a way that they are undetectable at macroscopic levels. The compactification process is analogous to rolling a two-dimensional sheet of paper into a linelike tube or crumpling it into a pointlike ball.

These concepts, and supersymmetry in particular, helped Witten discern deep commonalities underlying nature's forces. Yet neither approach could overcome the central barrier to

complete unification: the incompatibility of quantum theories, which explain electromagnetism and the nuclear forces, and general relativity, the theory of gravity. Witten calls this "the single biggest puzzle in physics."

Meanwhile superstring theory had been developing independently of Witten. The basic string concept emerged in the late 1960s, when Gabriele Veneziano of CERN and Yoichiro Nambu of the University of Chicago proposed that stringlike particles could give rise to the strong force. Interest in the theory faded when other physicists showed that it would only work in 26 dimensions. Then, in the 1970s, Schwarz of Caltech and others combined the string concept with supersymmetry to form the much more ambitious superstring theory. They contended that just as the vibrations of violin strings give rise to different notes, so could the vibrations of superstrings give rise to the mass-carrying particles and to all the forces acting on them—including gravity.

Although superstring theory required only 10 dimensions rather than 26, that was still six too many for most physicists. Skeptics were also put off by the fact that the theory dealt with energies and distance scales far beyond the range of any earthly particle accelerator. Moreover, in the late 1970s the attention of the physics community was focused on the astonishing confirmation of quantum chromodynamics by experiments in particle accelerators.

Witten first learned of superstring theory in 1975, but his initial attempts to understand it (others might be gratified to know) were frustrated by the "opaque" literature. In 1982, however, a review paper by Schwarz helped Witten grasp a crucial fact: superstring theory does not simply allow for the possibility of gravity; it demands it. Witten calls this realization "the greatest intellectual thrill of my life."

Although Witten began to work on the theory, he initially resisted proselytizing on its behalf. "I was reluctant to encourage young people to become involved in what I've always thought of as a very long term proposition," he says. In a 1984 paper, moreover, he and Luis Alvarez-Gaume of Princeton pointed out a problem that had to be overcome by superstrings and any other unification theory that included gravity. The problem had to do with anomalies, or mathematical inconsistencies, that arose when such theories tried to account for an asymmetric type of nuclear decay.

Later that year, however, Schwarz and Michael B. Green of Queen Mary College, London, published a paper

demonstrating that the anomalies could be eliminated in superstring theory. By this time, Witten's epiphany was complete. "It was very clear that if I didn't spend my life concentrating on string theory, I would simply be missing my life's calling," he says. He began pronouncing the theory a "miracle" and predicting that it would "dominate physics for the next 50 years." More important, he issued a flood of papers on the theory—19 in 1985 alone.

Much of Witten's early work focused on creating a model that was a reasonable facsimile of the real world. Drawing on his experience with Kaluza-Klein theory, Witten and several other theorists described how the six extra dimensions might have become compactified. They then suggested how the hidden dimensions might affect particle interactions at energies within the range of current accelerators. Many other investigators, at CERN and elsewhere, are continuing to work in this vein.

Yet Witten became increasingly convinced that to forge superstrings into a truly predictive model, theorists had to uncover its "core geometric principles." These principles, he says, may be analogous to the non-Euclidean geometry invented by the 19th-century mathematician Bernhard Riemann. Lacking this technique for modeling curvatures in space and time, Witten contends, Einstein might not have been able to formulate the theory of relativity. "What we have learned in the 20th century," Witten comments, "is that the great ideas in physics have geometric foundations."

His pursuit of these issues has taken Witten deep into topology, the study of the fundamental geometric properties of objects, regardless of their particular size or shape. In the eyes of a topologist, for example, a donut and a single-handled coffee mug are equivalent because each has only one hole. A donut and a banana are not equivalent, because one would have to tear the donut to turn it into a banana.

Witten eventually created a technique, called topological quantum field theory, that combines topology with the mathematical tools of quantum field theory. The technique, Witten says, allows him to uncover even deeper symmetries than the traditional methods of physics. Recently Witten applied the technique to work on knot theory done by Vaughn F. R. Jones, a mathematician at the University of California at Berkeley [see "Knot Theory and Statistical Mechanics," by Vaughn F. R. Jones; SCIENTIFIC AMERICAN, November 1990].

One of the major issues in knot theory concerns whether seemingly dissimi-

lar knots are in fact topologically equivalent—that is, whether one can be transformed into the other without being cut. Jones had shown how to accomplish this task with a relatively straightforward polynomial equation from statistical mechanics, which describes heat transport or other general properties in a complex system.

Subsequently, Witten showed that with topological quantum field theory one could find even deeper unities, or symmetries, between knots than with Jones's technique. One could also discern symmetries underlying knots in curved, non-Euclidean spaces—and even underlying the spaces themselves. Witten calls this his "single most satisfying piece of work." As a result of his finding, Witten and Jones shared the 1990 Fields Medal, the most prestigious prize in mathematics.

Witten bristles a bit in response to suggestions that in his fascination with such mathematical issues, he has lost touch with real-world physics. He points out that in 1985 he wrote a paper with Mark Goodman of Princeton proposing how to detect exotic particles, including the sparticles predicted by supersymmetry, that might account for the "dark matter" many cosmologists think dominates the structure of the universe. The paper has inspired various dark-matter experiments.

In the mid-1980s he also wrote several papers on cosmic strings, extremely long, massive and hypothetical objects that could have catalyzed the formation of galaxies. Witten proposed that cosmic strings might be superconductors and suggested ways they could be detected by astronomers. More recently, Witten helped Frank Wilczek, a colleague at the Institute for Advanced Study, elaborate a theory that explains high-temperature superconductivity as a by-product of unusual particles called anyons.

But Witten says he has "learned the hard way" to trust his instincts about what is important. "Physicists generally, including those who are enthusiastic about string theory, greatly underestimate the depths and subtlety of its underlying geometric ideas and the importance for human thought of unearthing them." Witten recites this line slowly, abstractedly, as if at this very moment he is trying to see into the twisted heart of things. Once those ideas stand revealed, will that be the end of physics? Witten declines to guess. "You would never look at physics again in the same way," he says. "But how you would look at it, and what the unsolved problems would look like, I have no idea."—*John Horgan*

SCIENTIFIC AMERICAN

Firearms, Violence and Public Policy

More guns mean more deaths from crime and accidents. Yet in the U.S. many laws that regulate firearms are misdirected and ineffective. What we don't know about gun control can hurt us

by Franklin E. Zimring

Even though the U.S. has many more gun-control laws than any other nation, Americans are more likely to be victims of gun-related violence. We have no hope of greatly improving this situation as long as we continue to construct gun-control policy on a weak foundation of facts.

Since the 1960s some social and behavioral scientists have been investigating how violence is related to firearms. But their conclusions have largely been ignored. Neither supporters nor opponents of gun-control laws have felt any great need to cite facts. Strong emotions have kept the conflicting parties at each other's throats.

Recent congressional debates on gun control have also relied on undocumented assertions. The 1991 Brady bill aims to reduce the violence associated with guns by making handgun purchases subject to waiting periods and police notification. The congressional debate over these issues is no more informed than it was during the deliberations that led to the Gun Control Act of 1968. A distaste for facts is also evident in the emphasis Americans have placed on research: much more money is spent on newspaper advertisements about gun control than on research about firearms and violence.

During the past three decades, investigators have learned about the relation between guns and the death rate from violent crime, accidents and suicides. They recorded the kinds of firearms used in these situations, and they measured how gun-control laws have influenced crime and the ownership of guns.

The accumulating evidence provides reasons for discomfort on both sides of the political struggle. Ample data confirm that as guns become more available, people are more likely to die during violent crimes—a connection that opponents of gun control have tried to deny. Research also shows that many laws do not significantly diminish the number of guns used in violence, although many advocates of gun control have assumed they would.

Lawmakers have agreed to many measures that try to keep guns out of the hands of criminals. But these regulations have not brought violence down to a level that most Americans can tolerate. At the same time, we have not confronted the controversial issue of reducing the 35 million handguns that play the greatest role in violence caused by firearms.

The Federal Bureau of Investigation reports that in 1990 criminals in the U.S. committed more than 1.8 million acts of homicide, robbery, forcible rape and serious assault. More than 28 percent of these violent crimes involved firearms. And, most notably, guns were the cause of death in 64 percent of the 23,000 homicides. Although the statistics portray the magnitude of gun-related violence, they provide no information about why or how the flow of firearms should be regulated.

The issue of gun control hinges on whether the death rate from violence would subside if people were forced to abandon firearms and choose other weapons such as knives. As Philip J. Cook and Daniel Nagin of Duke University once asked, "Does the weapon matter?"

The answer is not obvious. In 1958 Marvin E. Wolfgang of the University of Pennsylvania claimed that the choice of weapon did not make much difference in criminal homicide. He argued that if guns were unavailable, assailants would wield other deadly weapons to achieve their goal. His analysis assumed that most people who commit criminal homicide have a single-minded and unconditional intention to kill.

A decade after Wolfgang's work, I found evidence that the weapon used is important in determining whether a violent assault will lead to death. I reviewed Chicago police records of more than 16,000 violent assaults, both fatal and nonfatal, to determine the outcomes of attacks with guns and knives, the two weapons most commonly used

FRANKLIN E. ZIMRING is professor of law and heads the Earl Warren Legal Institute of the University of California, Berkeley. For the past two decades, he has investigated the relation between firearms and violence, and he has helped shape gun-control policy in the U.S. For two years, Zimring directed research for the Task Force on Firearms of the National Commission on the Causes and Prevention of Violence.

in criminal homicide [*see box on next page*]. The study showed that most attackers seemed to stop short of ensuring the victim's death. In eight out of 10 assaults with guns, the attacker shot the victim only once. Attackers also inflicted one wound in seven out of 10 cases in which the victim died.

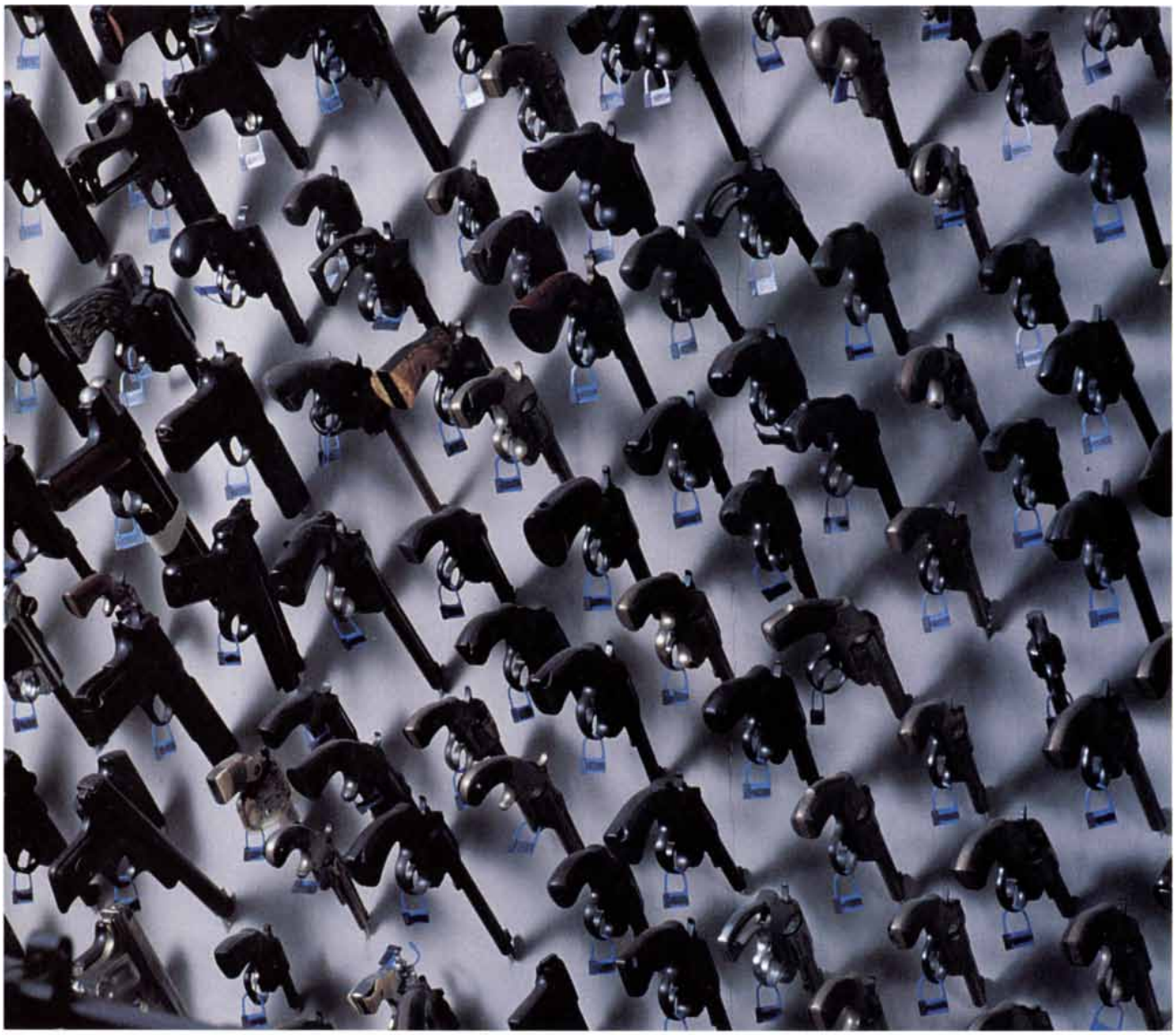
In general, fatality seemed to be an almost accidental outcome of a large number of assaults committed with guns or knives. The Chicago study indicated that most nonfatal attacks, like most homicides, resulted in wounds to vital parts of the body. According to the study, homicide and serious assaults involved the same kinds of motives (mostly spontaneous arguments), and they occurred in the same places at the same times.

One important difference among assaults was the weapon used: an assault with a gun was five times more likely to result in a fatality than an assault with a knife. And assaults with guns to all vital body areas—head, neck, shoulders, chest and abdomen—were many times more likely to kill than attacks to the same locations with knives. In 1990 more than 23 percent of the one million serious assaults involved guns, whereas knives played a role in 20 percent of assaults.

Do attackers who carry guns simply have more lethal intentions than assailants who use knives? The Chicago data indicated otherwise: compared with assaults with guns, attacks with knives were more likely to result in multiple wounds, and they were equally likely to damage a part of the body where death can result. The five-to-one difference in death rate thus seems to stem from the greater dangerousness of the firearm as a weapon, what is known as an instrumentality effect.

If such instrumentality effects are large, a shift from guns to knives would cause a drop in the homicide rate even if the total rate of violent assault did not change. Other studies have corroborated the presence of large instrumentality effects in urban violence.

When Hans Zeisel of the University of Chicago and other sociologists compared assaults in cities that had different mixes of guns and knifes, they found a difference in death rate between the two kinds of assaults. Zeisel, who examined New York and Houston,



HANDGUNS were involved in half of the 23,000 homicides committed with firearms in the U.S. last year. Although federal law forbids felons from acquiring any of the supply of 35 million handguns in the U.S., it has not reduced the violence associated with handguns to a tolerable level. Is restricting handguns America's only option?

wrote, "If the level of gun attacks in Houston were reduced from 42 percent to New York's level of 24 percent, 322 gun attacks would have been knife attacks. At present, these 322 gun attacks resulted in 63 fatalities.... If they were knife attacks, roughly 12 fatalities would result—a reduction from 20 deaths per 100 attacks to four per 100."

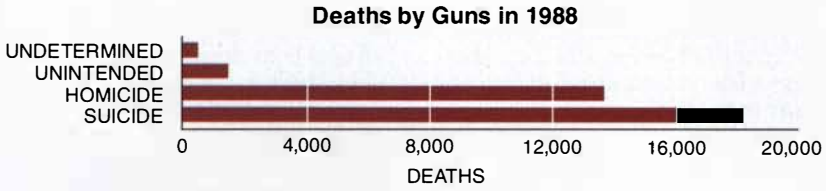
Although all guns are deadly, some types of firearms are more harmful than others because they are more likely to be used in crime and violence. In the U.S. handguns—small, concealable weapons that can be fired with one hand—account for one third of the 120 to 150 million firearms estimated to be owned by civilians. Handguns are used in more than 75 percent of firearm-related homicides and more than 80 percent of firearm-related robberies. On the average, rifles and shotguns are seven times less likely than handguns to be used in criminal violence. In states and cities that have made special efforts to restrict handguns, the major problem is still illegal handguns rather than long guns. (No one has recorded how many crimes are committed with what are called military assault weapons. These semiautomatic rifles and handguns require a separate trigger pull for each bullet discharged but can fire many rounds of ammunition quickly.)

To determine whether the type of gun influences the outcome of violent assaults, my colleagues and I conducted a follow-up study in Chicago in 1972. Not surprisingly, we found that a large-caliber gun was twice as likely to kill as a small-caliber gun in cases in which the guns inflicted the same number of wounds to the same part of the body. This statistic indicated that mortality from criminal violence is strongly correlated with the dangerousness of the weapon, not just the attacker's intent.

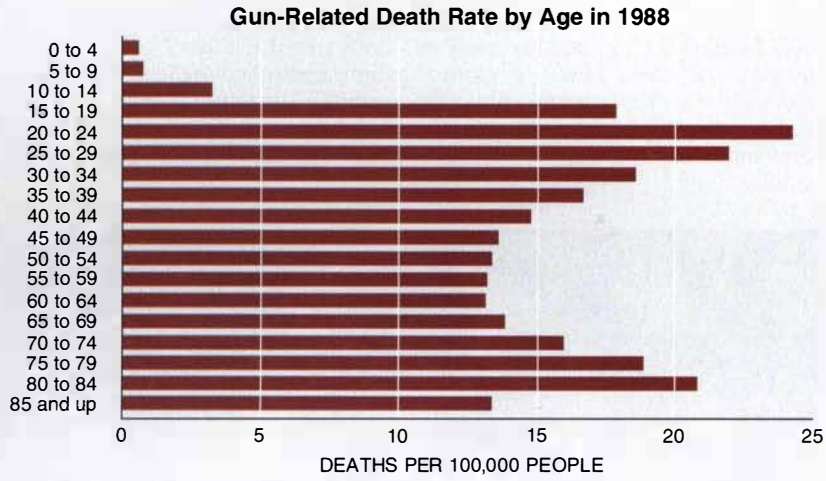
Some sociologists have criticized the conclusion that the motivations behind most homicides are ambiguous. They argue that the similarity between fatal and nonfatal assaults might be superficial. Furthermore, they maintain that attackers who wield guns may have substantially different intentions from criminals who use knives. But so far no evidence derived from data on assault has been presented that argues against instrumentality effects in homicide.

The role of weapons is largely unknown in another kind of assault, forcible rape. More than 100,000 cases are reported to the police every year, and roughly 8,000 of them involve firearms. Although a change in firearm use could prevent many rapes at gunpoint,

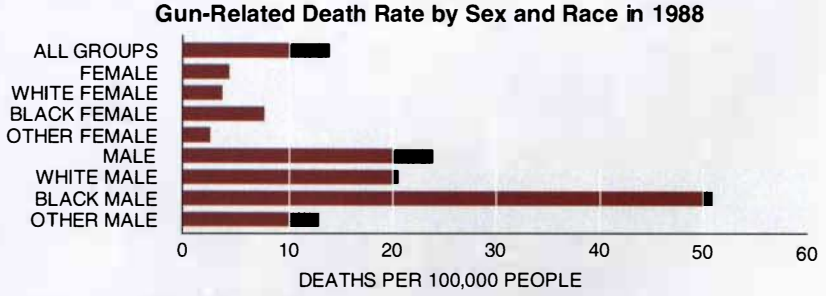
Weapons and Violence



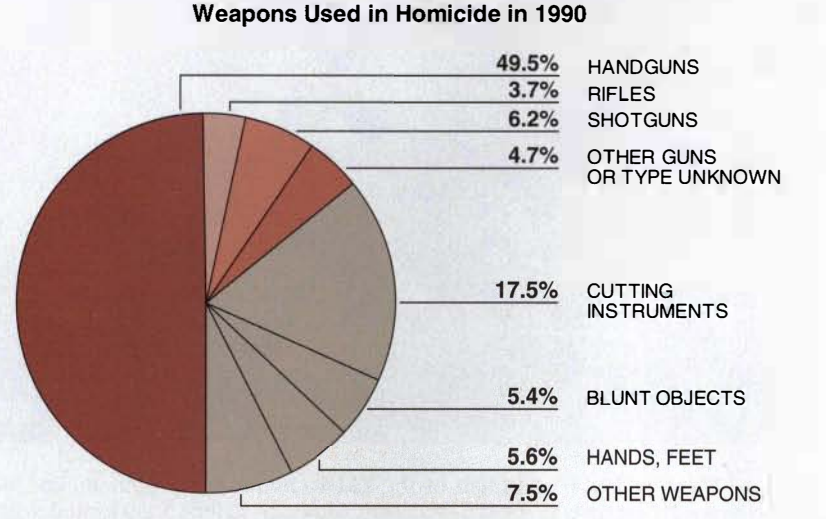
SOURCE: National Center for Health Statistics



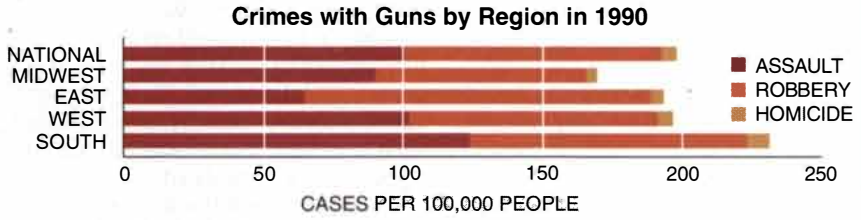
SOURCE: National Center for Health Statistics



SOURCE: National Center for Health Statistics

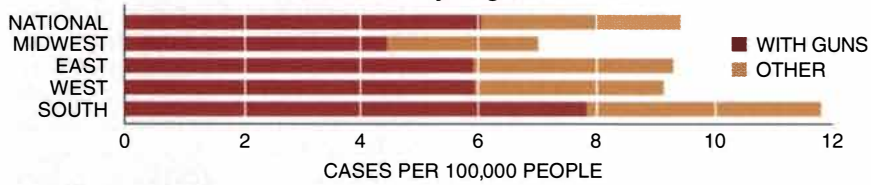


SOURCE: FBI Uniform Crime Reports



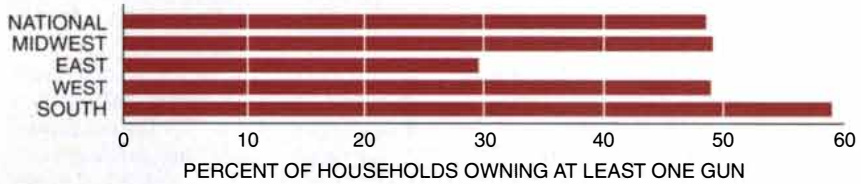
SOURCE: FBI Uniform Crime Reports

Homicides by Region in 1990



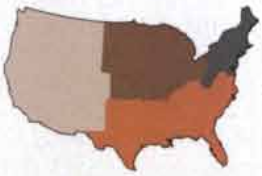
SOURCE: FBI Uniform Crime Reports

Households with Guns by Region in 1991



SOURCE: Gallup Poll

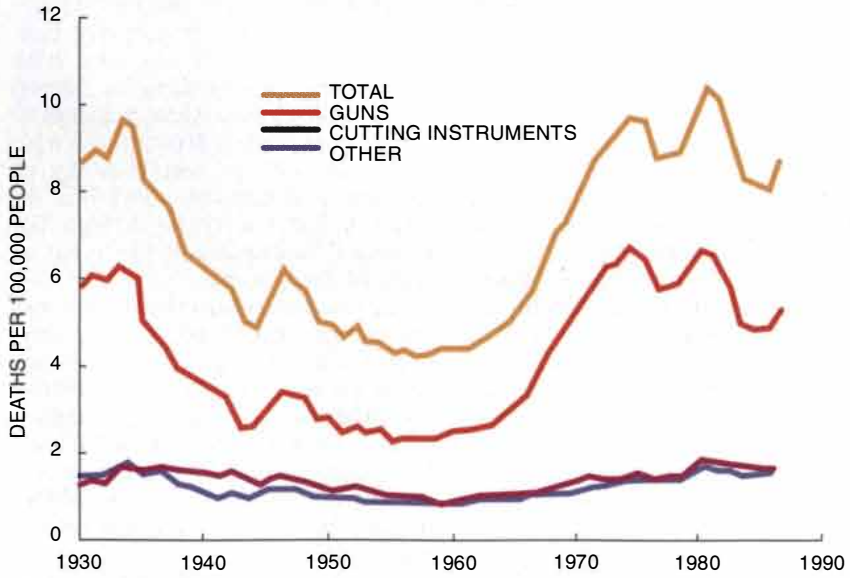
Type of Firearm in Households with Guns in 1991



	NATIONAL	MIDWEST	EAST	WEST	SOUTH
HANDGUN	47%	33%	48%	54%	52%
SHOTGUN	51%	56%	53%	36%	55%
RIFLE	50%	46%	59%	52%	49%

SOURCE: Gallup Poll

Death Rates from Homicide by Year and Method



SOURCE: The Injury Fact Book

most rapists use personal force or other means. No studies have yet determined the different rates of death from forcible rape involving different types of weapons.

A central issue in the gun-control debate is robbery. Each year the FBI tallies more than 500,000 robberies and 3,000 deaths of robbery victims. (This type of crime produces by far the most killings of strangers.)

In 1990 criminals carried firearms in 37 percent of 640,000 robberies reported to the police. Criminals do not rely heavily on firearms for robberies of vulnerable individuals on the street. But they often carry guns to rob stores. My colleagues and I found that in Chicago firearms are involved in two thirds of robberies of commercial establishments but fewer than two fifths of street robberies.

Because robberies need involve only the threat of injury, it cannot be assumed that the choice of weapon in a robbery will influence the outcome in the same way as the choice in assaults does. Only in recent years have investigators published special reports on the influence of guns on the death rate from robbery.

The National Crime Survey reports that crime victims are less likely to resist robbers who carry guns than those who wield other weapons. Apparently, a firearm makes the threat of force by a robber conspicuously credible. The risk of any victim injury is therefore lower when a robber has a gun instead of, say, a knife.

But in cases where robberies result in injuries, guns are far more deadly than other weapons. James Zuehl and I at the University of Chicago found that the death rate is three times higher for robberies at gunpoint than for robberies with knives, the next most dangerous robbery weapon. Cook has found that in areas where gun ownership is unusually prevalent, death rates from robbery are high.

These findings independently confirm the instrumentality effects found in assaults because robbers presumably do not select their weapons with the intent of injuring their victim. Robberies with firearms, like assaults with guns, contribute greatly to the crime-related death rate, independent of the motivations of the criminal.

One question not yet answered is how much the availability of guns influences the rate of robberies. Presumably, when guns are widely available, robberies become easier to commit and increase in frequency. Yet in one study of U.S. cities, Cook did not find a corre-

lation between gun availability and total robbery rate. No one has yet investigated the more specific comparison of commercial robbery rates.

Very few scholars have studied how firearms are related to two important forms of noncriminal violence: suicide and accidents. In the U.S. death by self-inflicted gunshot wounds accounts for over half the more than 30,000 suicides committed every year. Firearms are no more effective a means of suicide than are such methods as hanging or jumping. But when firearms are not at hand, people most frequently attempt suicide by drug overdoses. If guns were not available to people with suicidal intentions, how many attempts might be redirected to less lethal means and how many lives would be saved? This question deserves far more attention than it has received.

Accidents involving firearms claim about 2,000 lives every year in the U.S., a much lower death rate than that for homicides or suicides. About 60 percent of accidental fatalities occur in or around the home. The population groups that have the highest fatality rates from accidents are male children and adolescents, who are generally inexperienced in using guns and are often tempted to play with them. Clearly, if young people did not have access to guns, the death rate would drop.

Two other findings about the misuse of firearms have some significance when considering strategies of firearm control. First, the percentage of gun-related crimes in an area is related to the proportion of owners of firearms in that area. In 1969 the Task Force on Firearms demonstrated that if a city ranks high in gun use for one kind of crime, such as aggravated assault, the use of guns tends to be high for other types of crime, such as robbery. Cook has also found that high rates of gun use in suicide are positively correlated with rates of gun use in violent crime. As more firearms are available to the civilian population, more guns are also available for misuse.

Second, in the 1970s the Federal Bureau of Alcohol, Tobacco and Firearms conducted studies of handguns confiscated on city streets. The bureau showed that handguns are most likely to be misused in the first few years after the weapons are introduced to the civilian market. In four U.S. cities studied by the bureau, handguns three years old or younger constituted half of those confiscated on the streets but less than a quarter of the total number owned by civilians. Consequently, handguns play their greatest role in street crime when they are a few years old, even though they often function for many years.

All gun-control measures seek to reduce the influence of the use of firearms in crimes and violence. The ideal gun-control measure would prevent all crime and violence involving guns without interfering with their legitimate use in contemporary life. More realistically, regulations should reduce the problems caused by illegitimate uses as much as possible while minimizing the restraints on legitimate uses of guns. To achieve this goal, lawmakers have proposed many different types of regulations [*see table on page 54*]. The laws are usually based on one of three strategies: prohibiting high-risk uses of firearms, keeping guns out of the hands of high-risk users or banning high-risk firearms.

Regulations that prohibit high-risk uses of firearms include "place and manner" prohibitions and extra penalties for unlawful use. Most gun laws in the U.S. seek to regulate the place and manner in which firearms may be used. They prohibit such high-risk uses as carrying firearms in a motor vehicle or discharging a firearm in populated areas or concealing weapons on one's person. Place-and-manner laws attempt to deter high-risk behavior with guns, and they rely on the police to intervene before violence occurs. These laws prevent violence only to the extent that the police can discover and arrest persons who violate such laws. How much violence is deterred because guns are not carried is not known. But the rate of gun-related violence would almost certainly increase if carrying loaded weapons became widespread.

High-risk behavior with guns can also be discouraged by instituting particularly stiff penalties for criminals who use firearms. Such laws attempt to deter the would-be robber or attacker from using a gun during a crime. More than half of the states have passed such laws. This approach is popular with gun owners because the penalties concern only gun-related crime and place no restrictions on firearm ownership.

Extra-penalty laws are limited, however, in their potential to curb firearm violence. To reduce the number of crimes involving guns, such laws would have to discourage persons who would not be deterred by the already severe penalties for robbery and assault. Can the threat of additional punishment succeed? A robber might be deterred from using a gun if the penalty for robbery with a gun were several times greater than that of robbery without a gun. But punishment for robbery is already severe, at least as set forth in the statutes. In addition, a robber who carries a gun should not be punished so harshly that the additional penalty he risks if he injures or kills his victim is relatively small. It seems unlikely, therefore, that such laws have much more potential to deter gun use in robberies.

The issue of additional deterrence is also complicated for the crime of assault with a gun because the person who attacks someone with a gun is already risking the maximum punishment if the victim dies. A series of studies—the most notable by Colin Loftin and his associates at the University of Maryland—find that at best extra-punishment laws reduce gun-related crime by a small amount.

Extra punishment and place-and-manner regulation do have a role in comprehensive firearm control. But they cannot be used as the primary controls in a system that aspires to influence the death rate from violence substantially.

Laws attempting to discourage high-risk uses of firearms can be more effective when they are combined with selective ownership prohibitions. This second gun-control strategy seeks to deny high-risk users access to firearms. Usually the law considers high-risk users to include convicted felons, minors, adjudicated mental incompetents, drug addicts and fugitives from justice. The theory is that high-risk users should not be allowed to own firearms, because the societal damage they cause through violence outweighs the social value of their interests in using guns legitimately.

The federal government and nearly every state prohibit some type of high-risk ownership. But many of these laws do not make a person prove his eligibility to own a gun before obtaining one. The ownership ban is supposed to be effective because the ineligible person will be subject to criminal penalties if caught possessing a firearm. Although such a law will not keep guns out of the hands of high-risk users who lie about their status, it is hoped high-risk users will be deterred by the threat of criminal punishment.

The ownership bans represent some improvement over simply passing stiffer penalties for gun-related crimes because the law attempts to separate the potential criminal from his gun before he commits a crime with it. And if such laws could reduce the number of gun owners by excluding the people subject to the prohibition, they would indeed reduce violence.

But it is no easy matter to keep guns away from a group of "bad guys" while

HUNTING WITH GUNS became an American pastime long before the 1850s when lithographer Frances F. Palmer created "Woodcock Shooting" (*above*) for the publishing firm of Currier & Ives. During the past 140 years we have not changed our attitudes toward the legitimate uses of guns, but we have grown more concerned about gun-related violence.

allowing a larger group of "good guys" to own millions of them. Furthermore, bans are not very effective when the purchaser does not have to prove his eligibility as a gun purchaser. For instance, a convicted felon, who is prohibited from acquiring a gun under federal law, can easily obtain one if he is willing to lie about his record.

To prevent high-risk groups from freely obtaining firearms, many state and local governments now attempt to identify ineligible gun buyers before they acquire weapons. The screening system included in the Brady bill permits the police to determine whether a prospective gun purchaser has a criminal record. If the check turns up nothing, or even if it is not done, the purchaser can obtain the gun.

Screening systems are more effective than simple bans because screening prevents some high-risk users from obtaining guns even if they are willing to break the law. But because the screening systems mean additional costs and delays to all who wish to buy firearms, gun-owner groups usually oppose screening.

Screening systems and ownership prohibitions have a limited effect on homicides because most are committed by people who were not previously convicted of a felony and therefore who can purchase a gun legally. Still, screening systems are helpful to the extent that they keep firearms away from convicted felons and minors.

Many systems of owner screening suffer from another limitation. They affect only the purchase of weapons from commercial retailers. The measures do not apply to transactions between private citizens, who are continually selling and passing on some 35 million handguns. Prospective criminals can easily steal or purchase these guns.

To prevent the transfer of guns from private citizens to ineligible users, a few states require owners of firearms to register their weapons. In such schemes, owners must give information identifying each firearm they own. Gun accountability works like car registration, whereas owner screening is analogous to driver's licenses. But owner licensing schemes will have a limited impact on violence because they leave a large arsenal of firearms to which potential criminals have easy access.

The dangers of widespread access to firearms can be alleviated, in part, by regulating guns that are particularly dangerous. Laws that limit the supply of high-risk weapons can complement the strategy of decreasing high-risk uses and users. Supply-reduction laws strive to make the most dangerous guns so scarce that potential criminals cannot obtain them easily. The laws are justified on the grounds that the harm caused by owning and using these particularly dangerous firearms is greater than the benefit, even in the noncriminal population.

Supply-reduction laws usually deny citizens the opportunity to own a high-risk firearm unless they prove a special need for the weapon or belong to an exempt group, such as the military or the police. The idea behind supply reduction is that gun-related violence will decrease significantly only if the firearms most often used in violence are not available to most people.

The first weapons singled out for high-risk classification were machine guns and sawed-off shotguns as specified in the National Firearms Act of 1934. More than three decades later federal law put special restrictions on destructive devices of military origin and forbade the importation of handguns that were classified as unsuitable for sporting purposes. Several current proposals are designed to prohibit civilian ownership of military assault weapons.

Some state and local governments also single out the handgun as a high-risk firearm and require special licenses to purchase handguns. The system in New York City is the most prominent example of this approach. City residents must demonstrate a special need for a handgun, whereas in permissive licens-

Strategies for Gun Control

STRATEGY	COST AND INCONVENIENCE	EFFECT ON LEGITIMATE USES OF GUNS	POTENTIAL NUMBER OF LIVES SAVED	PROBLEMS
STOP DANGEROUS USES Extra punishment for crimes commited with guns and laws that regulate the proper use of guns	MINOR	MINOR	FEW	Weak threat, no logistical barrier to prevent gun misuse
STOP DANGEROUS USERS Laws that prohibit convicted felons from buying guns	MINOR	MINOR	FEW	Ineffective against those willing to lie about criminal record
Gun licenses for those who are not convicted felons and gun registration	SOME	SOME	SOME	Difficult to monitor tens of millions of guns
REDUCE SUPPLY OF FIREARMS Restrictions on ownership of exotic firearms	MINOR	MINOR	SOME	Little control over guns obtained illegally
Restrictions on ownership of handguns	MAJOR	MAJOR	MANY	Costly, interferes with legitimate uses, little control over guns obtained illegally

ing systems only high-risk users are disqualified. Gun-control supporters frequently advocate restricting the supply of handguns on a nationwide basis.

But a decrease in handguns will reduce firearm violence only if other guns are not substituted. It has been difficult so far to determine what effect restrictive licensing has had. Cities and states that have such licensing schemes have found that long guns are not replacing handguns as crime weapons. But one reason for this lack of substitution is that the laws have failed, so far, to reduce the supply of handguns significantly.

It is all too easy to move guns across state and city borders from areas where guns are easily available to places where firearm supplies are regulated. Law enforcement agencies in Massachusetts and New York City have shown that, in jurisdictions that have tight controls, more than 80 percent of guns confiscated by police were originally acquired out of state. Current federal law is trying to inhibit the flow of firearms between states, but the problem remains substantial.

In most northeastern cities with restrictive handgun laws, the use of guns in violent crime is lower than that in other U.S. cities. In Washington, D.C., the Federal Bureau of Alcohol, Tobacco and Firearms strictly enforced federal laws for nine months in 1976, and it reported success in reducing the flow of firearms. But the effort was not sustained. Washington currently experiences high rates of gun-related violence despite a restrictive handgun law.

Analysts do not have sufficient information to make predictions about the impact of gun control in the U.S. Some have compared the U.S. with such countries as England, France, Germany and Holland. They have discovered that nations with much lower rates of firearm ownership have lower rates of crimes involving guns. But these Western nations also have much lower rates of violent crimes that do not involve firearms.

Indeed, it is the very high rate of violence in the U.S. that makes the costs of gun use so large. The U.S. has both a "crime problem" and a "gun problem," and each exacerbates the other. No Western nation has ever instituted strict controls under conditions similar to those in the U.S.

Although most citizens support such measures as owner screening, public opinion is sharply divided on laws that would restrict the ownership of handguns to persons with special needs. If the U.S. does not reduce handguns and current trends continue, it faces the prospect that the number of handguns in circulation will grow from 35 million to more than 50 million within 50 years. A national program limiting the availability of handguns would cost many billions of dollars and meet much resistance from citizens. These costs would likely be greatest in the early years of the program. The benefits of supply reduction would emerge slowly because efforts to diminish the availability of handguns would probably have a cumulative impact over time.

At the heart of the debate over handgun restrictions is a disagreement about the character of American life in the 21st century. Roughly half of Americans believe that strict handgun control is not worth the hardship of changing policy in the U.S. They assume that the weapons can remain a part of American life for the indefinite future. But just as many Americans see the removal of the current stockpile of handguns as a necessary down payment on the American future. They regard free availability of handguns as a severe threat to urban life. American policy on handgun control will ultimately depend on which of these attitudes prevails.

FURTHER READING

FIREARMS AND VIOLENCE IN AMERICAN LIFE: A STAFF REPORT SUBMITTED TO THE NATIONAL COMMISSION ON THE CAUSES AND PREVENTION OF VIOLENCE. George D. Newton and Franklin E. Zimring. U.S. Government Printing Office, 1969.

UNDER THE GUN: WEAPONS, CRIME, AND VIOLENCE IN AMERICA. James D. Wright, Peter H. Rossi and Kathleen Daly. Aldine de Gruyter, 1983.

THE CITIZEN'S GUIDE TO GUN CONTROL. Franklin E. Zimring and Gordon Hawkins. Macmillan Publishing, 1987.

CRIME IN THE UNITED STATES: UNIFORM CRIME REPORTS, 1990. Printed annually by the Federal Bureau of Investigation. U.S. Government Printing Office, 1991-282-076/45217.

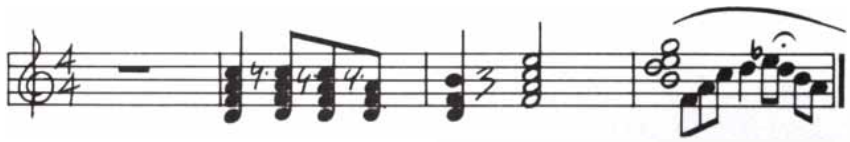
THE INJURY FACT BOOK. Susan P. Baker, Brian O'Neill, Marvin J. Ginsburg and Guohua Li. Oxford University Press (in press).

MAN: In high school, Sharon Simpson used to call me "Pinhead." So at the ten-year class reunion I drove up in my new Eagle Talon with all-wheel drive.

OFF-CAMERA VOICE: And what did she say?

MAN: "Nice car, Pinhead."

(MUSIC)

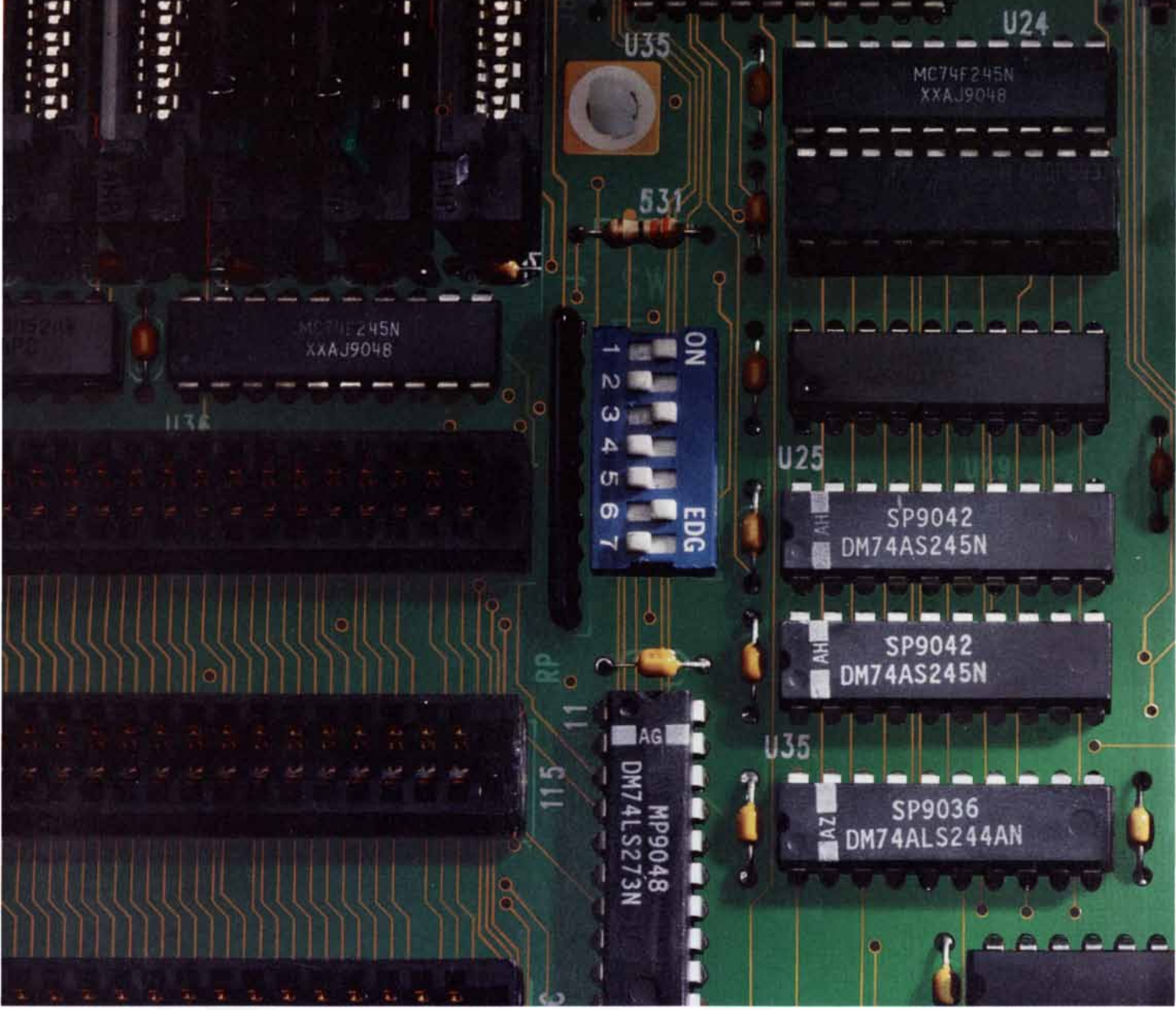


Advantage: Eagle

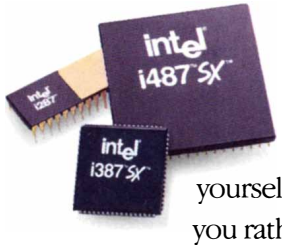
Eagle Talon TSi with available all-wheel drive and 195 horsepower turbocharged, intercooled engine is backed by Chrysler's new Owner's Choice Protection Plan for 1992 models. This plan includes a 3-year/36,000-mile bumper-to-bumper warranty with 5-year/60,000-mile powertrain protection, or a 12-month/12,000-mile basic warranty with 7-year/70,000-mile powertrain protection. Your choice. And with no deductible. See limited warranties, restrictions, and details at dealer. Excludes normal maintenance, adjustments, and wear items. For more information about Eagle Talon, or how to buy or lease one, call 1-800-JEEP-EAGLE. Buckle up for safety.

OWNER'S CHOICE PROTECTION PLAN 7/70 POWERTRAIN, OR 3/36 BUMPER-TO-BUMPER

OFFICIAL SPONSOR OF THE 1992 U.S. OLYMPIC TEAM 36 USC 380

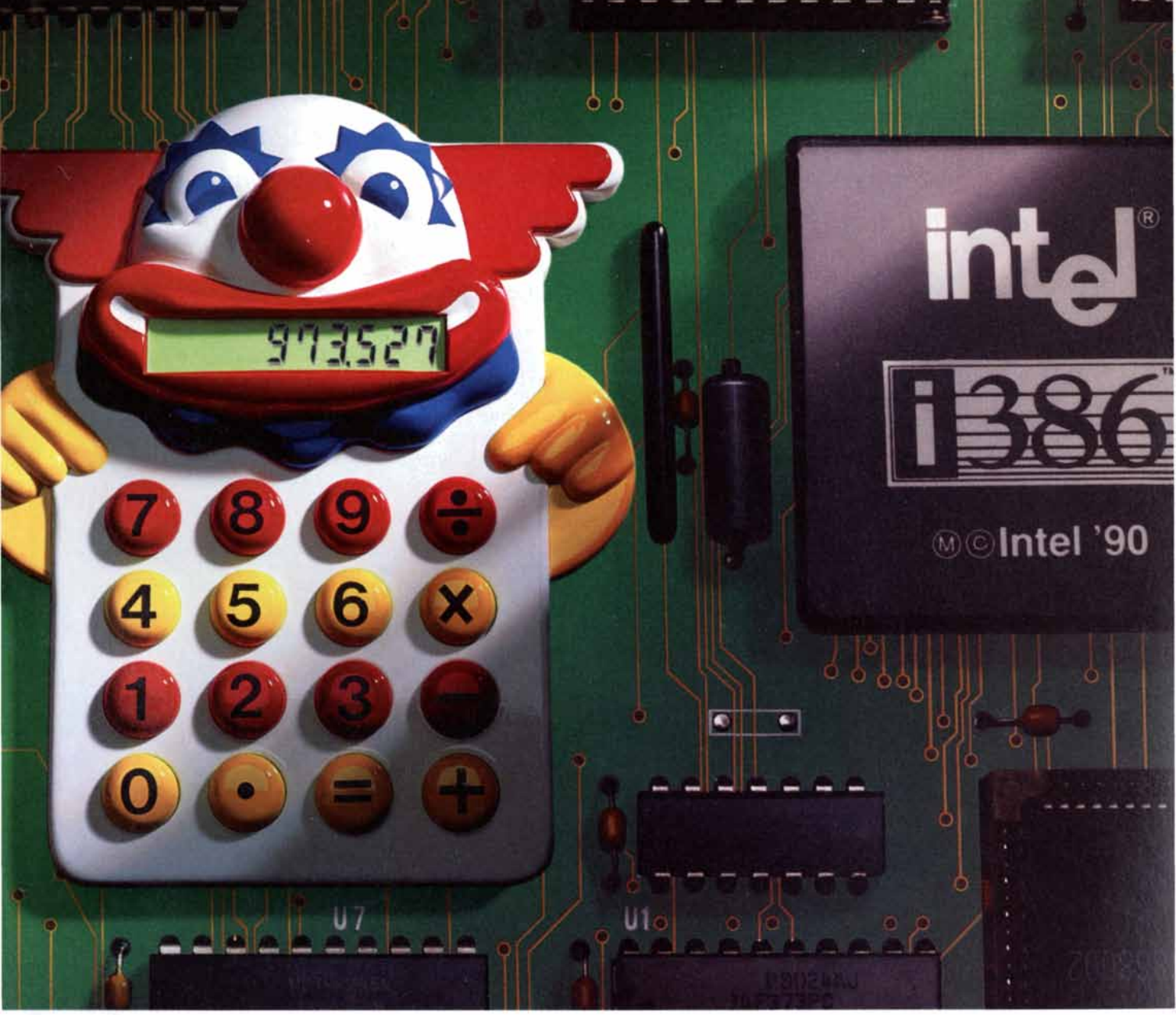


Ask for genuine Intel or who knows what yo



If you need a math coprocessor to speed your power applications, ask yourself this question: Which would you rather have sitting next to your Intel microprocessor — an Intel Math CoProcessor or something you may know nothing about? Because if you don't specify Intel, that's basically what you're getting — a big question mark. With Intel, however, there's simply no question. You're getting quality.

That's because Intel has the longest track record with math coprocessors. In fact, we've manufactured



Math CoProcessors, u'll have to count on.

and sold millions more than all the others combined. And we've tested every one of them with the most exhaustive battery of tests in the industry. All to assure you absolute reliability.

So ask for Intel Math CoProcessors. Or there's no calculating what you'll end up with.

For a free information packet, including our new low prices, call **(800)538-3373.**

intel®

The Computer Inside.™

The Nuclear Equation of State

By describing how nuclear matter behaves at various temperatures and pressures, the equation offers a glimpse of new phases of matter and conditions inside a supernova or during the big bang

by Hans Gutbrod and Horst Stöcker

No container on the earth can withstand the force of an exploding star. A successful effort to reproduce the big bang would destroy all matter as we know it—including the experimenters. Scientists who study astrophysics and cosmology therefore must often inform observation with speculation about how matter behaves under extreme pressure and at high temperatures and densities. Indeed, matter under such conditions can become quite exotic, perhaps consisting solely of a large number of free quarks. But by using powerful accelerators and sophisticated theoretical models, physicists at leading research centers around the world are exploring such ultradense matter quantitatively. Specifically, they are probing its nuclear equation of state.

An equation of state describes how the pressure of matter varies under different regimes of density and temperature. In the world more accessible to our senses, the equation details the macroscopic behavior of the four states, or phases, of matter. For instance, at sea level, where the pressure is one standard atmosphere, water becomes a solid below zero degrees Celsius and turns into a vapor above 100 degrees C. At still higher temperatures, the water vapor dissociates into its constituent hydrogen and oxygen ions, or atoms stripped of their electrons, forming a state known as plasma.

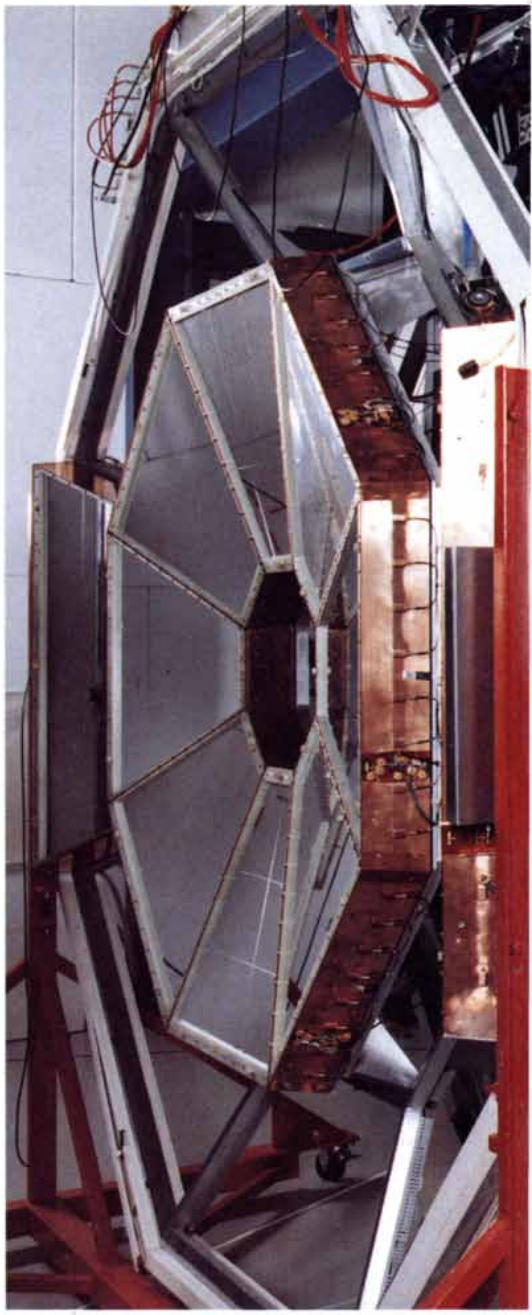
Nuclear matter exhibits similar "phases." By analogy, massive nuclei, which are made up of many nucleons (protons and neutrons), would represent the liquid phase. A solid phase could be represented by condensates of the nuclear fluid, such as particles called density isomers. The vapor phase would consist of a dilute gas of light nuclei and free nucleons. An assembly of free quarks and gluons corresponds to the plasma phase. Quarks are the fundamental building blocks of the class of particles called hadrons. Hadrons include protons and neutrons (made up of three quarks each) and other, less familiar, particles, such as pions (which consist of a quark and an antiquark). Gluons bind quarks together.

Simple experiments using ordinary laboratory apparatus can describe the phases of water and hence determine its equation of state. Such data are perhaps best expressed in terms of a phase diagram, which plots density against temperature. The phases of water are bounded by curves, which include such data points as zero and 100 degrees C at one atmosphere.

The different phase transitions in the nuclear equation of state can be represented by a similar plot. The exact values for the boundary curves, however, are much more difficult to determine than are those for water. Nucleons cling strongly to one another, "boiling" only when the temperature reaches 50 to 100 billion kelvins [see "Hot Nuclear Matter," by Walter Greiner and Horst Stöcker; SCIENTIFIC AMERICAN, January 1985].

To probe the nuclear equation of state—that is, the dependence of the pressure on the temperature and density—researchers first developed theoretical models. Such mathematical de-

HANS GUTBROD and HORST STÖCKER approach nuclear physics from an experimental and theoretical perspective, respectively. Gutbrod, a scientist at the European laboratory for particle physics (CERN) near Geneva and at the Institute for Heavy-Ion Research (GSI) in Darmstadt, received his Ph.D. from the Max Planck Institute for Nuclear Physics in Heidelberg. He has been involved in a number of key developments in nuclear physics, including the discoveries of flow phenomena and nuclear shock waves. Stöcker is professor of physics at the University of Frankfurt, where he received his Ph.D. He holds a joint appointment at the GSI and serves as a consultant to the Lawrence Berkeley Laboratory, the Brookhaven National Laboratory and the Lawrence Livermore National Laboratory.



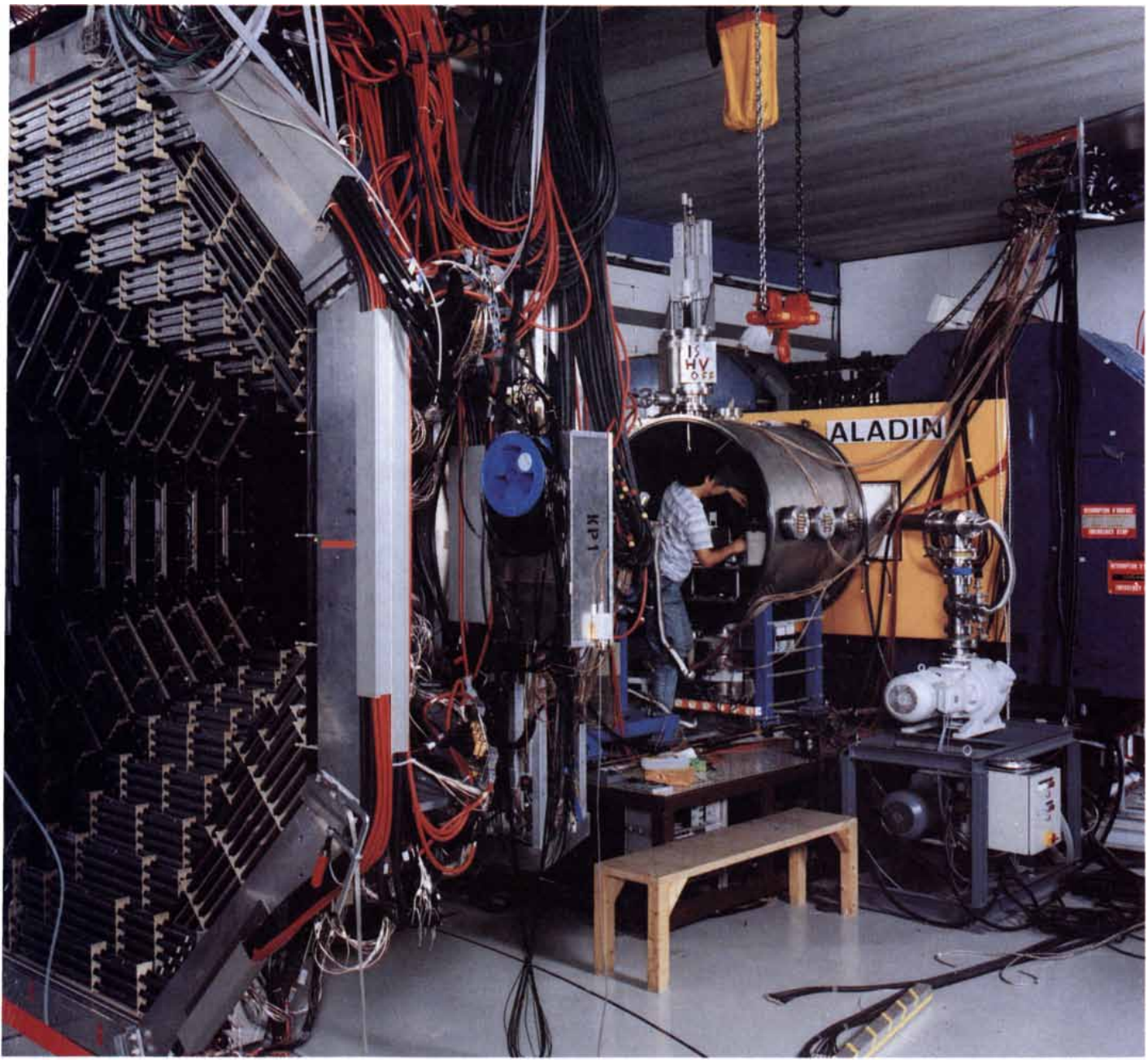
PARTICLE DETECTOR at the Institute for Heavy-Ion Research records hundreds of collision events each second. The particles produced by a beam of

scriptions usually start with the energy of each proton and neutron in the nucleus. The energy depends on the density and temperature within the nucleus. The pressure in the system is expected to rise as the density and the energy per particle increase.

Such straightforward model calculations are based on the one and only point of the equation of state precisely known so far. This point is at the ground state of the nuclei, where the pressure and temperature are zero and the nuclear density is about 3×10^{14} grams per cubic centimeter. The calculations then extrapolate to finite temperatures and densities, assuming certain parameters that remain constant.

But these parameters may not hold under extremely hot and dense (or even dilute) conditions, such as those that prevail when some stars die. As the stellar material collapses inward, the nuclear density at the center begins to rise, eventually exceeding the density within a typical nucleus by a factor of five to 10. The infalling matter hits the hard core and forms what will be the newly born neutron star. At the same time, a shock wave travels outward and results in an outburst of matter and light from the star, observed from the earth as a supernova [see "How a Supernova Explodes," by Hans A. Bethe and Gerald Brown; SCIENTIFIC AMERICAN, May 1985].

Hans A. Bethe of Cornell University, Gerald Brown of the State University of New York at Stony Brook and Sid Kahana of Brookhaven National Laboratory in Upton, N.Y., and their colleagues



heavy ions colliding with a target (*not shown*) travels from left to right. The two octagonal structures (*left*) detect fragments emitted perpendicularly with respect to the direction in which the beam of particles travels. The beam enters the Aladin spectrometer (*right*), shown here being prepared for an experiment. Aladin detects the particles in the beam that have simply glanced off the target. The measurements help researchers to decipher the nuclear equation of state.

argued that powerful supernova explosions could be achieved only if the nuclear equation of state is extremely "soft"—that is, only if nuclear matter is rather easily compressed, so that an increase in density would result in only a slight increase in pressure.

But can a supersoft equation of state stabilize neutron stars and thus prevent them from becoming black holes? Although neutron stars with fairly low masses have indeed been observed, Norman K. Glendenning of the Lawrence Berkeley Laboratory has demonstrated that the nuclear equation of state must be much stiffer—that is, the pressure and the energy per particle must rise more quickly than simplified supernova calculations would indicate.

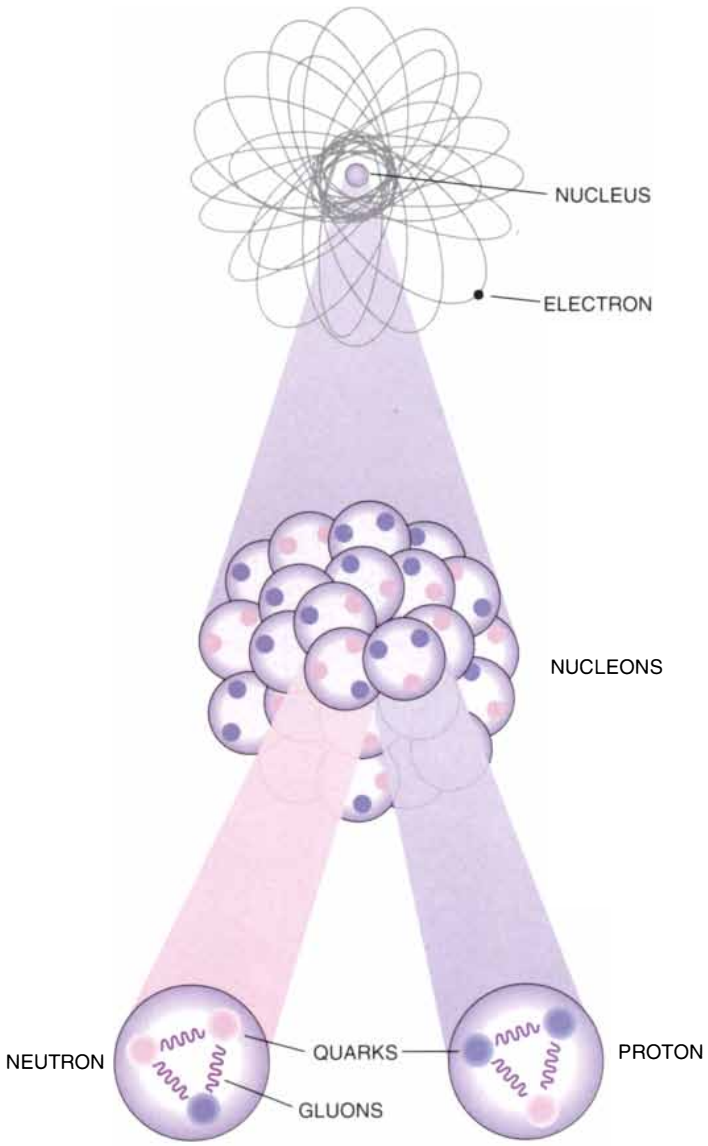
Clearly, the nuclear equation of state in this and other situations would remain elusive without experimental data. How can a laboratory produce conditions similar to those that occur when stars die or those that prevailed in the universe shortly after the big bang? The answer is to produce a "little bang."

Efforts to create such a dense, hot state of matter began in the late 1960s and early 1970s. At that time, Werner Scheid and Walter Greiner and their colleagues at the University of Frankfurt, as well as George F. Chapline, Michael H. Johnson, Edward Teller and Morton S. Weiss of the Lawrence Livermore National Laboratory, offered a proposal. The collision between two heavy nuclei accelerated to a speed of more than about one quarter the speed of light can generate a high-density shock wave. Such a shock wave would compress nuclear matter, achieving densities and temperatures high enough to test and improve predictions about the nuclear equation of state.

Because the researchers imagined that nuclei were much like droplets of water, their description is called the hydrodynamic model. A collision between nuclear droplets essentially boils them into the vapor phase, which comprises many particles and light nuclei collectively referred to as fragments.

The Structure of Matter

Atoms, which constitute all matter, consist of electrons and nuclei. Nuclei are built up of protons and neutrons, collectively called nucleons. Nucleons consist of quarks. Two up quarks (each with +2/3 electric charge) and a down quark (-1/3 charge) make up the positively charged proton; the neutron, a neutral particle, consists of an up quark and two down quarks. Gluons bind quarks together.



The first and for a long time the only experimental device to test the idea of creating the nuclear shock waves was the Bevalac, built in 1974 at the Lawrence Berkeley Laboratory (LBL). The Bevalac imparts energy to a beam of ions, accelerating them. Subsequent improvements in technology have enabled the Bevalac to accelerate ions to energies between one and two billion electron volts (GeV), depending on the mass of the nuclei. Such energies correspond to nearly 95 percent the speed of light, far exceeding the speed of sound. The accelerated beam then collides with nuclei in a stationary target. The resulting fragments provide information about the temperature and pressure in the collision—and thus about the nuclear equation of state.

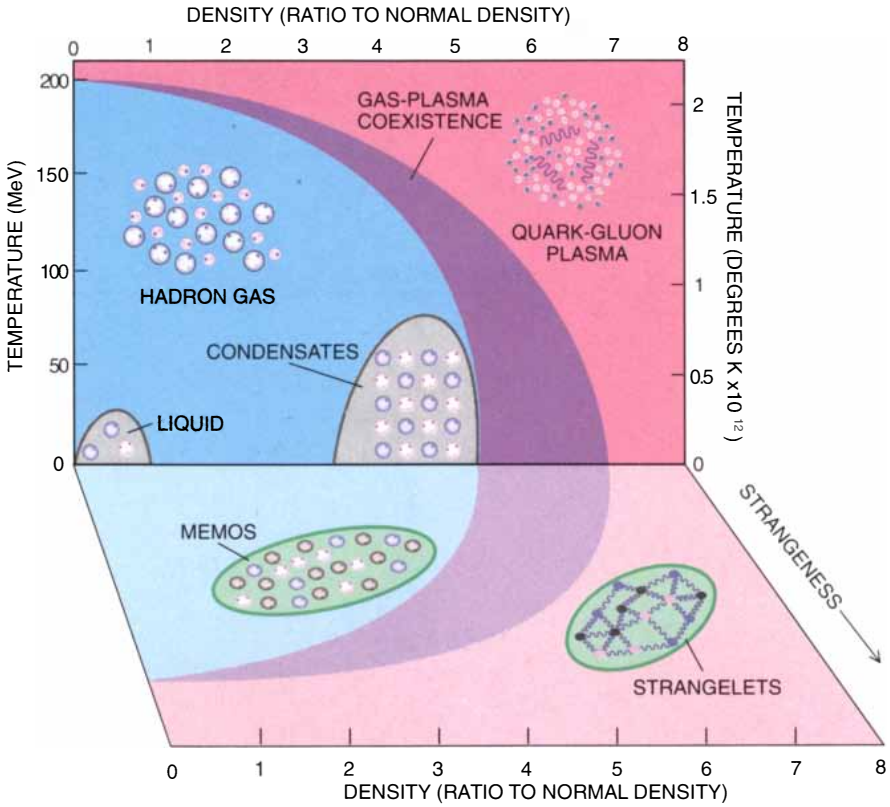
Photographs made during the collision provided the initial means of detecting the fragments. The images consisted of a number of tracks left by particles as they travel through a medium [see "Tracking and Imaging Elementary Particles," by Horst Breuker, Hans Drevermann, Christoph Grab, Alphonse A. Rademakers and Howard Stone; SCIENTIFIC AMERICAN, August]. Erwin Schopper of the University of Frankfurt, Harold H. Heckman of the University of California at Berkeley and Ingvar Otterlund of the University of Lund in Sweden led several experimental groups in pioneering this research technique.

Yet because analyzing data visually is rather slow, it has been difficult to achieve convincing, statistically significant results. Electronic detectors improved the speed of data analysis, but they could identify only one or few par-

ticles for every collision. In addition, they could not distinguish less violent (and less interesting) glancing impacts from head-on collisions. Investigators therefore developed more complex detection concepts. In particular, they found that one can identify central events by measuring the total number of emitted particles: if many particles result from the impact, then the projectile and target have collided head-on.

One instrument that can identify central collisions is the streamer chamber at LBL. Reinhard Stock of the University of Frankfurt, Rudolf Bock of the Institute for Heavy-Ion Research (GSI) in Darmstadt and the late Howard G. Pugh from the LBL headed a collaboration between the LBL and the GSI that pioneered the chamber's use. Several cameras photograph each interesting reaction. An analysis of the track curvature and thickness reveals the momenta and charge of each fragment.

A GSI-LBL group headed by Arthur M. Poskanzer, Hans-Georg Ritter and one of us (Gutbrod) focused on another instrument. We developed the first electronic detector in nuclear physics that recorded information over an entire spherical area. This detector, called the Plastic Ball spectrometer, collects data at a high rate, identifying the mass, charge, energy and angle of emission of nearly all the charged particles from many events.



PHASE DIAGRAM renders the nuclear equation of state in graphic form. Nuclear matter in its normal phase resembles a liquid. Increasing the temperature or density "boils" nuclei into the hadron gas phase. Under extreme density but low temperature, nucleons could become "frozen," forming condensates. Further heating or compression may produce the plasma phase, which would consist of free quarks and gluons. The gas and plasma phases may exist simultaneously over a wide region. Particles that have strange quarks, such as multistrange, metastable objects ("memos") and strangelets, may also form.

The capabilities of the Plastic Ball spectrometer turned out to be very important for the study of central collisions between massive nuclei. Some theoreticians had strongly criticized the predictions made by the Frankfurt school. They believed that the nuclei would not stop but instead would pass through one another. Shock waves and high-density matter would therefore not form in the laboratory.

But in early experiments, the Plastic Ball group observed that nuclei do indeed stop each other at energies as high as 2 GeV per nucleon. The clearest proof came in 1990, when analysis of data taken five years before showed that initial projectile and target nuclei completely disappeared; the resulting pile of nuclear debris was virtually at rest where the center of mass of the two nuclei would be expected to lie.

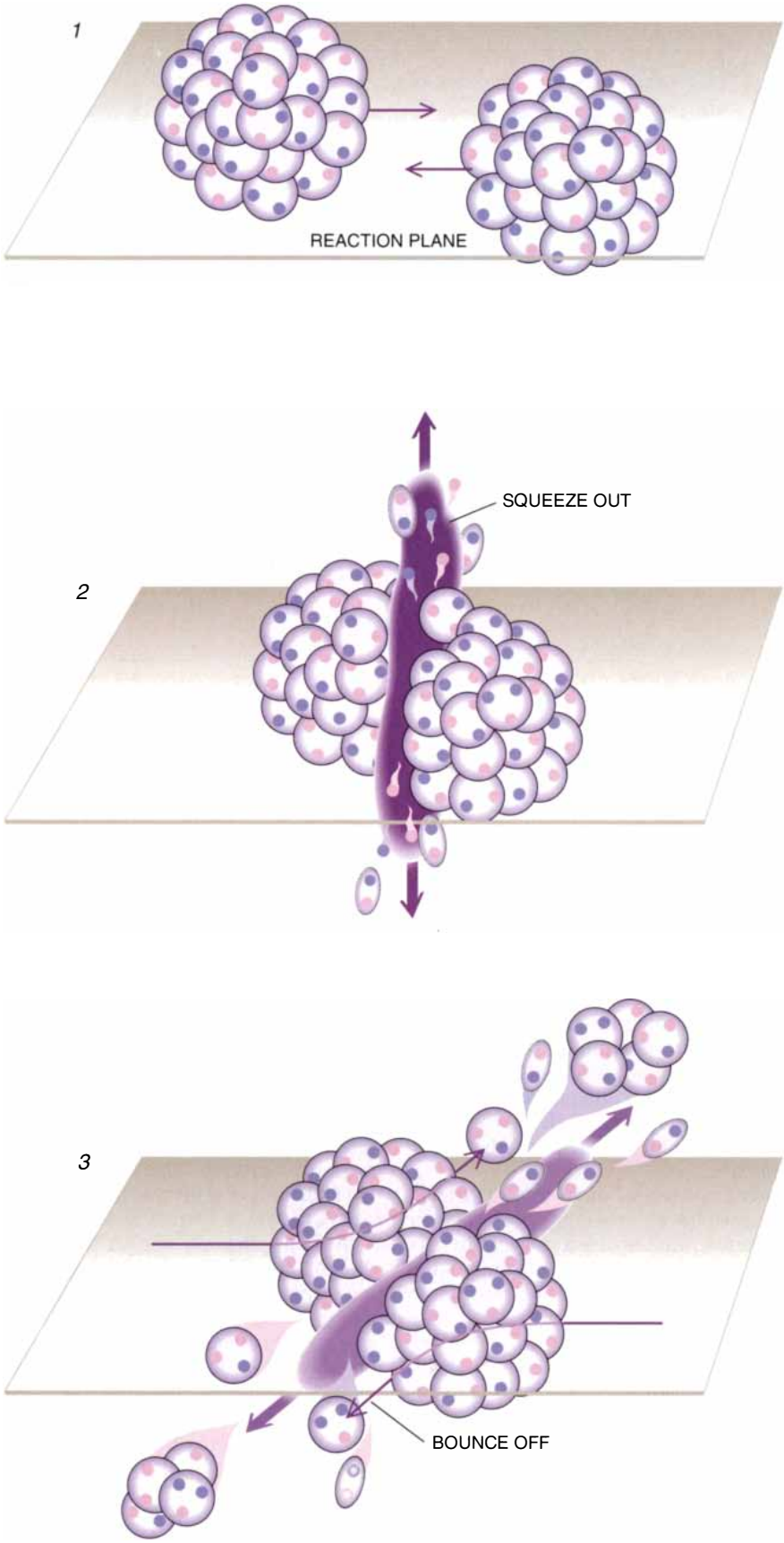
The formation of nuclear shock waves should lead to a second observable outcome: a collision would produce substantial numbers of pions. Pions belong to a family of moderately massive particles called mesons; they bind nucleons to one another. In the 1970s Scheid, Greiner and one of us (Stöcker) suggested that if the hydrodynamic shock-wave model is correct, then the yield of pions should rise linearly as the bombarding energy increases. The rate of increase would be a direct measure of the nuclear equation of state. The GSI-LBL streamer chamber group tested this theory in the early 1980s and indeed observed the predicted linear rise.

The number of pions and their energy spectra captured by the streamer chamber data enabled the researchers to calculate the effective temperature in the stopped heap of nuclear matter. Both sets of calculations yielded temperature values of up to 1.5×10^{12} kelvins (150 million electron volts [MeV]) in the collision zone, the highest nuclear temperatures ever measured in the laboratory. Plotting the temperature against the density of the colliding nuclei would yield a phase diagram, which characterizes the nuclear equation of state. We have not yet been able to measure directly the density reached in the collision, however. Theoretical corrections substantially complicate quantitative determination from these data.

Using the hydrodynamic model, Gerd Buchwald, Gerhard Graebner and Joachim Maruhn of the University of Frankfurt formulated a third prediction that would prove that nuclear shock waves form: the fragments of the collided, stopped nuclei should move together in a sideward, "transverse" direction with respect to the original direction of motion. The more head-on the collisions, the greater the collective sideward flow.

Nucleons that do not hit the target directly should bounce off the dense, hot central region. They retain their motion in the plane of impact, much as a cue ball in billiards does after a glancing shot. This bounce-off motion acts as a barometer: it measures the pressure exerted by the dense, central matter. Investigators determine the pressure quantitatively by examining the sideward flow: the greater the flow, the higher the pressure.

Early Plastic Ball experiments used beams of calcium ions. The particles emitted from the collision of such light ions did not flow sideways, as theory had predicted. But experiments



COLLIDING NUCLEI provide the key mechanism for investigating the nuclear equation of state. Here two gold nuclei collide slightly off center (*1*). In the hot compression zone formed by the shock wave, matter composed of protons, pions and other light particles (perhaps quarks and gluons) is squeezed out at right angles to the reaction plane (*2*). The remnants of the two nuclei then bounce off each other (*3*).

with heavier beams, consisting of niobium and gold, revealed sharp, sideward patterns, even in off-center collisions. The data confirmed the hydrodynamic model calculations.

Still, the dependence of the data on the mass of the projectile and target—in particular, the observed small flow in the calcium system—needed explanation. Arnold R. Bodmer, now at the University of Illinois at Chicago, and Bernd Schürmann of the University of Munich suggested that nonequilibrium effects, such as nuclear viscosity and a small but finite mean free path (the average distance a particle travels before colliding with another particle), might be responsible.

Hans Kruse, Barbara V. Jacak, Joseph Molitoris and Stöcker, then at Michigan State University, first demonstrated that this is indeed quantitatively possible. Other theorists later generalized the model, treating the nucleons as small wave packets that move under the mutual influence of other nucleons. The generalizations made by these theorists—Jörg Aichelin of the University of Heidelberg and Albrecht Rosenhauer, Georg Peilert, Christoph Hartnack, Maria Berenguer, Walter Greiner and Stöcker of the University of Frankfurt—are referred to as the quantum molecular dynamics model.

Like the hydrodynamic model, the quantum molecular dynamics model predicts collective sideward flow. But the model also explains why heavy nuclei create more sideward flow than do lighter ones. Compression in the shock zone is greater when heavy ions collide than it is when light ions do. As the pressure in the shock zone increases, so does the force the shock exerts on the surrounding matter from which the fragments form. An observer sees the momentum transfer as an increased flow of fragments in the scattering plane. Plastic Ball experiments analyzed by Karl-Heinz Kampert of the University of Münster have confirmed the theory propounded by Peilert and his colleagues.

The Frankfurt theorists anticipated a fourth experimentally observable effect in the collision of nuclei: matter in the central portion should be squeezed so strongly that some of it should be ejected in a direction perpendicular to the scattering plane. The phenomenon is called off-plane squeeze out. In 1989 Kampert and Hans-Rudi Schmidt of the GSI analyzed Plastic Ball data and confirmed the phenomenon's existence. In addition, they found that the magnitude of the squeeze out increases linearly as the mass of the colliding nuclei rises. The result supports

the prediction made by Hartnack and his co-workers and proves unambiguously that the central, compressed region is sufficiently dense to study the equation of state.

As we anticipated, the compressed nuclear matter has yielded much information on the hot, dense nuclear medium. Current data indicate that the equation of state seems to "stiffen" continuously: the energy of each nucleon rises as density increases. The temperatures achieved are so high that heavy nuclei do in fact boil and become a vapor of light nuclei and free nucleons. These results have been obtained only for moderately high densities, corresponding to maximum energy densities equal to about two to three times that in normal nuclear matter. But a complete description of the nuclear equation of state—in particular, the exploration of the plasma phase—requires observations over much greater energy densities than existing accelerators can achieve.

Very heavy nuclei must collide at energies of much more than 2 GeV to create the plasma phase. Such high energies are sometimes referred to as ultrarelativistic, because the velocity of the nuclei approaches the speed of light. Calculations must then consider the effects of relativity. Even the GSI's new Schwerionen Synchrotron (SIS) accelerator, the most powerful heavy-ion accelerator operating today, does not come close to the ultrarelativistic regime. In the summer of 1990, SIS successfully accelerated ions of gold—one of nature's heavier elements. The device can produce densities that exceed normal nuclear density by a factor of two to three. Yet this density is probably only about one half to one quarter of that needed to mimic the matter of collapsed stars.

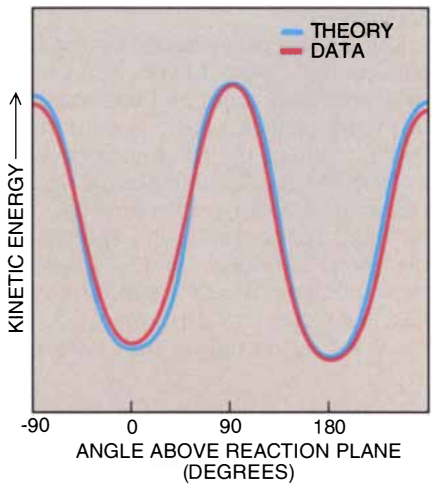
In the plasma state we hope to create at ultrarelativistic energies, quarks and gluons would no longer remain confined to the interior of hadrons but instead would propagate over distances the size of the nucleus. In 1978 Chapline and Michael Nauenberg of the University of California at Santa Cruz and, independently, Walter Greiner and Stöcker predicted that the extreme densities and temperatures necessary for quark matter production can be reached in heavy-ion collisions of about 10 GeV per nucleon, but only for a fleeting moment.

Workers would like very much to collide very heavy nuclei at such energies. We have a poor understanding of what happens when two nucleons (or more generally two hadrons) collide so violently. The tremendous amount of kinetic energy in an inelastic collision not only should show up in the energy and momentum of the originally colliding constituents but also should create particle and antiparticle pairs. For instance, if the kinetic energy exceeds 270 MeV, then a positive and a negative pion might form.

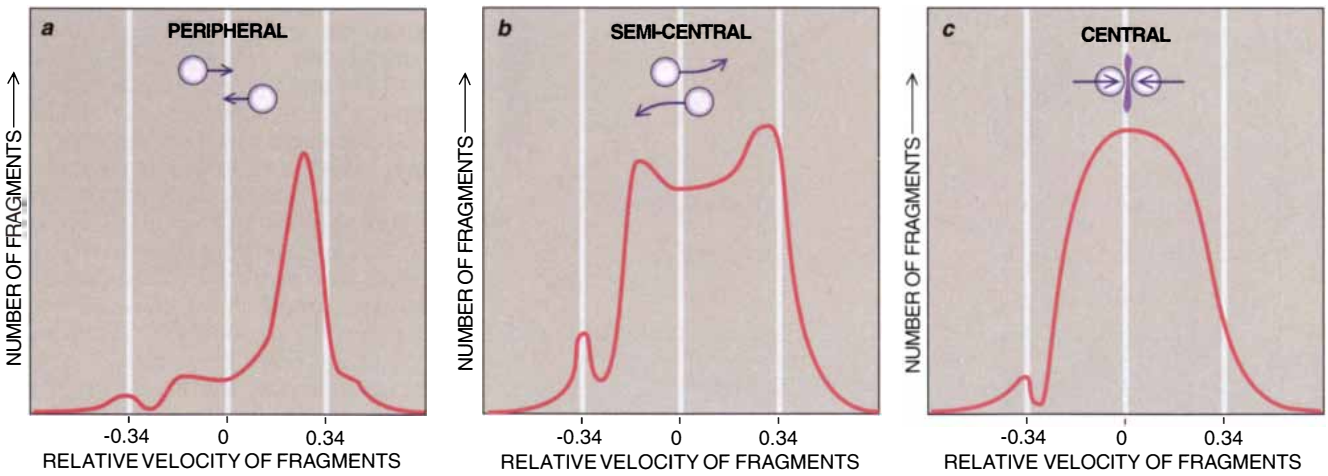
To describe such ultrarelativistic collisions, theorists at the University of Lund and the University of Saclay have envisioned the collision of hadrons as the excitation of two massive strings rather than the collision of liquid drops. The strings extend like rubber bands between the quarks that have been hit and those that have not. The string stores potential energy as the end points move apart at high velocity.

Eventually, the strings break and fragment into quark-antiquark pairs. The number of such pairs reflects the transfer of energy and momentum in the collision. The quark pairs quickly develop into hadrons (predominantly pions). In a nuclear collision, many strings and hadrons might exist simultaneously. If so, the system could develop into the quark-gluon plasma. Hadrons would then form much as they did shortly after the big bang: they would condense out of the quark-gluon plasma in a phase transition only after expansion reduced the energy density—that is, when the system cooled.

To predict the experimental outcomes of the relativistic heavy-ion collisions, Heinz Sorge, Andreas von Keitz, Raffaele Mattiello, Andre Jahns, Luke Winkelmann, Thomas Schönfeld, Walter Greiner and Stöcker of the University of Frankfurt developed a relativistic extension of the quantum molecular dynamics model. They pre-



SQUEEZE OUT is a phenomenon that has only recently been confirmed by experimental data. It manifests itself as sharp increases in the kinetic energy of the particles emitted 90 degrees above and below the reaction plane.



SEVERITY OF IMPACT determines if two nuclei (each here moving at 0.34 the speed of light) stop each other. The velocities of the resulting fragments tend toward zero as collisions become more direct (*a* and *b*). In a head-on collision, virtually all the mass is at rest and concentrated where the center of mass of the two nuclei would be expected to lie (*c*).

dicted that experimenters should observe substantially more stopping and much higher densities in collisions between heavy nuclei, such as gold ions, than they would in collisions between light ions—for example, those of silicon. Their calculations also imply that collective flow should be observable even at these ultrarelativistic energies. The predicted bounce-off and squeeze-out effects would also provide details on the equation of state of the quark-gluon plasma: they could also serve here as barometers to measure the pressure in the ultradense quark matter.

The theoretical models have achieved an extraordinary level of refinement. In addition to being able to predict experimental outcomes, they can be used to describe the transient stages of a collision, especially the early phases when the matter is most strongly heated and compressed.

But existing accelerators cannot impart enough energy to very heavy ions (which contain about 200 nucleons) to form the plasma phase. Even if they could, many of the conditions needed to form the quark-gluon plasma would continue to challenge experimental verification. For instance, the spectra of the pions produced in the collision should signal plasma formation. But pions are subject to a complicated dynamic evolution during the expansion of the system, a process that could distort the information they proffer. Furthermore, researchers do not know how to determine the exact size of the reaction zone at maximum compression.

Because of the limitations of current heavy-ion accelerators, some physicists are searching for the plasma through the collision of light ions, which consist of about 30 nucleons. Measurements indicate energy densities of about 10 to 20 times the normal nuclear energy density—close to the critical energy density considered necessary to form the quark-gluon plasma.

But GSI-LBL results immediately raise the question of whether the quark-gluon plasma can form when light ions collide. Light ions are most likely not thick enough to stop one another and thereby form a state of dense matter. But even if quark-gluon plasma can form in light systems, it will expand quickly, cool and undergo a phase transition into hadronic matter. One must therefore find signals that come only from the transient existence of the quark-gluon plasma. Such signals would be more difficult to detect in light-ion systems than in heavy-ion ones because of high background noise.

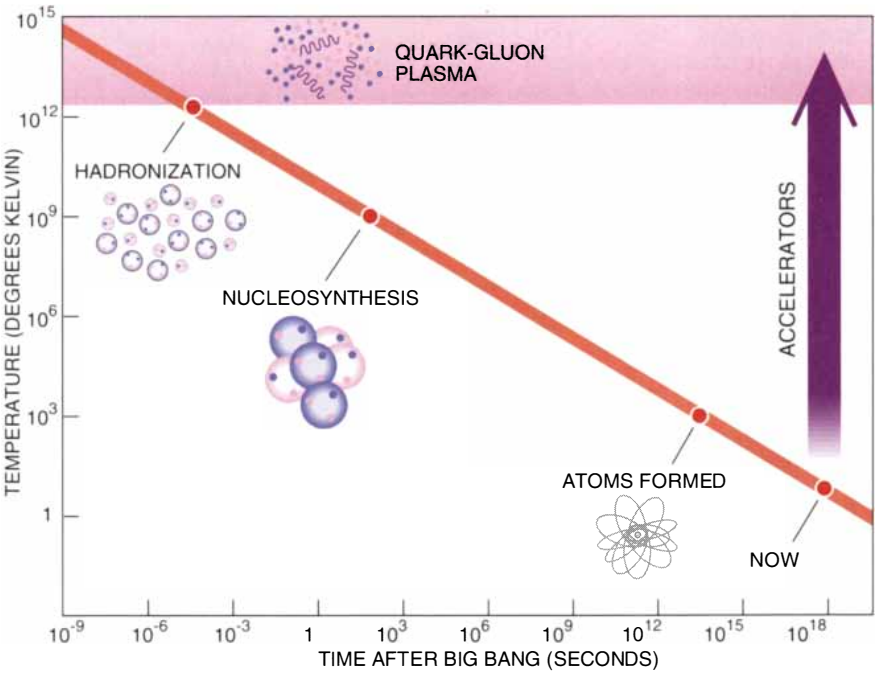
Some researchers would prefer to search for the quark matter phase via weakly interacting particles, such as photons or pairs of leptons (electrons, muons and their antimatter counterparts). The photons and leptons form in the very hottest part of the system and penetrate the surrounding material without reacting. They thus provide a direct signal of the initial plasma phase as described by the nuclear equation of state.

Such particles have offered some tantalizing hints of the production of the plasma phase. In 1988 a collaboration at CERN detected a dramatic drop in the production of muon-antimuon pairs, a plasma signal predicted by Tetsuo Matsui of the Massachusetts Institute of Technology and Helmut Satz of CERN. The particles result from the decay of certain hadrons. But, as shown by Sean Gavin and Miklos Gyulassy of LBL, the formation of an ultradense hadron fluid rather than the quark-gluon plasma can account for the observed drop-off. More conclusive evidence may come from workers at CERN, who are currently searching for the direct photons in collisions between a light nucleus (sulfur) and a heavy target (gold).

In the mid-1980s Peter Koch of the University of Regensburg, Johann Rafelski of the University of Arizona and Bernd Müller of Duke University proposed that another potential signal of the quark-gluon plasma may exist. If the plasma actually forms, then the abundant thermal gluons will most probably fuse to form many pairs of strange and antistrange quarks.

Because single quarks cannot exist in an unbound state, they would coalesce into abundant numbers of strange and antistrange elementary particles called kaons and hyperons. The enhanced production of such entities would signal the presence of a quark-gluon plasma.

Using the facilities at Brookhaven National Laboratory, a Japanese-American collaboration of workers from 10 different institutions recently reported enhanced strangeness production. The group bombarded a gold target with a 15-GeV-per-nucleon silicon beam. It found a strongly increased production of kaons. Further evidence of enhanced strangeness came from two international collaborations working at CERN. The groups, headed by Emanuele Quercigh of CERN and Stock, observed a large increase in the number of antihyperon particles produced. All these observations can be explained by the formation of the quark-gluon plasma. Theoretical studies by Mattiello, Sorge, Walter Greiner and Stöcker, however, show that the enhanced kaon yield and spectra can be understood if an ultradense hadron fluid rather than the quark-gluon plasma forms. The antihyperon production, however, has stumped many



TEMPERATURE OF THE UNIVERSE has been falling since the big bang. During the first microsecond, all matter is thought to have existed as quark-gluon plasma. As the universe expanded and cooled, more complex matter condensed out of the plasma, eventually forming the atoms observable today. Accelerators now under construction should be able to heat nuclei to 2×10^{12} kelvins (200 million electron volts [MeV]), perhaps creating the much sought after primordial quark matter.

PARTICLE TRACKS produced in a collision between two nuclei yield information about the dynamics of the interaction. Such head-on collisions (*left part of image*) produce hundreds of charged particles, mostly protons and pions.

theorists and remains an intriguing puzzle.

Other, even more exotic, particles may form from the plasma phase. Their presence would unambiguously confirm the existence of the plasma and provide a manifestation of the nuclear equation of state in a hitherto unexplored realm. For instance, the plasma phase might yield globules that consist of a large number of up and down quarks. In that case, one would speak of multiquark droplets. The droplets must be relatively massive; otherwise all nuclei would be unstable and decay. But if the droplets also contain strange quarks, they may be stable. Sui Chin and Arthur K. Kerman of M.I.T. and Larry D. McLerran and James D. Bjorken, then at Stanford University, proposed two reasons for the stability of such "strangelets."

First, the lowest available quantum states are occupied. Therefore, any potential transition to these states runs afoul of the Pauli exclusion principle, which holds that every quantum state can be occupied only by a single particle. (It is the quantum analogue of Archimedes' principle: where there is one body there cannot be another.) The exclusion principle suppresses the decay of a strange quark into an up quark.

The second explanation involves an increase in the binding energy of the system. Conceivably, strange quark matter could be bound strongly enough to become absolutely stable. This condition, proposed by Bodmer, Edward Witten of Princeton University and Edward Farhi and Robert L. Jaffe of M.I.T., would represent the true ground state of matter—that is, the zero point of the nuclear equation of state.

But such a conclusion depends on a particular choice of parameters in the unknown equation of state of quark matter. Other theories hold that a quark droplet may be only metastable. The droplet could decay slowly, emitting a nucleon and a pion. Even longer lifetimes could be possible if the energy of the strangelet is such that it inhibits the decay into a proton and a pion. Neither form of strangelet stability can be ruled out from present knowledge.

In 1987 Carsten Greiner of the University of Erlangen, Dirk Rischke of the University of Frankfurt, Koch and Stöcker discovered a mechanism that could produce strange quark matter droplets in heavy-ion collisions. These predictions have spurred collaborations headed by Jack Sandweis of Yale University, Henry J. Crawford of the LBL and P. Buford Price of the University of California at Berkeley, and a CERN group.

Furthermore, an abundant production of strangeness opens up the intriguing possibility of producing metastable exotic multistrange objects, or "memos." Such objects would consist predominantly of hyperons. Both memos and strangelets would be readily detected. Unlike normal nuclei, they are thought to be electrically neutral or negatively charged. As a result, their motions through a magnetic field would be easily distinguished from normal nuclei.

Physicists are looking forward to several exciting opportunities to probe the nuclear equation of state and the plasma phase. The SIS facility at the GSI has initiated a new research program. Scientists worldwide have even begun to discuss plans to create accelerators that can operate at ultrarelativistic energies. The Booster project at Brookhaven is being commissioned and will enable the Alternating Gradient Synchrotron (AGS) to collide gold ions at 11 GeV per nucleon by 1994. Brookhaven is also constructing the Relativistic Heavy-Ion Collider (RHIC). This device should be ready by 1997 and will accelerate gold ions to 20 trillion electron volts (TeV) per nucleon.

CERN has proposed building by 1994 an injector that accelerates lead beams to 200 GeV per nucleon. The laboratory's leadership also intends to install a new accelerator, the Large Hadron Collider (LHC), inside the Large Electron-Positron collider tunnel at CERN. There collisions between lead ions will have the equivalent of more than 12,000 TeV per nucleon. A completely unknown regime of hadronic matter would become accessible for study.

Even peripheral collisions at these energies hold great promise. Recently Mathias Grabiak, Gerhard Soff and Walter Greiner of the University of Frankfurt and others suggested such ultrarelativistic collisions of heavy ions can produce the searched for Higgs particle. Theorists believe the particle generates the masses of other fundamental particles, such as the W and Z bosons. Its discovery would provide the mathematical consistency necessary in the standard model of the forces in nature [see "The Higgs Boson," by Martinus J. G. Veltman; SCIENTIFIC AMERICAN, November 1986].

Clearly, the tremendous progress in studying the nuclear equation of state has been possible only through close collaboration among theorists, experimentalists and engineers. Future high-energy collision experiments may produce the plasma phase of nuclear matter as well as exotic particles never before observed.

FURTHER READING

FROM QUARKS TO THE COSMOS: TOOLS OF DISCOVERY. Leon M. Lederman and David N. Schramm. Scientific American Library, 1989.

NUCLEAR EQUATION OF STATE, Part A: DISCOVERY OF NUCLEAR SHOCK WAVES AND THE EOS. NUCLEAR EQUATION OF STATE, Part B: QCD AND THE FORMATION OF THE QUARK-GLUON PLASMA. Edited by Walter Greiner and Horst Stöcker. Plenum Publishing, 1990.

SIMULATING HOT QUARK MATTER. Jean Potvin in *American Scientist*, Vol. 79, No. 2, pages 116–129; March–April 1991.

ANNOUNCING A TIRE THAT MAY LAST AS LONG AS YOU OWN YOUR CAR.

**Our new all-season radial is backed by an 80,000 mile treadwear limited warranty. See dealer for details.*

Extra rubber, extra steel, gives you extra protection against road hazards.

Revolutionary new design ensures even wear for even higher mileage.

Rest easy. With the smooth, quiet, comfortable ride we're famous for.

Once again Michelin raises the bar of excellence. With the highest mileage passenger tire we've ever made.

The XH4 maintains Michelin's renowned all-season performance from the first mile to the last.

More miles. More value. More Michelin.

MICHELIN®
BECAUSE SO MUCH IS RIDING ON YOUR TIRES.®

CONGRATULATIONS, IT'S A MICHELIN.

BACKED BY AN 80,000 MILE WARRANTY.*

Amyloid Protein and Alzheimer's Disease

When this protein fragment accumulates excessively in the brain, Alzheimer's disease may be the result. Understanding how that fragment forms could be the key to a treatment

by Dennis J. Selkoe

Imagine surmounting life's great and small hurdles only to face at the end the relentless and devastating loss of one's most human qualities: reasoning, abstraction, language and memory. Such a fate now awaits millions of individuals in races and ethnic groups worldwide. The dramatic rise in life expectancy during this century, primarily through the cure of infectious diseases, has enabled many of us to reach an age at which degenerative diseases of the brain–particularly Alzheimer's disease–become common.

Most of us were raised on the notion that a grandparent who became confused and forgetful had "hardening of the arteries." Dementia, the failure of cognitive ability, was widely assumed to be a natural accompaniment of old age. But neuropathological studies during the past two decades have shown that the brain lesions first described by the Bavarian psychiatrist Alois Alzheimer in 1907—senile plaques and neurofibrillary tangles—are the most common basis for late-life dementia in many developed countries.

DENNIS J. SELKOE is co-director of the Center for Neurologic Diseases at Brigham and Women's Hospital in Boston and professor of neurology and neuroscience at Harvard Medical School. He received his bachelor's degree from Columbia University and his medical degree from the University of Virginia. After research training at the National Institutes of Health, Selkoe completed a fellowship in biochemistry and neuronal cell biology at Harvard Medical School and Children's Hospital in Boston. In 1978 he founded an independent research laboratory for the study of Alzheimer's disease and related neurobiological phenomena. In 1988 the NIH presented Selkoe with its Leadership and Excellence in Alzheimer's Disease Award.

The loss of memory, judgment and emotional stability that Alzheimer's disease inflicts on its victims occurs gradually and inexorably, usually leading to death in a severely debilitated, immobile state between four and 12 years after onset. The cost to American society for diagnosing and managing Alzheimer's disease, primarily for custodial care, is currently estimated at more than $80 billion annually. No treatment that retards the progression of the disease is known.

How should scientists interested in deciphering and ultimately blocking this complex disorder begin their attack? The answer appears increasingly to lie in understanding the genesis of its hallmark: the so-called senile plaques that occur in huge numbers within the cerebral cortex, the hippocampus, the amygdala and other brain areas essential for cognitive function. Recent discoveries indicate that such work is indeed providing powerful clues about the earliest events in the mechanism of the disease.

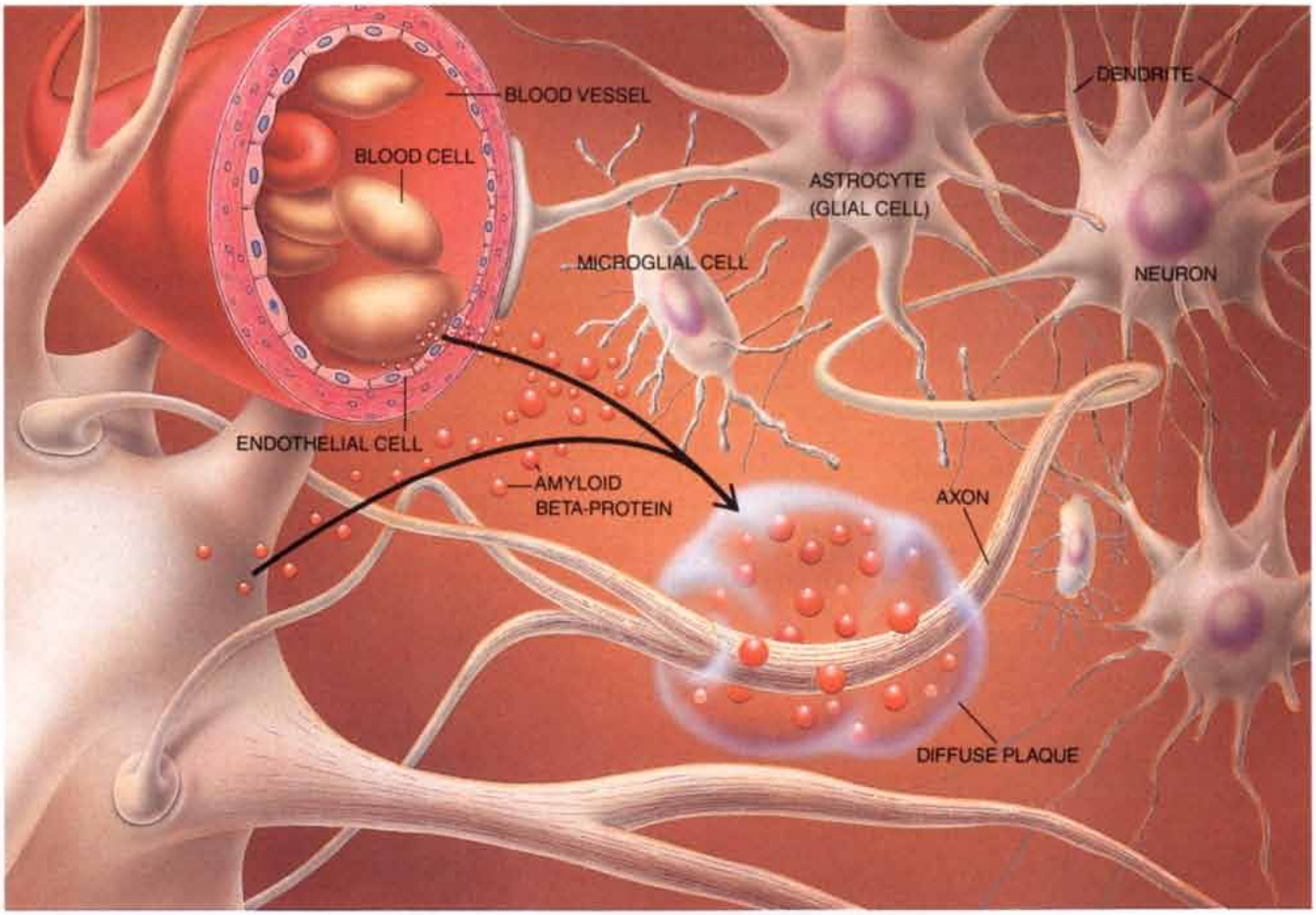
Peering through the microscope at the brain of his first patient, Alzheimer wrote prophetically: "Scattered through the entire cortex, especially in the upper layers, one found miliary foci that were caused by the deposition of a peculiar substance in the cerebral cortex." Evidence emerging from many laboratories during the past seven years indicates that this "peculiar substance" is a protein fragment, approximately 40 amino acids long, referred to as the amyloid beta-protein. It arises from cleavage by enzymes of a much larger precursor protein encoded by a gene on human chromosome 21.

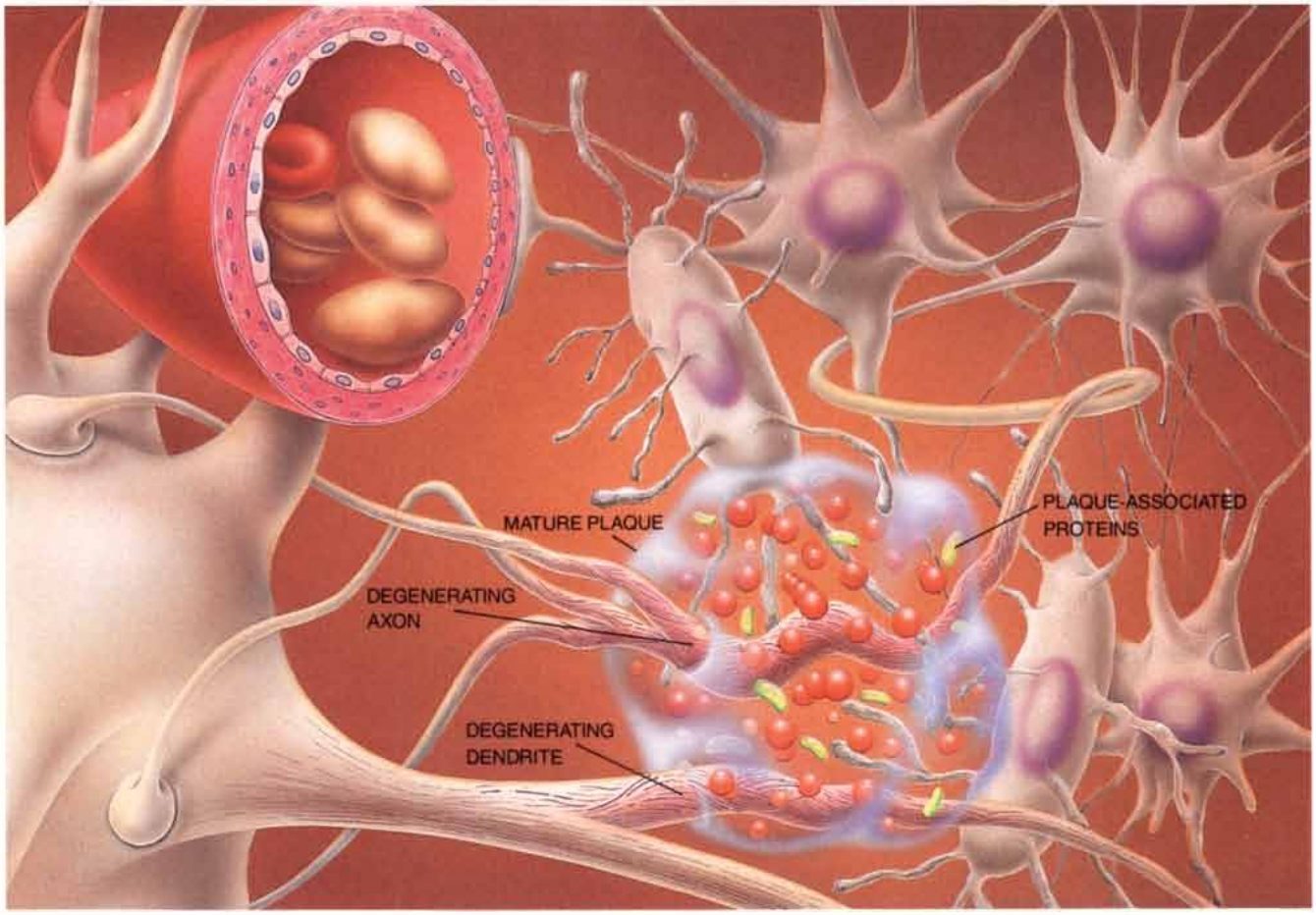
The study of the amyloid beta-protein has helped to clarify the genetic basis of Alzheimer's disease. It has long been known that at least some cases are caused by genetic abnormalities: in some families, roughly one half of each generation acquires this familial type of Alzheimer's disease. People with Down's syndrome, who are born with three copies of chromosome 21 instead of the normal two, almost always acquire the brain lesions typical of Alzheimer's disease prematurely, by their forties or fifties. The behavior and mental abilities of many Down's syndrome patients seem to decline further at about the same time.

Discoveries about the genetic mechanisms regulating the accumulation of amyloid beta-protein are helping us understand why past attempts to treat Alzheimer's disease have been so fruitless—and they are steering us toward new treatments that may be significantly more successful.

The story of the beta-protein in Alzheimer's disease begins with the neuropathologist. Even before the time of Alzheimer, pathologists knew that the human cerebral cortex sometimes contained variable numbers of spherical plaques. These plaques consisted of altered axons and dendrites (the long, tapering ends of neurons, collectively called neurites) surrounding an extracellular mass of thin filaments. Un-

ROLE OF AMYLOID BETA-PROTEIN in causing Alzheimer's disease is illustrated according to one hypothesis. Amyloid beta-protein, which may be released by neurons, by glial cells or by cells in the circulatory system, accumulates in the nervous system as innumerable diffuse plaques (*top*). The beta-protein, various glial cells and other proteins that become embedded in the plaque matrix as it matures may then gradually cause nearby neurons to degenerate (*bottom*).





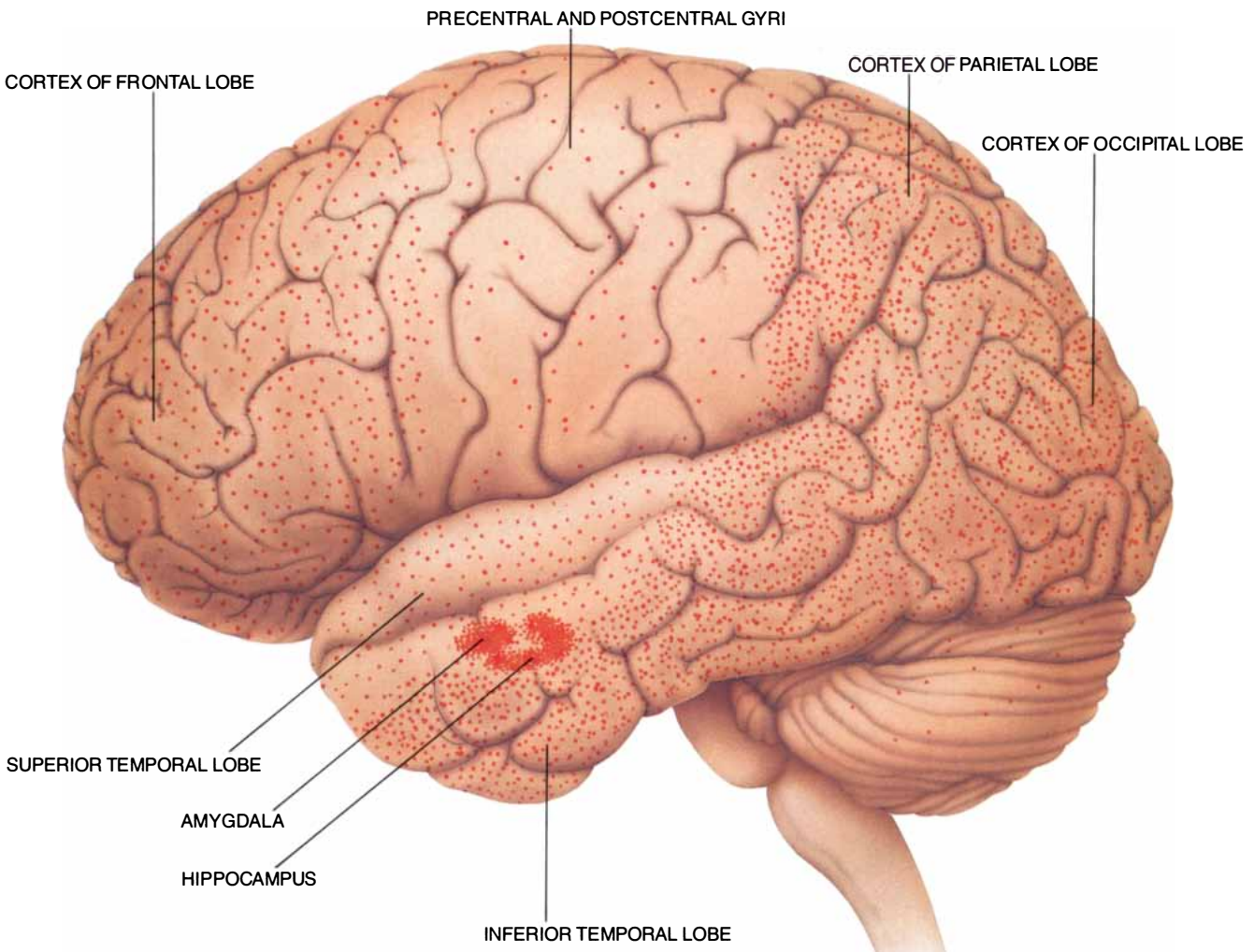
der the microscope, these filaments resembled similar extracellular deposits that accumulated in other organs in a variety of unrelated diseases.

In 1853 the great German pathologist Rudolf Virchow called such deposits "amyloid," an unfortunate term because it implied that the deposits were made of a starchlike substance. Chemical studies have shown that the principal constituents of the amyloid filaments are actually proteins and that the identity of the proteins differs among the various diseases marked by the deposition of amyloid. The common thread among these disparate diseases, or amyloidoses, is that they are characterized by innumerable extracellular deposits of normal or mutated protein fragments. Moreover, the protein subunits are always folded in a particular three-dimensional pattern called a beta-pleated sheet.

The senile plaque is a complex, slowly evolving structure, and the time required to generate fully formed, "mature" plaques may be years or even decades. In addition to the central core of amyloid beta-protein and the surrounding abnormal neurites, mature plaques contain two types of altered glial cells. (Glial cells normally associate with neurons and perform many supportive and protective functions.) In the center of the mature plaques, one usually observes microglial cells, which are scavengers capable of responding to inflammation or the destruction of nervous system tissue in many brain disorders. Around the outside of the plaque are so-called reactive astrocytes, glial cells that are often found in injured brain areas.

Along with the senile plaques, brain tissue affected by Alzheimer's disease is characterized by variable numbers of neurofibrillary tangles: dense bundles of abnormal fibers in the cytoplasm of certain neurons. These fibers, or paired helical filaments, are not made of the amyloid beta-protein. Instead they appear to be composed of a modified form of a normally occurring neuronal protein called tau. Like the amyloid plaques, the neurofibrillary tangles are not specific for Alzheimer's disease. The tangles occur in a dozen or more chronic diseases of the human brain.

Most of us who live into our late seventies will develop at least a few senile plaques and neurofibrillary tangles, particularly in the hippocampus and other brain regions important for memory. For the most part, the distinction between normal brain aging and Alzheimer's disease is quantitative rather than qualitative. Usually patients with progressive dementia of the Alzheimer type have moderately or markedly more mature neuritic plaques and neurofibrillary tangles than age-matched nondemented people do. Elucidation of the genesis of the plaques and tangles in Alzheimer's disease should therefore tell us a lot about the highly similar le-



BRAIN OF AN ALZHEIMER PATIENT characteristically shows numerous plaques of amyloid beta-protein (indicated by stippling) in specific brain areas. These plaques become centers for the degeneration of neurons.

sions that underlie in part the more subtle changes in memory and cognition affecting some otherwise healthy septuagenarians.

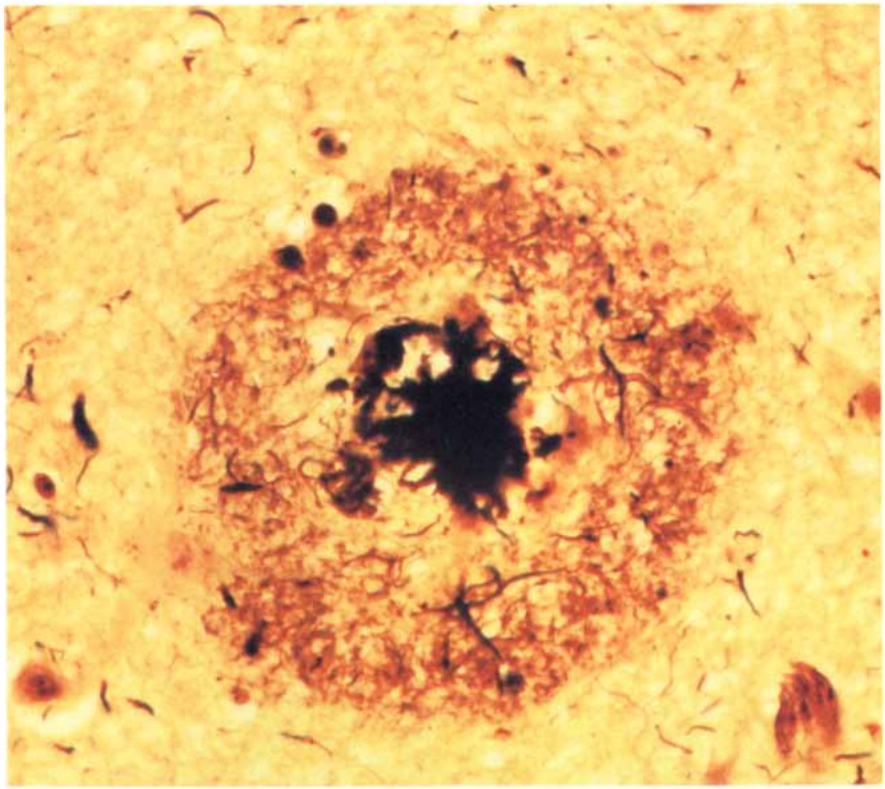
Because large numbers of amyloid-bearing neuritic plaques in brain regions critical for intellectual function are an invariant feature of Alzheimer's disease, we need to understand the nature of the amyloid protein. In 1984 George G. Glenner and Caine W. Wong of the University of California at San Diego first isolated amyloid from blood vessels in the meninges (connective tissue surrounding the brain) of Alzheimer patients. When they solubilized the isolated amyloid, they found it was composed of a small protein, dubbed the "beta protein," that had a novel amino acid sequence.

Shortly thereafter, Colin L. Masters of the University of Western Australia, Konrad Beyreuther of the University of Cologne and their colleagues isolated the amyloid cores of senile plaques. The core protein had the same size and amino acid composition as the meningovascular beta-protein, and antibodies against one reacted with the other.

But plaque cores isolated by Carmela Abraham and me at that time yielded beta-protein that had a "blocked" first amino acid and thus could not be sequenced, in contrast to the vessel-derived beta-protein that had been successfully sequenced by Glenner and Wong. This and other findings suggested to us that the mature plaque core contained beta-protein that was chemically modified as compared with that found in blood vessels. It appears that both vascular and plaque amyloid are composed of the beta-protein but that its precise chemical form in the two sites may differ slightly.

After a protein is purified and sequenced, the next step in its characterization is often the cloning of the complementary DNA molecule that embodies the genetic code for synthesizing the protein. In the case of amyloid beta-protein, this was accomplished independently by four laboratories in early 1987. One of these groups—Jin Kang, Beyreuther, Benno Müller-Hill and their colleagues at the University of Cologne—had the good fortune of isolating a stretch of DNA that contained the entire coding sequence for the protein. This sequence demonstrated that amyloid beta-protein was just a small fragment, 40 or so amino acids, out of a 695–amino acid protein, which is now generally referred to as the beta-amyloid precursor protein (beta-APP).

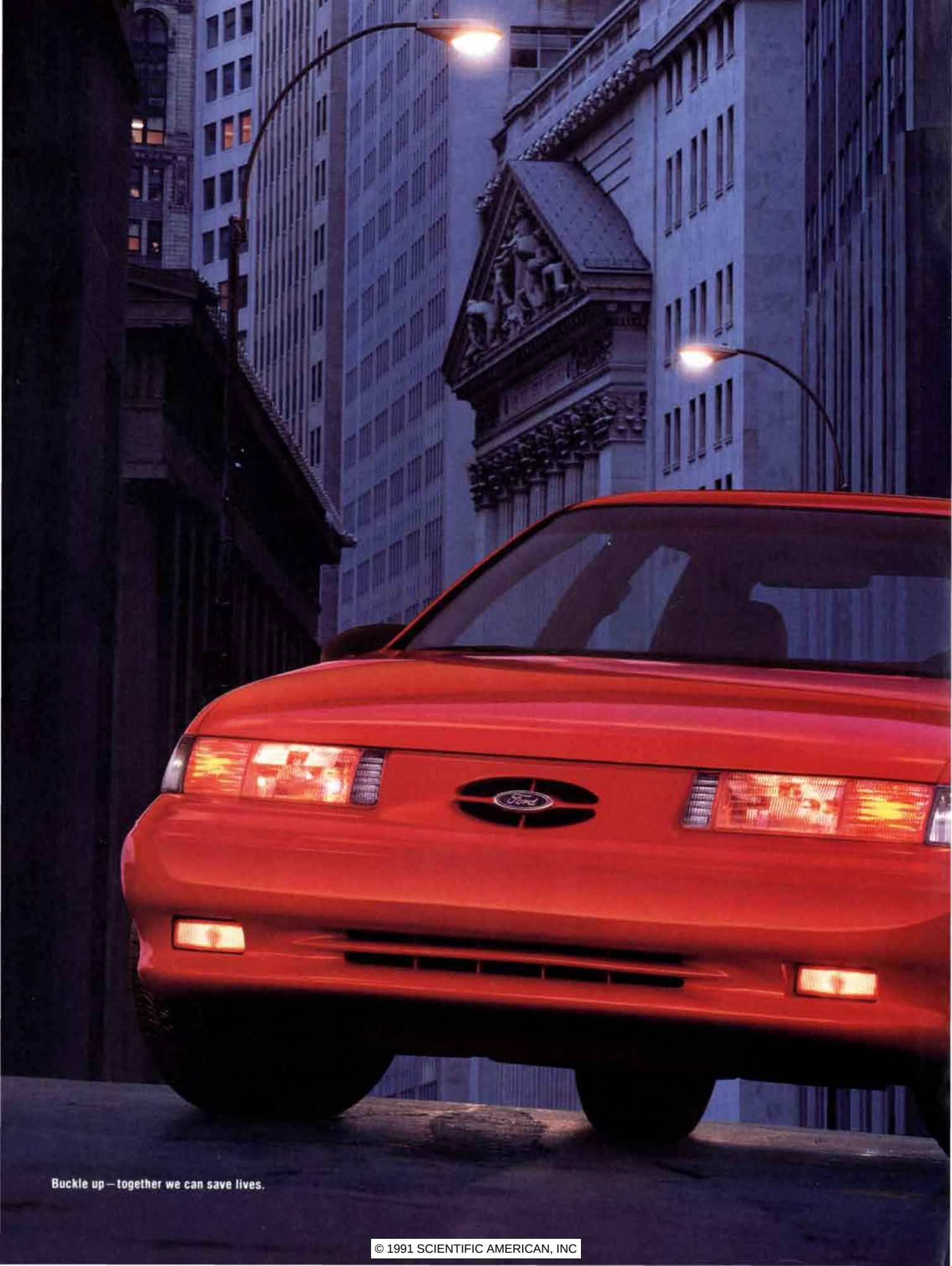
The presumed structure of beta-APP included one region (from amino acid 625 to 648) that could anchor the molecule to cell membranes. To everyone's surprise, the small amyloid beta-protein fragment comprised amino acids 597 through 636—the 28 amino acids just outside the membrane-spanning domain and the first 12 amino acids within the membrane. That finding presented a conundrum that has still not been resolved: How can a segment of beta-APP that normally anchors it firmly to cell membranes appear in the extracellular space as amyloid? Or to put the question another way, how can the enzymes that snip amyloid beta-protein out of its large precursor gain access to that transmembrane region? The answer may have important implications for blocking the amyloidosis of Alzheimer's disease.

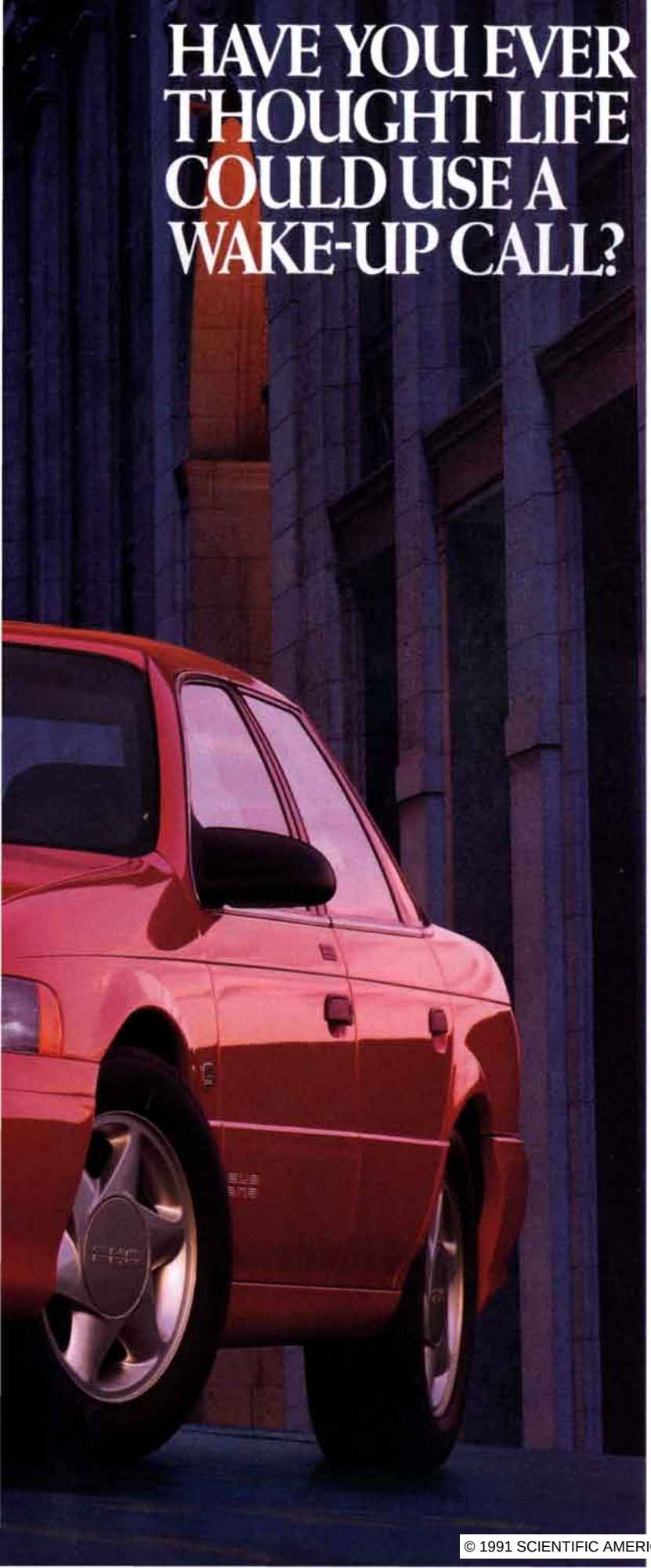


PLAQUE OF AMYLOID BETA-PROTEIN in the brain of an Alzheimer patient is visible as a black globular mass in this stained image. The plaque is surrounded by a halo of abnormal neurites (axons and dendrites) and degenerating neural cell bodies that appear darker than the normal neurons.

Perhaps even more exciting than knowing the amino acid structure of the amyloid precursor was the finding, made by each of the four laboratories that had cloned beta-APP, that the precursor was encoded by a gene located on chromosome 21. This discovery gave rise to a kind of global "aha!" experience in the Alzheimer field. We could suddenly fathom why people with Down's syndrome, who are born with an extra copy of that chromosome, routinely developed beta-amyloid deposits at a relatively early age. At about the same time, a collaborative group led by Peter St. George-Hyslop, Rudolph Tanzi and James F. Gusella of the Massachusetts General Hospital found that at least one form of familial Alzheimer's disease (FAD) appeared to be caused by a genetic defect that was also located somewhere on chromosome 21.

At first it seemed that the gene for beta-APP might itself be the one responsible for early-onset FAD. Yet as so often happens in science, the situation is more complex. Several studies of FAD patients in late 1987 and early 1988 failed to reveal either defects or duplications in the protein-coding region of the beta-APP gene; also, a long stretch of DNA appeared to separate the beta-APP gene from the approximate site of the early-onset FAD defect on chromosome 21. Moreover, Gerard D. Schellenberg and his collaborators at the University of Washington and Margaret A. Pericak-Vance and her colleagues at Duke University Medical School were unable to demonstrate linkage of Alzheimer's disease to markers on chromosome 21 in several different families, including some with late-onset (older than 65 years) forms of the disease.





HAVE YOU DRIVEN A FORD LATELY?

THE 24-VALVE, 220 HORSEPOWER FORD TAURUS SHO.
It's a wake-up call in a world of sleepy sedans. The 24-valve, 220 horsepower shot of adrenalin called Ford Taurus SHO.

When its Super High Output V-6 comes alive, a special sport-tuned suspension and anti-lock brakes sync up with a *new, improved five-speed shifter* to create a responsive and invigorating performance.

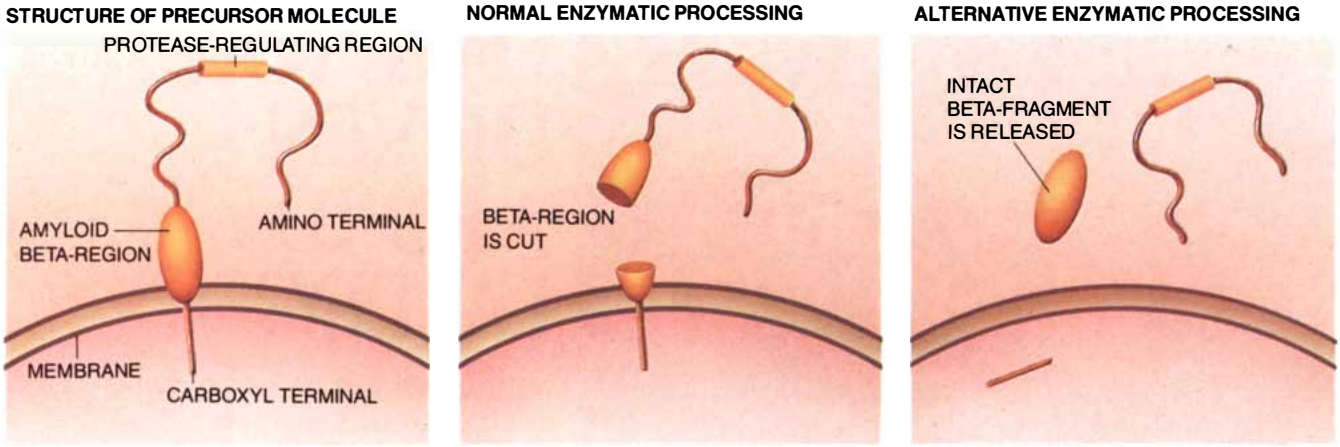
REFINED TO BE EVEN MORE REFINED.
All it takes is some rapid eye movement around the cockpit to see that it's all new. Its dash flows into the doors in a seemingly seamless manner. SHO has performance gauges, articulated sport seats (leather trim is an option), a driver-side air bag (a supplemental restraint system to be used with your safety belt), and a newly available front passenger-side air bag. Outside, *ground effects enhance a new shape* to create a dramatic look. One to be noticed both coming and going.

Take your daily routine down a new route. The 1992 Ford Taurus SHO.

NEW BUMPER-TO-BUMPER WARRANTY.
New 36-month/36,000-mile limited warranty requires *no deductible.* See your dealer for a copy.

NEW TAURUS
SHO

Ford



BETA–AMYLOID PRECURSOR PROTEIN (beta-APP) is a membrane-spanning molecule that can be processed in more than one way (*left*). Normally, enzymes cut beta-APP in the middle of the beta-region (*center*). Sometimes enzymes liberate the intact beta-fragment by cutting only at its ends (*right*). Those fragments can then be secreted and accumulate in tissues. Some forms of beta-APP contain a protease regulator, which may be involved in the protein's normal function.

These and other studies have led to the conclusion that FAD is genetically heterogeneous—it can apparently be caused by many different genetic defects on different chromosomes.

This realization should perhaps come as no surprise. Alzheimer's disease is very common, occurs in many different ethnic groups and is closely similar to the normal process of brain aging. It is reasonable that the syndrome we call Alzheimer's disease could arise in somewhat different forms from many distinct genetic alterations. All these alterations, however, appear to act through a critical common mechanism involving the increased deposition of amyloid beta-protein.

A new and potentially exciting development in the search for the genes that cause FAD was reported recently by Alison Goate, John Hardy and their collaborators at St. Mary's Hospital in London. They found two families in which everyone who had Alzheimer's disease also had a particular DNA pattern in the beta-APP gene. When the investigators sequenced the DNA encoding the amyloid beta-protein region of beta-APP, they found a mutation that resulted in a switch of amino acid 642 (out of 695) from a valine to an isoleucine. Because no mutation in the beta-APP gene had previously been found in either normal or FAD subjects, that change is probably not an irrelevant chance event but rather the cause of Alzheimer's disease in these two families. In recent months at least six other families prone to FAD that have a DNA mutation at position 642 have been found.

Two major conclusions emerge. First, one specific molecular cause of Alzheimer's disease has been identified. Second, it is now clear that beta-amyloid deposition can arise directly from a mutation in beta-APP, without any other preexisting cellular or molecular defect. An idea that some of us have long held—that beta-amyloid abnormalities can truly initiate some forms of Alzheimer's disease—has moved from speculation to reality.

Some cases of Alzheimer's disease appear to occur sporadically, that is, without any known familial predisposition. It is difficult, however, to reach a firm conclusion in this regard about a late-onset illness, because family members bearing a faulty gene may have died of other causes before the symptoms appeared. Environmental factors are likely to influence the onset of Alzheimer's disease. One piece of evidence supporting this opinion is the observation that identical twins may manifest Alzheimer symptoms at considerably different ages. Unfortunately, the search for environmental factors that can trigger the disease has been resoundingly unsuccessful to date, although debate still swirls around the unsettled role of aluminum as a contributing factor. One possible environmental trigger noted in a small minority of patients is a history of major head trauma, although how trauma could accelerate beta-amyloid deposition is unclear.

While investigations of the genetic defects underlying familial Alzheimer's disease have gone forward, steady progress in characterizing beta-APP itself and the role of the beta-protein in the disease has also continued. On the molecular biological front, several laboratories discovered alternative DNA sequences encoding beta-APP—sequences that contain either one or two additional coding segments at position 289 of the original 695-amino acid form. One of these two inserts encodes a stretch of amino acids that has the ability to bind to and inhibit proteases, the enzymes that cut proteins into smaller fragments. Because of that discovery, we can guess for the first time about one normal function of beta-APP: it may be an inhibitory molecule that regulates the activity of proteases.

On the protein chemistry front, Marcia Podlisny and I at Harvard, in collaboration with Tilman Oltersdorf and Lawrence C. Fritz of Athena Neurosciences in South San Francisco, identified and characterized beta-APP in brain and other human tissues and in cultured cells. In all the tissues that we examined, we detected a stable fragment of beta-APP that contained one end (the carboxyl terminus) of the molecule as well as part or all of the critical amyloid beta-protein region. Several laboratories subsequently showed that the other part of beta-APP, the one with the amino terminus, is shed into the fluid outside of cells, including cerebrospinal fluid and plasma. In 1990 Fred Esch and his co-workers at Athena Neurosciences showed that this normally occurring fragmentation of beta-APP resulted from a cleavage of the full-length precursor at amino acid 16 within the amyloid beta-protein region.

Their finding tells us that the as yet unidentified protease that normally cuts at this site prevents the formation of intact amyloid beta-protein. The beta-amyloid deposition that occurs during aging and in Alzheimer's disease must therefore involve an alternative proteolytic pathway—one that cleaves beta-APP at the beginning and end of the

amyloid beta-protein region. Researchers are now avidly seeking those alternative proteases, because drugs designed to inhibit them should decrease or prevent the deposition of amyloid beta-protein.

I have placed great emphasis on understanding the normal and abnormal processing of beta-APP. Yet what is the evidence that amyloid beta-protein deposition precedes the pathology of Alzheimer's disease rather than follows it? Since the beginning of the century, neuropathologists have argued about whether the amyloid in the core of the senile plaque was produced by the neurites in the plaque periphery as they degenerated or whether its appearance preceded and caused neuritic changes. Some researchers wondered whether the amyloid might have come from healthy neurons or from glial cells or even from nearby blood vessels.

In 1988 and 1989 several investigators noticed that amorphous, nonfilamentous deposits of amyloid beta-protein occur in Alzheimer brain tissue and that such diffuse, or "pre-amyloid," plaques are actually much more abundant than the classic neuritic plaques. Using antibodies against amyloid beta-protein as highly sensitive probes, researchers have detected such diffuse plaques not only in the brain areas implicated in the symptoms of the disease, such as the cerebral cortex, but also in other areas, such as the thalamus and cerebellum. Significantly, most diffuse beta-protein deposits contain few or no degenerating neurites or reactive glial cells. Electron microscopic examinations by Haruyasu Yamaguchi of Gunma University in Japan have revealed that much of the tissue within the diffuse plaque is indistinguishable from surrounding normal brain tissue.

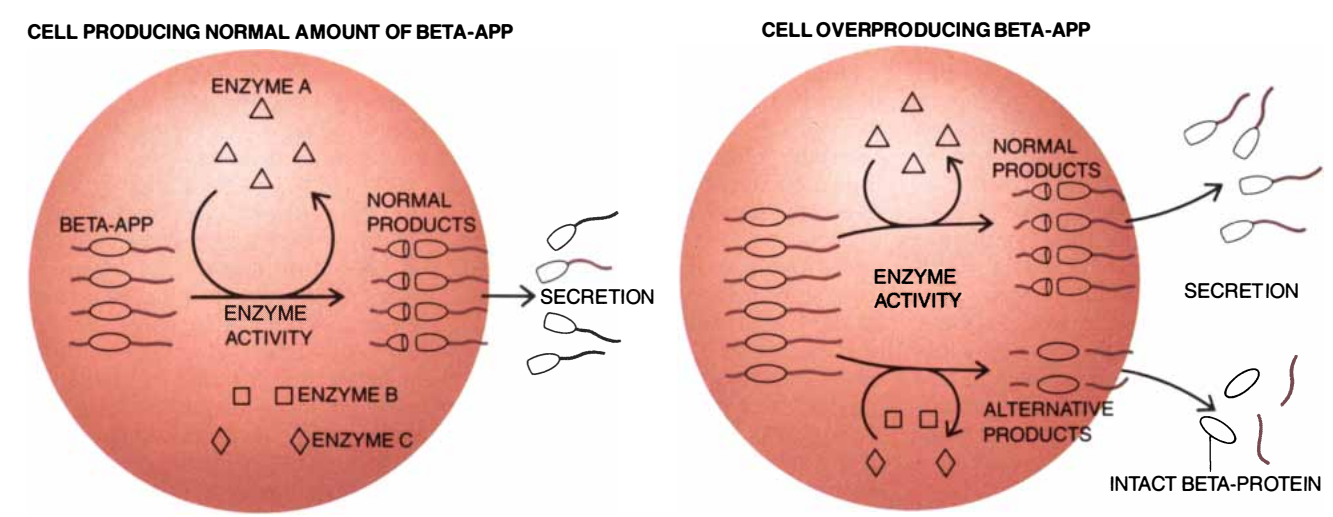
Several laboratories have now also reported that some patients with Down's syndrome who die in their teens or twenties seem to have many diffuse plaques in the absence of mature neuritic plaques, neurofibrillary tangles or other signs of cellular pathology. Because virtually all such patients would ultimately have exhibited these full-blown lesions, one may conclude that diffuse amyloid beta-protein deposits precede Alzheimer-type neuropathology in Down's syndrome and, by implication, in Alzheimer's disease itself. It is likely that only a minority of the diffuse beta-protein deposits gradually progresses to involve the surrounding neurites and glia. For unknown reasons, this maturation seems to occur much more commonly in the symptom-producing cerebral cortex than in, for example, the symptom-free cerebellum.

The observation of diffuse plaques suggested that amyloid beta-protein deposition preceded the alteration of neurons and other brain cells. Such data, as well as the occurrence of amyloid beta-protein in the walls of meningeal blood vessels outside brain tissue, led my Harvard Medical School colleagues Catharine L. Joachim and Hiroshi Mori and me to search for amyloid beta-protein deposition in organs other than the brain. In early 1989 we found small deposits in and around selected blood vessels of the skin, intestine and certain other tissues from some Alzheimer patients and aged control subjects; these deposits reacted specifically with antibodies against amyloid beta-protein.

The detection of those deposits provided the first evidence that the process underlying beta-amyloid deposition in Alzheimer's disease may not be restricted to the brain. The predilection of these extracerebral deposits to occur near blood vessels also strengthens the parallels between Alzheimer's disease and certain systemic amyloidoses that we know have a circulatory origin. Most important, the deposition of small amounts of amyloid beta-protein in peripheral blood vessels in the absence of any preceding neuronal injury—indeed in the absence of local neurons and glial cells—supports the hypothesis that the release and accumulation of amyloid beta-protein precede rather than follow neuronal degeneration in the brain.

Some of the clearest evidence that amyloid beta-protein deposition can be the primary event initiating an illness has emerged from studies of a rare genetic disorder found in two villages in the Netherlands. Patients in the affected families die in midlife from cerebral hemorrhages caused by severe amyloid deposition in innumerable blood vessels, hence the name "hereditary cerebral hemorrhage with amyloidosis of the Dutch type" (HCHWA-Dutch). Blas Frangione and his collaborators at New York University Medical Center and the University of Leiden found that the deposited protein in this disease was in fact amyloid beta-protein.

Then, in 1990, Efrat Levy and Frangione of New York University, Mark Carman of Athena Neurosciences, Sjoerd van Duinen of the University of Leiden



OVERPRODUCTION OF BETA-APP could encourage the liberation of the amyloid beta-fragment. Ordinarily, only the enzymes that cut within the amyloid beta-region may react with beta-APP (*left*). When excess beta-APP appears, however, other enzymes that release the amyloid beta-fragment by cutting at its ends may also have the opportunity to act (*right*). This mechanism may occur in people with Down's syndrome, who carry an extra copy of the beta-APP gene.

YELLOW DEPOSITS IN STAINED BLOOD VESSELS contain amyloid beta-protein. The discovery of this protein in the vascular system suggests that the amyloid protein accumulating in the brain may originate in the bloodstream.

and their co-workers discovered a DNA mutation in the beta-APP gene that caused the substitution of the amino acid glutamine for a glutamic acid at position 22 within the amyloid beta-protein. Exactly why this unique mutation leads to such severe cerebrovascular deposition is not clear. Nevertheless, genetic analyses of affected and unaffected siblings by Christine Van Broeckhoven of the University of Antwerp and her collaborators confirm that the mutation is the disease-causing defect.

The Dutch patients also seem to show numerous beta-protein deposits in the cerebral cortex that closely resemble the pre-amyloid plaques of Alzheimer's disease and Down's syndrome. For unknown reasons, those deposits do not seem to affect the surrounding neurites and glia: no neurofibrillary tangles or other neuronal alterations appear, and the patients do not suffer Alzheimer-type dementia. The characterization of HCHWA-Dutch nonetheless demonstrated for the first time that a mutation in beta-APP can cause diffuse amyloid beta-protein plaques and cerebrovascular beta-amyloidosis. The work provides a compelling example of a principle of medical research: the study of a rare disease in a small patient group can provide critical insights into common pathological processes that affect the entire population.

Even if one assumes that these various observations support the early deposition of amyloid beta-protein in Alzheimer's disease, what real evidence is there that the beta-protein is biologically active? Researchers must determine whether amyloid beta-protein is itself injurious to neurons or whether it serves as a matrix to which other more active molecules bind. At the moment, only fragmentary information about the biological effects of amyloid beta-protein is available. In 1988 Janet S. Whitson and Carl W. Cotman of the University of California at Irvine, in collaboration with me, noticed that the first 28 amino acids of the amyloid beta-protein had trophic (survival-promoting) effects on cultures of rat hippocampal neurons. Subsequently, Cotman and his colleagues demonstrated that the entire beta-protein molecule had both neurite-promoting effects and indirect neurotoxic effects.

More recently, Bruce A. Yankner and his colleagues at the Children's Hospital in Boston have reported that, under the right conditions, low doses of the full-length amyloid beta-protein can enhance the survival of freshly cultured rat neurons. When those investigators raised the dose modestly and allowed the cultures to age for four days, however, they observed toxic effects on some of the neurons. They localized these biological effects to one segment of the amyloid beta-protein (amino acids 25 through 35). The amino acid sequence of this molecular region is very similar to that of a naturally occurring brain peptide called substance P. If the neurotoxic effects observed in Yankner's study are confirmed, the argument that amyloid beta-protein is directly responsible for the neuronal pathology found in neuritic plaques will be substantially strengthened.

A major obstacle to deciphering how the beta-protein is released from its large precursor and causes neurons and glial cells to change is the lack of a close, convenient animal model for Alzheimer's disease. Inexpensive laboratory rodents do not spontaneously develop cerebral amyloid deposits with age, although expensive primates do. In 1987 my colleagues and I, in collaboration with Linda Cork and Donald Price of the Johns Hopkins School of Medicine, showed that antibodies made against purified human beta-protein reacted strongly with the amyloid found in the senile plaques and blood vessels of old monkeys, dogs and several other mammals. Additional studies in Price's laboratory and ours have yielded evidence that monkeys undergo an age-related beta-amyloidosis strikingly similar to that of humans. When drugs that block either the formation or the activity of the beta-protein are developed, it will be important to test their efficacy and safety in primates.

A more practical animal model is now being developed by injecting fragments of human DNA into fertilized mouse ova. Several laboratories have recently reported the production and breeding of these transgenic mice, which incorporate part of the human beta-APP gene into their genomes. In one example, transgenic mice developed by Barbara Cordell and her colleagues at California Biotechnology exhibit deposits of beta-protein that show some of the features of diffuse plaques. Whether such deposits will lead to alterations of the surrounding neurons and glial cells that are analogous to those of Alzheimer's disease remains to be seen. Various strains of transgenic mice expressing the mutant beta-APP molecules will be particularly important for studying the formation of beta-protein deposits and their inhibition by drugs.

How can we integrate the rapidly emerging knowledge about the nature and effects of amyloid beta-protein and its precursor into a dynamic model of how at least some forms of Alzheimer's disease work? I speculate that a group of distinct but related defects on chromosome 21, including DNA mutations, deletions and perhaps rearrangements, can result in either structurally abnormal beta-APP proteins or deregulation of the transcription of the beta-APP gene.

In families with the latter type of defect, the DNA that is altered presumably controls how much or what form of

TRADITION ~ TECHNOLOGY

JOIN SCIENTIFIC AMERICAN IN THE FAR EAST AND DISCOVER THE DIFFERENCE BETWEEN TRADITION AND TECHNOLOGY IN JAPAN, SOUTH KOREA AND CHINA.

Far East Kaleidoscope • May 6-23, 1992 • Aboard the new 80-passenger Aurora I

Explore the serene gardens, graceful temples, and soaring castles of Japan. Experience the excitement of modern Seoul, and examine the fascinating archaeological excavations at Kyong-ju, 7th-century Korea's royal city. Spend four days discovering the Imperial splendors of the Forbidden City, the Ming Tombs and the Great Wall. And witness the achievements of contemporary Japan, South Korea, and China.

SCIENTIFIC AMERICAN EXPEDITIONS

WHAT DO YOU GET FROM SCIENTIFIC AMERICAN EXPEDITIONS?

Provocative travel programs. Expert lecturers.
Engaging travel companions who share your interests.
And the chance to discover new worlds.

For a FREE brochure call Toll-Free
1-800-367-6766
or return this coupon to Scientific American Expeditions.

Send To: Scientific American Expeditions 132 East 70th Street New York, NY 10021

Name _____ Address _____

City _____ State _____ Zip _____

YES, please send me information about the **Far East Kaliedescope** as well as ☐ Costa Rica ☐ Antarctica.

messenger RNA is transcribed from the beta-APP gene. (Messenger RNAs are the crucial intermediary molecules in the translation of proteins from genes.) The control of gene transcription is a highly complex process, and the regulatory elements within the large beta-APP gene are not yet well understood. For example, a DNA alteration on chromosome 21 that causes some forms of FAD might interact adversely with the DNA in the regulatory regions of the beta-APP gene; alternatively, a mutation might occur directly within one of the regulatory regions. In the model I propose, the ultimate result of those DNA defects that occur outside the protein coding region of the beta-APP gene is to enhance—sometimes subtly, sometimes robustly—the amount or type of beta-APP proteins synthesized in at least some cell types. The result might be an overproduction of beta-APP somewhat reminiscent of that in Down's syndrome, which occurs because of an extra copy of the beta-APP gene.

The identity of the cell types producing the beta-APP molecules that release the amyloid beta-protein is unknown. In my view, circulating blood cells (such as platelets) and endothelial cells lining the blood vessel wall are among the most likely candidates, although neurons and glial cells are also reasonable possibilities. If these cells synthesized either excess or altered forms of beta-APP, some of those molecules might be broken down by an alternative enzymatic pathway, thereby liberating large fragments that contain the amyloid beta-protein.

Over time, I suspect, these fragments are further cleaved by proteases to release the intact amyloid beta-protein, which then accumulates in the extracellular spaces of the brain in the form of diffuse plaques. Because of local tissue factors in the cerebral cortex and other brain regions important for cognitive function, a minority of these diffuse plaques becomes increasingly filamentous and compact. The addition of so-called beta-amyloid-associated proteins—some of which have already been identified—and the activation of nearby microglia and astrocytes probably contribute to the maturation of the plaques.

At some point in this dynamic process, amyloid beta-protein or other molecules attracted to it begin to have both trophic and toxic effects on the surrounding axons and dendrites, resulting in a further evolution of some plaques to a neuritic form. In all probability, many other biochemical and structural changes accompany this phase, including a loss of synapses and a resultant decline in the cortical levels of acetylcholine and some other neurotransmitters. Some of the affected neurons would produce the masses of paired helical filaments that constitute the neurofibrillary tangles. As this complex and progressive cascade of molecular changes proceeds in areas such as the hippocampus and cerebral cortex, the patient slowly acquires symptoms of intellectual failure.

The mechanism of the disease I have outlined here is both speculative and simplified. Several observations cannot yet be placed into this scheme. The greatest challenge to students of Alzheimer's disease—and the greatest source of ongoing controversy—is the attempt to arrange the observed biochemical changes in a temporal sequence of pathogenesis. In proposing this model of Alzheimer's disease, particularly of the familial type, I have emphasized alterations in the regulation of the synthesis of beta-APP. The chromosome 21–linked form of FAD is strikingly similar to Down's syndrome, in which the elevated synthesis of beta-APP is a fact, not speculation.

Nevertheless, defects in the metabolism and degradation of beta-APP could also underlie some cases of Alzheimer's disease. For example, genetic defects could alter the enzymes that are normally responsible for attaching carbohydrate or phosphate molecules to beta-APP, or they could alter the proteases that normally break down beta-APP. In such cases, the gene making the defective enzyme could be located on any of the chromosomes, not just on chromosome 21. To date, no other chromosome has been clearly implicated as the site of a gene defect causing FAD, but Allen D. Roses and his colleagues at Duke University Medical Center have raised the possibility that in some families a gene on chromosome 19 may be involved.

In addition to searching for genetic alterations that may underlie different cases of FAD, it is important to pursue epidemiologic surveys that might uncover environmental factors that predispose individuals to Alzheimer's disease or that accelerate its course. At present, there is no compelling evidence that factors such as nutrition, educational level, occupation or emotional state influence the occurrence or progression of the disease.

The goal of all Alzheimer's disease research is the development of an effective therapy. In 1976 three laboratories in Great Britain published evidence that neurons synthesizing the neurotransmitter acetylcholine were severely compromised in the cerebral cortex of Alzheimer patients. Physicians subsquently attempted to increase the amount of acetylcholine by administering acetylcholine-enhancing drugs, largely to no avail. Today we know that Alzheimer's disease affects many different types of neurons and neurotransmitters, which is why the illness is difficult to treat by simply replacing the neurotransmitters.

The model for Alzheimer's disease that I have described points to new therapeutic strategies aimed at one or more crucial steps in its molecular progression. First, one might block the delivery to the cerebrum and its vasculature of those beta-APP molecules that are responsible for the amyloid deposits. That approach would be most feasible if, as I suspect, those proteins arrive by way of the bloodstream. Second, one could inhibit the proteases that liberate amyloid beta-protein by cleaving beta-APP. Third, one could retard the apparent maturation of amyloid beta-protein deposits into neuritic plaques, perhaps by interfering with the formation of the amyloid filaments that seems to accompany this change. Fourth, one could interfere with the activities of the microglia, astrocytes and other cells that contribute to the chronic inflammation around the neuritic plaques. And fifth, one might block the molecules on the surface of neurons that mediate the trophic and toxic effects of amyloid beta-protein and the proteins associated with it in the plaques.

None of these pharmacological targets will be easy to reach, but the number of investigators focusing on beta-amyloidosis and the current pace of discovery make it likely that inhibitors of one or more crucial steps in the development of the disease will emerge in the next few years. In view of the tragedy that Alzheimer's disease represents for its victims and for society, this feat cannot come a moment too soon.

FURTHER READING

ALZHEIMER'S DISEASE. R. Katzman in *New England Journal of Medicine,* Vol. 314, No. 15, pages 964–973; April 10, 1986.

THE PRECURSOR OF ALZHEIMER'S DISEASE AMYLOID A4 PROTEIN RESEMBLES A CELL-SURFACE RECEPTOR. Jin Kang et al. in *Nature,* Vol. 325, pages 733–736; February 19, 1987.

β-AMYLOID PRECURSOR PROTEIN OF ALZHEIMER'S DISEASE OCCURS AS 110- TO 135-KILODALTON MEMBRANE-ASSOCIATED PROTEINS IN NEURAL AND NONNEURAL TISSUES. Dennis J. Selkoe et al. in *Proceedings of the National Academy of Sciences,* Vol. 85, pages 7341–7345; October 1988.

SEGREGATION OF A MISSENSE MUTATION IN THE AMYLOID PRECURSOR PROTEIN GENE WITH FAMILIAL ALZHEIMER'S DISEASE. Alison Goate et al. in *Nature,* Vol. 349, pages 704–706; February 21, 1991.

THE MOLECULAR PATHOLOGY OF ALZHEIMER'S DISEASE. Dennis J. Selkoe in *Neuron,* Vol. 6, pages 487–498; April 1991.

Our pilots learn to fly like they learned to walk–one step at a time.

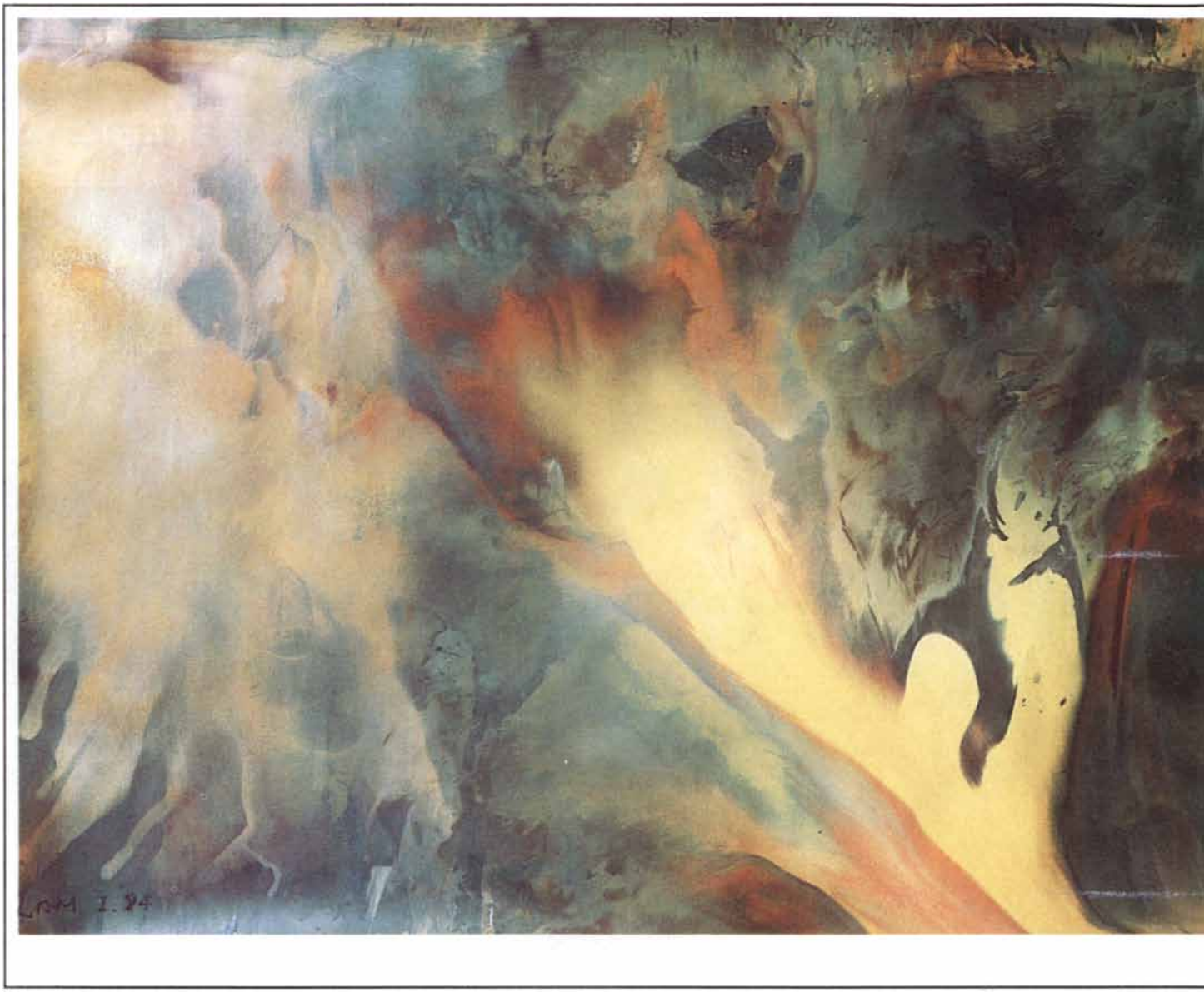
Phoenix, Arizona – Lufthansa Pilot Training Center

At Lufthansa, we believe a well-trained pilot is the most important safety device there is. So important, that we built our own flight schools in Bremen, Germany and Phoenix, Arizona. Just so our pilots are trained to the highest standards. But even before they can go through our rigorous program, they first have to get in. Each year we recruit 6,000 applicants to undergo tough psychological and intelligence testing. Only about 340 pass. And from there, it takes around 19 years of experience to finally become a captain on a Boeing 747. So you know your smooth flight on Lufthansa is no coincidence. It's the result of a passion, not to mention years and years of preparation.

A passion for perfection.SM

Lufthansa

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



Chromoskedasic Painting

It is a novel method for producing color images on black-and-white photographic paper without using pigments or dyes

by Dominic Man-Kit Lam and Bryant W. Rossiter

Passion—the chromoskedasic painting reproduced on these two pages—dazzles the eye with oranges, yellows, greens and blues. The hues are surprising because they were created by applying colorless chemicals to black-and-white photographic papers. Almost all paintings and color photographs consist of pigments that reflect certain parts of the visible spectrum of light and absorb others. But the original chromoskedasic painting contains no such pigments, only white paper and particles of silver.

These tiny particles produce colors not by reflecting or absorbing radiation but by scattering light. Particles of different sizes scatter different wavelengths of light, yielding various colors. The term "chromoskedasic" is derived from Greek roots meaning color by light scattering.

PASSION, 1984

The basic methods of chromoskedasic painting were discovered serendipitously one autumn evening in 1980. I (Lam) was developing black-and-white photographs of a retina, as part of a research project at Baylor College of Medicine. I noticed that some photographs were covered with patches of brown and yellow. Other photographers have undoubtedly observed this same effect. The colors usually appear because the photographic solutions were mixed improperly, because the solutions had deteriorated or because the photographic paper was defective. I wondered how colors could emerge from the use of black-and-white photographic paper and solutions, materials that do not contain dyes or pigments. Having seriously pursued painting for 20 years, I hoped to control the production of these colors and thus exploit the unusual characteristics of the new medium [see "Painting in Color without Pigments," THE AMATEUR SCIENTIST, page 136].

Without understanding in depth the mechanisms that generated the different colors, I systematically searched for the light and temporal conditions needed to produce different colors predictably on photographic paper. I also experimented with the concentrations of such solutions as activators, developers, stabilizers and fixers. By 1983 I had established a procedure for creating the primary colors—red, blue and yellow—and I could combine the three colors to create various shades. During the past eight years, I have refined the techniques by trying different kinds of papers and solu-

DOMINIC MAN-KIT LAM and BRYANT W. ROSSITER have investigated the techniques of chromoskedasic painting for the past three years. Lam is director of the Center for Biotechnology and professor of biotechnology, cell biology and ophthalmology at Baylor College of Medicine. His foremost interests are visual art, visual science and the development of a global network for biotechnology. Lam is a member of the President's Committee on the Arts and the Humanities. Rossiter is former director of the chemistry division and director of science and technology development at the Eastman Kodak Research Laboratories. He is also editor of the treatise *Physical Methods of Chemistry*.

SUNSET, 1985

ODE TO LINES AND CURVATURES, 1984

FELINE: FOR YEE YEE, 1990

SNOW MOUNTAIN, 1989

tions, and I have attained better control of colors, tones and composition.

Although my empirical approach led to a technique for producing color from black-and-white photographic materials, I had not given much thought to the physical and chemical mechanisms underlying such a process. In the summer of 1989 my co-author (Rossiter) came to visit me in Houston and took an interest in my paintings. He noticed the brilliant red and yellow hues, the occasional metallic sheen, the three-dimensional character of some objects and the unusual light stability. (Some paintings had been exposed to direct ambient sunlight for more than seven years and had not faded.) From these observations and others, he deduced that the colors in my paintings were a consequence of light scattering from tiny silver particles in the photographic paper. He coined the term "chromoskedasic."

The colors of most paintings rely on a process known as subtractive color. Blue pigments, for instance, reflect blue light while absorbing most other wavelengths. The colors in conventional photographs are also the result of a subtractive process. But instead of pigments, photographs require color-sensitive emulsions and developing agents that produce or release dyes [see "The First Color Photographs," by Grant B. Romer and Jeannette Delamoir; SCIENTIFIC AMERICAN, December 1989].

Most black-and-white photographic papers contain silver salts. Under the influence of light and chemicals, the salts decompose to form silver particles. Ordinarily, these particles merely absorb light, yielding tones of black. Those salts that are not exposed to light do not yield silver particles and

YELLOW MOUNTAIN, 1989

LOTUS POND, 1989

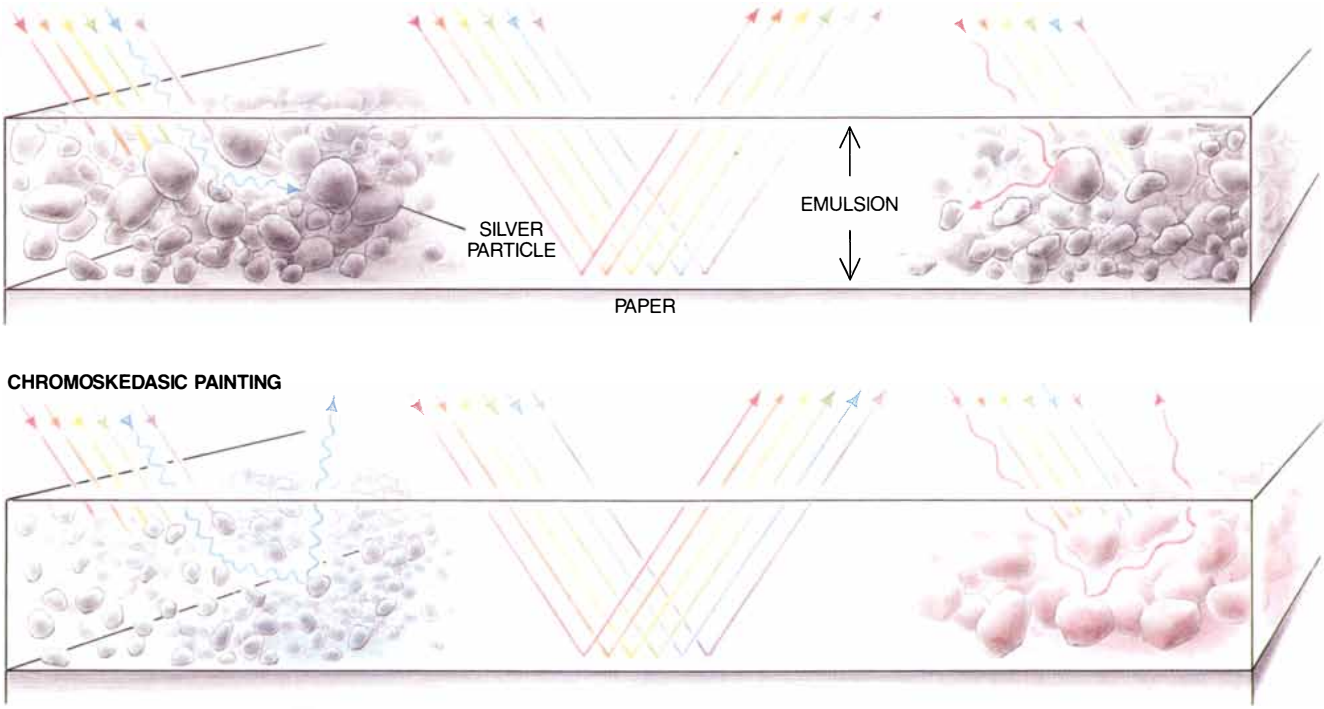
are washed away, allowing the white paper to show through. Chromoskedasic paintings are made using methods very similar to those employed in black-and-white photography. But chromoskedasic techniques require that the particle growth be carefully controlled through exposure to light and chemicals.

In chromoskedasic painting, the silver particles produce colors through a process known as Mie scattering. In 1908 Gustav Mie first described this scattering process in mathematical terms. He studied how light scatters through a medium consisting of spheres of similar size and electrical characteristics. He found that the wavelength of light scattered by such a medium depends on the size of the spheres [*see top illustration on opposite page*].

Mie's theory helped to explain why the sky is blue or why the sun appears red at dawn or sunset. The sky is blue, in part, because dust, water vapor and other particulate matter in the atmosphere are of such a size as to scatter light selectively from the blue region of the spectrum. Similarly, atmospheric particles of different sizes produce some of the brilliant colors of the rising or setting sun. (To be sure, many kinds of scattering, absorption and refraction processes contribute to the color of the sky.)

Particle formation in a complex, modern photographic emulsion is very different from the ideal conditions that Mie assumed. Chromoskedasic paintings consist of silver particles that vary greatly in shape and size. Nevertheless, the Mie equations predict the correlation between particle size and the colors created in chromoskedasic paintings.

Using electron microscopy, Donald L. Black of the Photoscience Research Division at Eastman Kodak in Rochester, N.Y., confirmed that regions of a particular color in a chromoskedasic painting contain silver particles that are relatively uniform in size. In agreement with theory, silver particles that produced yellow light were from 10 to 30 nanometers in diameter, whereas particles that scattered red light were 35 to 65 nanometers wide. Scientists still do not understand



CHROMOSKEDASIC PAINTINGS differ from black-and-white photographs in that the silver particles in the paintings are similar in diameter within a given region. A photograph (*shown in cross section at top*) is black in areas where particles absorb light. In a chromoskedasic painting (*bottom*), particles of a certain size scatter a particular color of light.

in detail how the size and shape of the particles influence the scattering of light and hence the colors of chromoskedasic paintings.

Many of the artistic possibilities of chromoskedasic painting have yet to be explored. Artistic expression is intimately related to materials and media. Watercolors are usually softer than oil paints, black-and-white photographs often bring out textures better than do color photographs, and marble sculpture is often more ponderous than steel. Chromoskedasic techniques bring out rich metallic colors that mingle, mix and wash over one another. The methods preserve the artist's expression for decades because the colors endure as silver particles firmly embedded in paper.

The techniques give artists many different ways to produce color images. They can manipulate a variety of darkroom conditions, such as light intensity, chemical concentrations and reaction time. They can precisely control the conditions to create representational paintings. Or they can allow nature to play with light and chemistry to generate abstract images. Chromoskedasic methods can also be easily integrated with conventional photographic techniques or with media such as acrylic, oil, watercolor and enamel paints. We hope chromoskedasic techniques will continue to inspire both artists and scientists.

VISION, 1986 (*enamel on chromoskedasic painting*)

FURTHER READING

THE THEORY OF THE PHOTOGRAPHIC PROCESS. Edited by T. H. James. Eastman Kodak Company, 1977.

INTRODUCTION TO PHOTOGRAPHIC THEORY: THE SILVER HALIDE PROCESS. B. H. Carroll and T. H. James. Wiley, 1980.

PHOTOGRAPHIC MATERIALS AND PROCESSES. Leslie Stroebel and Richard Zakia. Focal Press, 1987.

IMAGERY AND ILLUSION: THE PAINTINGS OF DOMINIC MAN-KIT LAM. Twinwood Press, 1990.

Microlasers

Millions of lasers measuring just a few millionths of a meter can now be etched on a single chip, offering a host of novel applications in optical communications and information processing

by Jack L. Jewell, James P. Harbison and Axel Scherer

Investigators have long dreamed of building an optical computer. Electronic switches and circuits would be replaced by a network of light through which individual impulses carry and process information. The idea is not only aesthetically appealing but also offers the promise of a machine that would be even faster and more versatile than the most powerful computer in use today.

The realization of the dream, however, has been thwarted by the lack of sufficiently tiny optical-signal processing devices. Like its electronic counterpart, the transistor, the basic building block of any optical processing or communications system must operate at low power and at high speed, and to do so it must be small. In the microelectronics industry, transistors having dimensions smaller than a micron (a millionth of a meter) are now routinely fabricated in numbers approaching tens of millions on a single semiconductor chip. Compared with this astonishing feat, the miniaturization of optical devices has lagged considerably.

The current state of affairs in commercially available optical technology is embodied by a device called the semiconductor diode laser, which is found in everything from compact-disc players to fiber-optic communications systems. Although the diode laser has revolutionized the storage and communication of information, the size of the device and its degree of integration are roughly comparable to that of the individually packaged transistors found in a late 1950s radio. A typical diode laser measures a few microns wide by several hundred microns long—making it several hundred times bigger than one of its microelectronic counterparts. Although such a device is orders of magnitude smaller than the familiar red helium-neon laser commonly used in supermarket bar-code scanners, it is simply too big to be useful in an optical computer.

Very recently, significant advances have been made in the miniaturization of diode lasers. In May 1989, in work that evolved from efforts to construct two-dimensional arrays of optical switches in hopes of eventually building optical computers, we fabricated more than one million microlasers, or micron-size lasers, on a single semiconductor chip about seven millimeters wide by eight millimeters long [*see illustration on opposite page*]. The microlasers were proposed by one of us (Jewell) and Sam McCall of AT&T Bell Laboratories. They were created at Bell Communications Research (Bellcore) by the other two of us and Leigh Florez. The devices range in size from one to five microns.

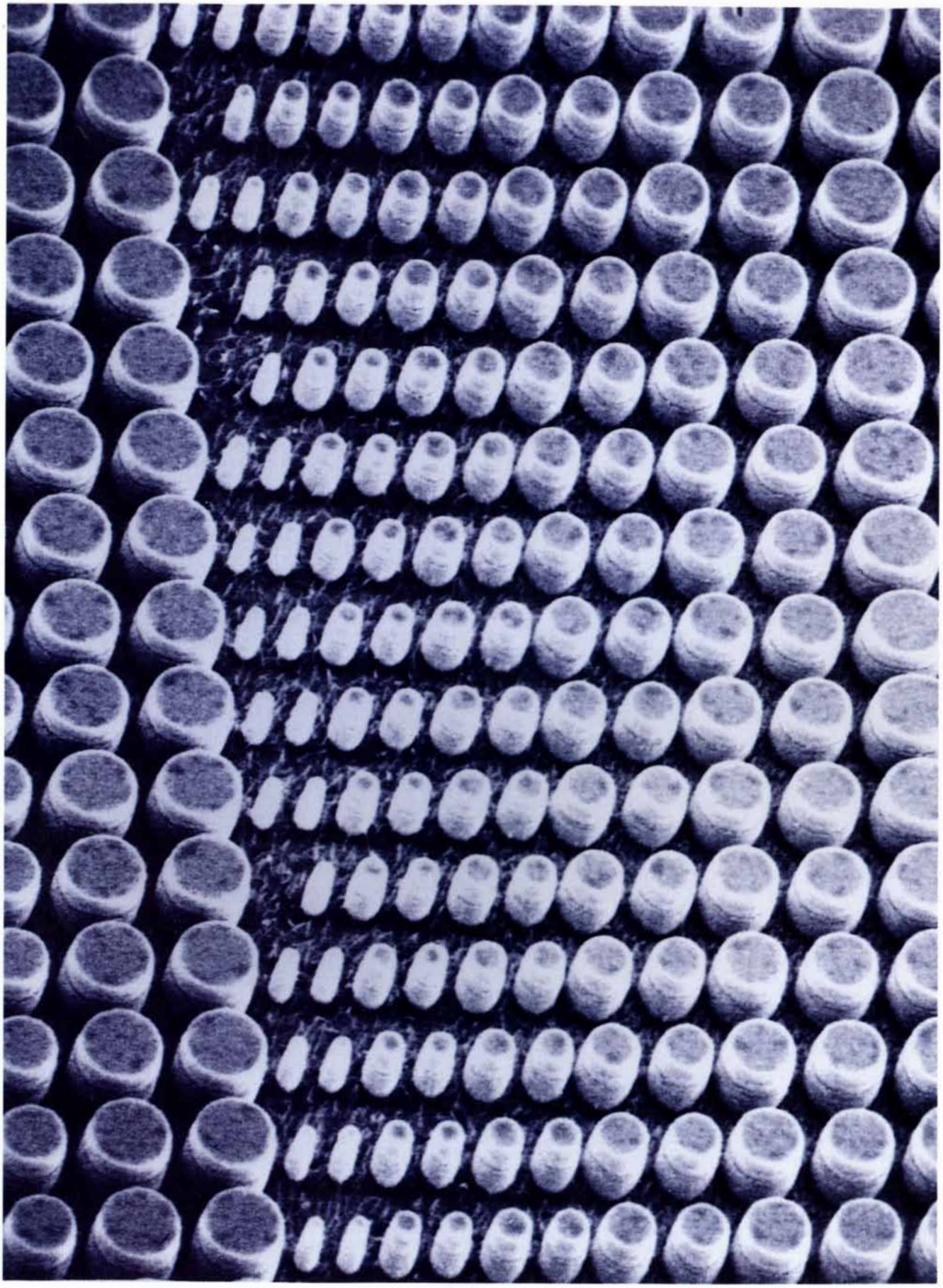
Such a scale is already two orders of magnitude smaller than that of conventional diode lasers. With additional work, the size can probably be brought down by another order of magnitude. Perhaps most exciting, as the microlaser approaches its practical lower limit in size (probably between one half and one quarter of a micron), it is believed that the quantum mechanical process of light emission can be radically altered in a way that will further enhance the performance of the device.

Microlasers are so new that at this time no one can predict whether they will make a significant impact on the commercial market. Moreover, the development of an optical computer is far from an easy task [see "The Optical Computer," by Eitan Abraham, Colin T. Seaton and S. Desmond Smith; SCIENTIFIC AMERICAN, February 1983]. Nevertheless, microlasers have generated enormous excitement for their potential applications in optical computers as well as in optical communications and information processing in general.

The principles of operation underlying a diode laser are the same as those for any laser [*see top of illustration on page 88*]. Atoms in a part of the laser called the amplifying medium—typically a solid, liquid or gas—are pumped, or energized, either electrically or with a source of electromagnetic radiation. When a light wave of a specific wavelength traveling through the amplifying medium encounters a pumped atom, it can induce the atom to release its energy in the form of a light wave at the same wavelength. The process is coherent, which is to say that the crests and troughs of the waves match up, and the intensity of the light increases. Mirrors on each end of the amplifying medium form a cavity, and they force the light to bounce back and forth many times through the medium, maximizing the increase in intensity.

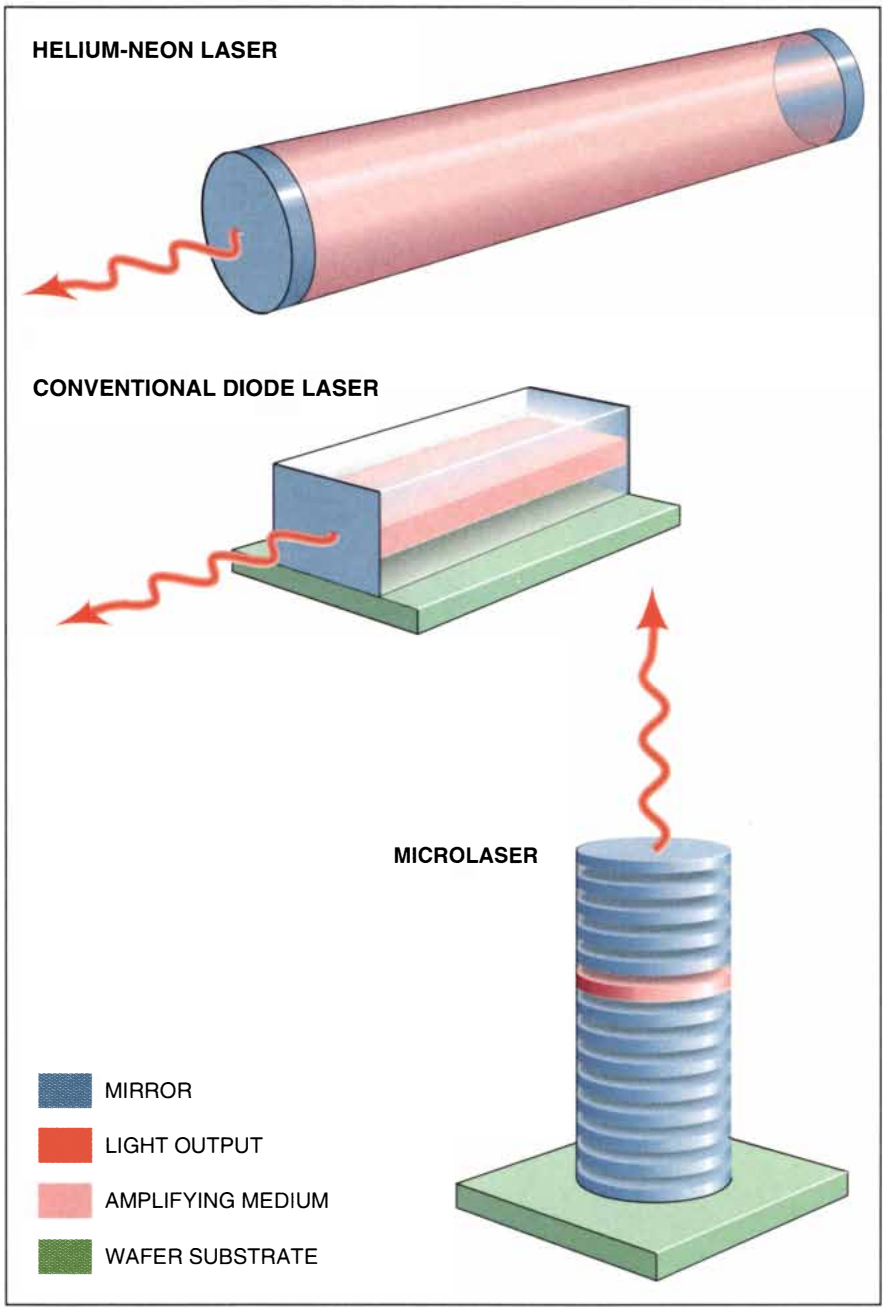
FIRST ARRAY OF MICROLASERS was completed in May 1989 by the authors. A tiny portion of the array, which contained more than one million microlasers, is shown in this scanning electron micrograph. The largest devices have a diameter of about five microns (millionths of a meter); the smallest are about a micron. Laser beams, which emerge from the circular-shaped ends, were generated in all the devices except the very smallest.

JACK L. JEWELL, JAMES P. HARBISON and AXEL SCHERER share an interest in optoelectronics. Jewell received a Ph.D. in optical sciences from the University of Arizona in 1984. Before co-founding Photonics Research, Inc., he was at AT&T Bell Laboratories. Harbison holds A.B., S.M. and Ph.D. degrees from Harvard University. Shortly after earning his doctorate in 1977, he moved to Bell Laboratories, joining Bellcore at its inception in 1984. Scherer is a member of the technical staff at Bellcore. He earned a Ph.D. from the New Mexico Institute of Mining and Technology in 1985.



The reflectivity of one or both of the mirrors is less than unity, with the result that some of the intensified light is able to escape from the cavity in the form of a laser beam.

In a diode laser the amplifying medium is a long, brick-shaped diode, a device that permits electric current to flow freely in one direction but blocks the flow of current in the other. The diode is made from a semiconducting material such as gallium arsenide. The amplifying medium of a conventional diode laser is about as long as the laser itself [*see middle of illustration on this page*]. Mirrors are formed on each end of the "brick" by simply cleaving the wafer: each interface between the semiconductor and air has a 30 percent reflectivity, which is more than sufficient for operation. When an electric current is made to flow parallel to the mirrors, the diode fills with electrons and holes (parts of the semiconductor lacking electrons). A short time later the electrons and holes recombine, and when they do, they emit light. The light is intensified by multiple reflections through the semiconductor, and a portion of it emerges from one of the mirrors in the form of a laser beam parallel to the underlying wafer substrate; that is, the beam comes out of the edge.



PRINCIPLES OF OPERATION are the same for a conventional helium-neon laser found in a supermarket bar-code scanner (*top*), a conventional semiconductor diode laser in a compact-disc player (*middle*) and one of the newly developed microlasers (*bottom*). Atoms in an amplifying medium (*pink*) are pumped, or energized, by electromagnetic radiation or current. When a light wave traveling through the medium encounters a pumped atom, it can induce the atom to release its energy in the form of a light wave at the same wavelength. The intensity of the light increases as it bounces back and forth between two mirrors (*blue*) positioned on each end of the medium. The reflectivity of one of the mirrors is less than unity, which allows some of the intensified light to escape in the form of a laser beam (*red*). The lasers are not drawn to scale: a helium-neon laser is about 100 to 1,000 times longer than a conventional diode laser, which in turn is about 100 times longer than a microlaser.

One characteristic that makes a microlaser so very different from a conventional diode laser is that the laser beam emerges perpendicular, instead of parallel, to the wafer substrate [*see bottom of illustration at left*]. Instead of the brick-shaped appearance of a conventional diode laser, a microlaser looks like a tiny Coke can; the laser beam shoots out of the top (or out of the bottom through the transparent wafer on which it has been grown). The difference in geometric construction means that more lasers per area can be packed on a wafer. Kenichi Iga and his colleagues at the Tokyo Institute of Technology first demonstrated the feasibility of such a surface-emitting (instead of edge-emitting) approach in the late 1970s.

Our own work in microlasers dates to about the time of Iga's breakthrough. Investigators at Bell Labs built surface-emitting lasers by sandwiching films of gallium arsenide that measured from one eighth of a micron to a few microns thick between mirrors and gluing the entire stack together. To sensible people, the notion of turning such a procedure into a practical technology was silly; fabrication bordered on black magic, and the devices were of poor quality. (Of course, the first transistors also had these "attributes.")

Steady improvements turned the tide. A microlaser, including mirrors, is now typically about six microns long, compared with the 250-micron length of a conventional diode laser. But what actually makes a microlaser "micro" is that its Coke-can geometry serves as a waveguide: an optical "hose" that prevents light from spreading outward as it propagates. (An optical fiber is a well-known example of a waveguide.) The typical diameter of a microlaser, as we mentioned earlier, is a few microns.

Even though microlasers are still in

Finding a Cure for Cancer is Driving Doctors up a Tree.

In the never-ending search
for a remedy for
ovarian cancer, doctors may
finally be barking up
the right tree.
The yew tree of the old-
growth forests in the
Pacific Northwest to be exact.

And on the next
Scientific American Frontiers
we'll show you how
doctors are trying to take a
rare drug called taxol
that's found in the yew bark and
produce it in laboratories.
We'll also look
at aerial robots, prairie
reclamation, aerodynamically
designed wheelchairs,
and slave ants.

So tune into
Scientific American Frontiers.
On PBS.* Sponsored by
GTE Corporation.
Who knows, you might even
find that there's
a cure for cancer growing right
in your own backyard.

*Check local listings for date and time.

SCIENTIFIC AMERICAN FRONTIERS

GTE

THE POWER IS ON

their infancy, their small size makes their power requirements comparable to those of the more mature conventional diode lasers. An important yardstick used to characterize the power requirements of any diode laser is the threshold current, the flow of electric current needed to make the diode just barely lase. The lowest threshold of a microlaser is 0.7 milliampere, compared with the 0.5-milliampere threshold of a conventional diode laser. With additional work, it is hoped that the threshold of a microlaser, because of its exceptionally small size, can be lowered to a few millionths of an ampere. At present, microlasers also fall somewhat short of conventional diode lasers in their bit rates—the number of zeros and ones that can be generated per second. At a bit rate of five billion per second, microlasers are not quite as fast as the best diode lasers, but rates approaching 100 billion per second are expected to be achieved eventually.

How are such exquisitely small, high-performance lasers made? Two crucial techniques make it possible. The first technique, molecular-beam epitaxy, allows the basic material of each laser to be built up from layers of semiconducting materials added one at a time. Individual lasers numbering in the thousands or millions can then be stamped out simultaneously by means of deep vertical etching, in conjunction with conventional photolithographic patterning.

Conceptually, molecular-beam epitaxy is quite straightforward. A semiconductor crystal is placed in a vacuum. Containers are filled with each of the chemical elements, such as indium, gallium, aluminum and arsenic, that will be used to build the desired semiconductor layers on top of the substrate. When an individual container is heated to the prop-

Fabricating and Testing the First Microlaser

Author Jack Jewell tells how it felt to do the groundbreaking work.

After months of strategizing and designing—and mostly anticipating—Axel Scherer picked up the laser wafer from Leigh Florez, who—working with Jim Harbison—had grown the structure by molecular-beam epitaxy. The wafer was two inches in diameter but only one fiftieth of an inch thick. We practically tiptoed with it back to Axel's lab. In spite of our care, however, within five minutes we managed to shatter the fragile disk into many odd-size pieces. Axel and I looked at each other and could think of nothing else to do but laugh. After all, we had planned to carefully cleave the wafer into smaller pieces anyway; our bad luck saved us time.

We picked up the larger pieces and started processing them. By midnight, we were ready to etch individual microlasers on the fragments. Then misfortune intervened in our favor once again. Axel told me that an important part of the etching apparatus, called a neutralizer, was not working and could not be fixed. Because the neutralizer sounds like some kind of gadget out of "Star Trek," we decided to pretend that it was not important and go on without it. By 2 A.M., Axel had etched several of our lasers, and we took the scanning electron micrograph shown on page 87. In the past, ugly devices that had looked like micro–trash cans had often worked quite well, whereas nice-looking devices had not worked at all. We agreed that the ones we had just made didn't stand a chance. The final step in the fabrication was to apply electrical contact points to the tops of the devices. For that, we would have to wait for the next day, since we knew of only one person who could do the job, Ray Martin (another Bellcore worker).

After a brief, fitful sleep, we asked Ray to spot-weld the contacts in place. At this point, anxiety really began to set in. We were rapidly approaching the transition from hopeful fabrication to the final test, which would surely spell failure and send us into months of troubleshooting. It didn't help that almost everyone who knew what we were trying to do thought we were more than a little daft. My wife (who was not one of those people) called to ask me if the lasers worked. I growled that we hadn't had time to test them yet and that they wouldn't work anyway, so don't bother me!

To test the lasers, we needed to touch the tops with a needlelike electrical probe. Trying to apply electric current through devices on my optical bench was a novel experience, because my entire career has been devoted to optics. Applying current to devices five microns tall and between one and five microns in diameter was definitely not for beginners. Axel and I took refuge in the fact that we were trying so many new things at once that no matter what happened we were sure to learn something. On touching the probe tip to one of the largest devices, we saw that current flowed in one direction, but not the other. Good! At least the device acted like a diode. We decided to set up a television camera to see if the desired infrared light was emitted. It was! The light was very dim up to about one or two milliamps of current and then increased dramatically, an encouraging sign. We also saw interference patterns, which are a characteristic of laser light.

"Our efforts would quickly degenerate into turning the translator knob and watching the blinking lights."

Something must be terribly wrong. The gadgets were behaving exactly like lasers. This could not be possible. We tried another device, then another and another. They all acted more or less the same, all over the chip. This could not be possible. It was too easy. Things don't work like that. Yet there we were.

For the next few weeks, it was hard to do any serious work. We found that with the aid of a device called a translator we could run the microlasers under the electrical probe at a rate of about a dozen a second. With the turn of a single knob, the lasers would blink on and off like a string of light bulbs. Over and over again, we went back and forth on the same devices, but they showed no signs of wearing out. Every time we decided to make some quantitative measurements, our efforts would quickly degenerate into turning the translator knob and watching the blinking lights, still in disbelief.

HE DOESN'T SELL VOWELS.
HE DOESN'T PACK A GUN.
HE DOESN'T SAY DUDE.
SO WHAT'S HE DOING ON TV?

To put it bluntly, Woodie Flowers isn't your typical TV star.

But, then again, Scientific American Frontiers isn't your typical TV show.

Sure, we try to solve mysteries. But they're not the type you see on your average police show.

They're some of the greatest mysteries of life. Everything from animal intelligence to curing cancer with tree bark.

And as a professor at MIT, nobody's more qualified to explore them with you than Woodie.

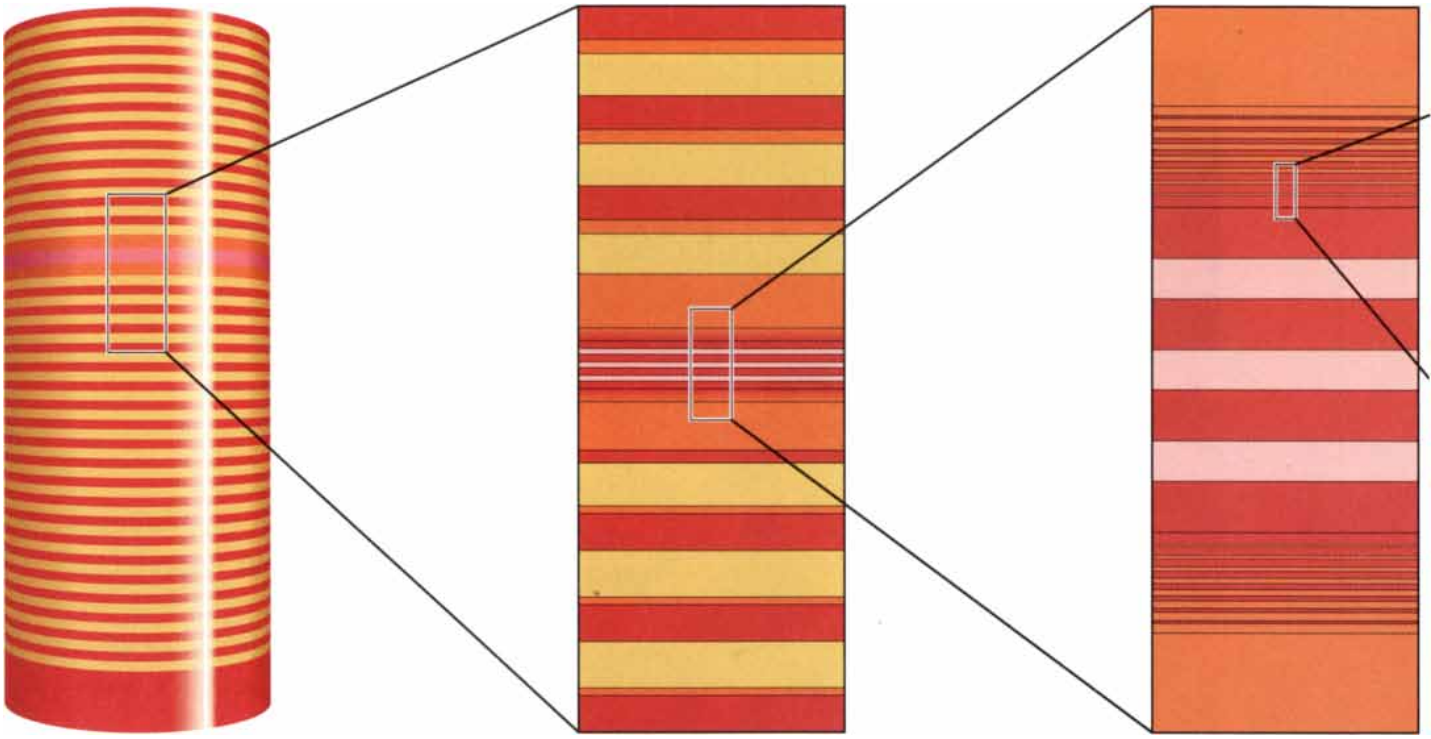
So tune into Scientific American Frontiers on PBS. Proudly sponsored by GTE Corporation. Hosted by Woodie Flowers.

He might not sell you any vowels, but with letters like PhD, BS and ME to his name, we think you'll be sold on him.

SCIENTIFIC AMERICAN FRONTIERS

GTE

THE POWER IS ON



SUCCESSIVE MAGNIFICATIONS of a microlaser reveal the large number of semiconducting layers—typically 500 or more—from which it is fabricated. In the most highly magnified drawing (*right*), individual atoms are visible: arsenic (*gray*), aluminum (*yellow*) and gallium (*red*). Layers made from them appear in the other three drawings: aluminum arsenide (*yellow*) and gallium arsenide (*red*), plus the alloy mixtures of aluminum gallium arsenide (*orange*) and indium

er temperature, the element within it begins to evaporate, streaming out the open end of the container and into the vacuum in much the way that steam escapes out the hole of a heated teapot. The result is a wide beam of molecules that sprays in the direction of the wafer. When desired, each beam of molecules can be blocked by a mechanical shutter.

The formation of multiple layers of different semiconductor materials is then accomplished by simply determining which chemical elements should be in which layer and opening the appropriate shutters. For example, layers of aluminum arsenide are formed by opening the aluminum and arsenic shutters simultaneously while keeping all the other shutters closed; similarly, layers of gallium arsenide are formed by opening the gallium and arsenic shutters. The thickness of each layer is determined by how long the shutters are held open. Molecular-beam epitaxy is such a refined technique that it is possible to grow an individual layer having the thickness of a single atom.

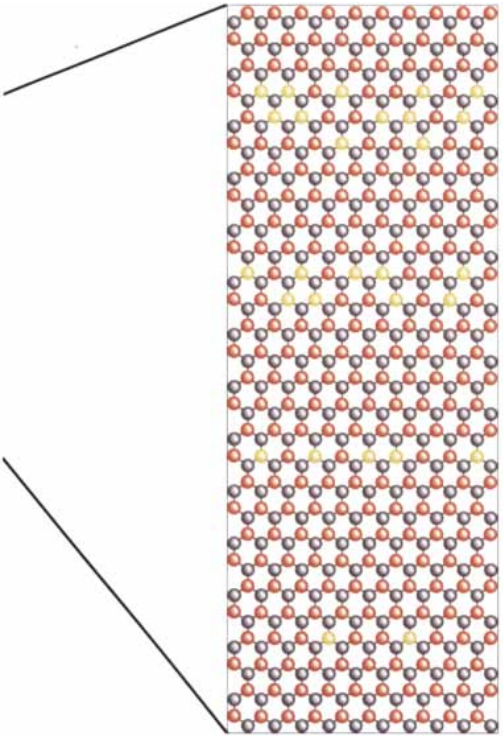
A typical microlaser may consist of more than 500 individual layers. To minimize the power required by the laser, one makes the amplifying medium quite thin, about a hundredth of a micron. The fact that it is so thin means that for the laser to work, light must bounce back and forth many more times than it must for a conventional diode laser. As a consequence, the mirrors must have a much higher reflectivity than the 30 percent reflectivity of the conventional diode laser. In fact, the reflectivity must be 99 percent or more.

Such a high reflectivity is attained by depositing in alternate layers two semiconductors, such as gallium arsenide and aluminum arsenide, that have different indices of refraction—that is, light travels at different speeds through them. The difference in the indices gives rise to a partial reflection at each interface, just as the difference in the indices of glass and air causes a partial reflection from a window. Although the reflection from each interface is only about 0.6 percent, a total reflectivity of greater than 99 percent can be achieved by stacking many alternate layers of the proper thickness.

Ironically, then, the bulk of a microlaser consists not of the approximately hundredth-micron-thick section containing the amplifying medium but rather of the mirrors, which together are four or five microns thick. The relative thickness of the mirrors is well justified, however, because an entire array of microlasers—mirrors and all—can be grown at once, instead of being painstakingly fabricated in the series of steps required to make a diode laser.

Just as important as the ability of molecular-beam epitaxy to form an entire microlaser in one step is the precision with which it can form the thickness of each layer. Careful control of layer thickness is particularly important in fabricating the mirrors. The technique for achieving the necessary control is based on the fact that an individual atomic layer becomes progressively rougher as it grows, until about half of it is in place. Then the layer begins to become progressively smoother until, when it is completely in place, it is atomically smooth.

One can determine how smooth a surface is by bouncing a beam of high-energy electrons off it: the smoother the surface, the more easily the electrons are reflected and detected on a screen on the other side. Hence, the intensity of the reflected beam oscillates in time as successive layers are grown. Accurate measurement of the frequency of the oscillation can precisely determine the time to deposit a single layer



gallium arsenide (*pink*). The actual amplifying medium is quite thin (*pink layer*); the bulk of the device consists of its mirrors (*red and yellow layers*).

of atoms. The resulting exact knowledge of the deposition rate can then be used to strictly control the layer's thickness. Timing of the mechanical shutters with a computer provides precise regulation.

Another advantage of molecular-beam epitaxy is that it enables one to modulate the electrical conductivity and reflectivity of the mirrors. The electric current for pumping the diode must pass through the mirrors, but unfortunately, the current does not flow easily in this instance. The electrical and optical properties cannot be optimized simultaneously, and so microlasers must be designed to accommodate the two competing demands as best as possible.

Once the wafers have been grown, conventional techniques for transferring patterns are used to carve the wafers into individual microlasers. An etch mask, consisting of a two-dimensional array of disks of etch-resistant material, is deposited on the wafer surface by vapor deposition and photolithography, a printing process that uses plates made according to a photographic image. Then, in a novel etching step, a collimated beam of xenon atoms, guided by the deposited mask, vertically etches out each Coke can–shaped microlaser. The resulting sidewalls are smooth, and the surface damage is minimal.

Our group at Bellcore has used these advanced techniques to make arrays of microlasers in which a few layers of one of the mirrors were deliberately varied in thickness. Consequently, the wavelength of each laser differs slightly from that of its neighbor—by an evenly spaced amount. Such arrays could have a major impact in the field of fiber-optic communications, where it is desirable to send multiple signals, each at a unique wavelength, down a single fiber.

The Bellcore group has also made entire arrays of microlasers that can be turned on and off by light pulses. The arrays could be incorporated in information-processing machines that operate by parallel processing, in which elements of a problem are solved simultaneously rather than in sequence [see "Advanced Computer Architectures," by Geoffrey C. Fox and Paul C. Messina; SCIENTIFIC AMERICAN, October 1987]. Related efforts at Nippon Electric Corporation, the University of California at Santa Barbara and Sandia National Laboratories, among others, are also meeting with success.

Future work in microlasers is likely to take two paths. First, the present generation of microlasers has an output at wavelengths of approximately one micron (infrared radiation). It would be desirable to increase the output wavelength to the region ranging from 1.3 to 1.5 microns, because these wavelengths are transmitted with the fewest losses in an optical fiber and are therefore of potential use in optical communications. To date, microlasers operating in the longer-wavelength region have not worked very well, primarily because of a lack of suitable semiconductor materials that can be grown into a stack of layers forming mirrors for those wavelengths. At present, instead of being five or six microns tall, an efficient microlaser working at 1.5 microns would have to be about 20 microns tall, which would take almost 24 hours to grow precisely using molecular-beam epitaxy. Such a structure would not be very practical, but alternative designs still to be demonstrated may prove sufficient.

Further miniaturization, which is also just now being explored, is the second likely line of future research. The road will be difficult; we are already close to reaching the small-size limit beyond which the lasers will no longer work. The reason is that the diameter of the laser must be comparable to the wavelength of the radiation it emits; if the diameter is smaller than the wavelength, it does not contain the light effectively. Admittedly, there is an added bonus of working with gallium arsenide: although the laser wavelength is about one micron in air, within the laser itself it is about three times shorter because of the high index of refraction of the semiconductor. At some size smaller than 0.3 micron, however, a microlaser will no longer work.

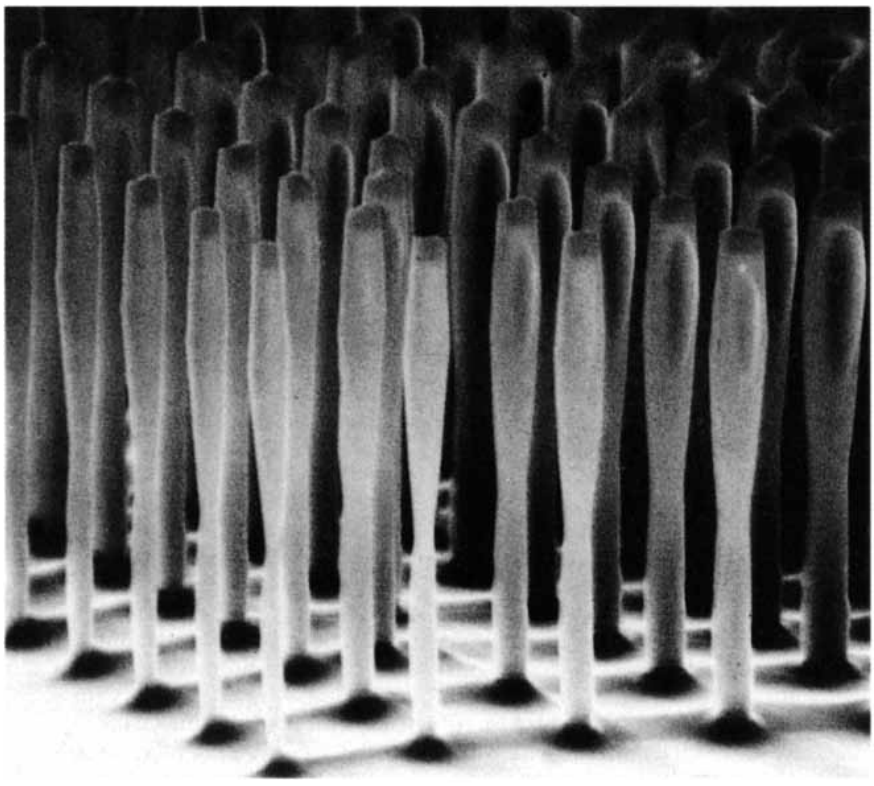
Why bother to shrink the size of microlasers from a relatively comfortable one micron to, say, a difficult one third of a micron? Further reduction could allow an order of magnitude more information to be communicated and processed with the same power requirements. Second, even more fundamental advantages in power and speed could be attained—advantages that are not predicted by simple scaling.

Virtually every laser is inefficient. The first flash of light emitted by the pumped atoms in the amplifying medium goes randomly in all directions. Most of the light does not even strike the laser mirrors, and it is lost. Only a tiny fraction, one part in 10,000 or so, travels in just the right direction to contribute to the working of the laser. It turns out, however, that when the cavity containing the amplifying medium is short and has a small diameter, the walls of the cavity itself, through reflections, can modify the direction in which the light is initially emitted. In a small cavity the light can be emitted only in certain allowed directions and at certain wavelengths. In an extremely small cavity the light is emitted only along the direction of the desired laser beam.

It would therefore appear that microlasers having a diameter of approximately one third of a micron may form the ultimate optoelectronic information-processing device: a small, fast and highly efficient laser requiring low power. Although the prime motivation for building such a device would be for information processing, other applications would undoubtedly be found. Panels of visible-light microlasers could provide efficient lighting, turning the applications full circle from exotic optoelectronic computer technology to a more down-to-earth lamp with which you might read this magazine.

We have experimented to see just how small microlasers can be made and still work. For the purposes of our tests, we pump the atoms in the amplifying

MULTILAYERED STRUCTURE of a microlaser is evident in a scanning electron micrograph. To show the structure, layers of aluminum arsenide have been etched preferentially, with the result that layers of gallium arsenide have larger diameters and project out. Such preferential etching is not done on a working microlaser.



WORLD'S SMALLEST MICROLASERS are about eight microns long and only half a micron in diameter. The authors have successfully tested them with bursts of light.

medium with focused bursts of light instead of electric current. It is easier to fabricate light-pumped devices than electrically pumped ones, and although optically pumped devices are never expected to have any commercial potential, our tests with them have helped convince us (and our managers) that an effort to make electrically pumped cavities of such size is not completely outrageous.

So far we have demonstrated that devices having a diameter as small as one half of a micron can work. Interestingly enough, our tests actually precede any accurate theoretical simulation. Modeling the complex interactions arising from the large number of semiconductor layers would consume huge amounts of computer time and might even require an optical computer!

Perhaps the most exciting aspect of microlaser research is the manner in which it spans the disciplines of fundamental physics, device physics, classical optics, semiconductor optics and optoelectronic and systems integration. Microlasers also challenge investigators' creativity to take advantage of the tremendous applications awaiting discovery. Communications and remote sensing are two areas in which optical systems appear to have fundamental advantages over electronic systems. As machine vision becomes increasingly important, arrays of inexpensive microlasers could allow the technology to be adapted to aid the sight-impaired. We believe the variety of applications of microlasers will be exceedingly wide, but at this point it is only a belief. We look forward to rereading our article 10 years from now to see where we were completely off target and where we were right on.

FURTHER READING

HETEROSTRUCTURE LASERS. H. C. Casey, Jr., and M. B. Panish. Academic Press, 1978.

PHYSICS OF SEMICONDUCTOR DEVICES. Simon M. Sze. John Wiley & Sons, 1981.

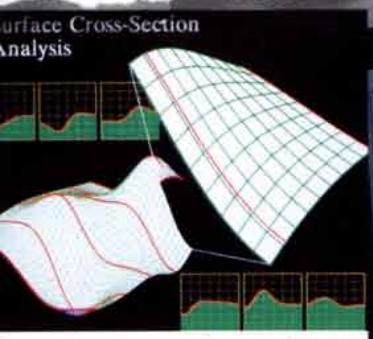
THE TECHNOLOGY AND PHYSICS OF MOLECULAR BEAM EPITAXY. Edited by E.H.C. Parker. Plenum Publishing, 1985.

DIMINISHING DIMENSIONS. Elizabeth Corcoran in *Scientific American,* Vol. 263, No. 5, pages 122–131; November 1990.

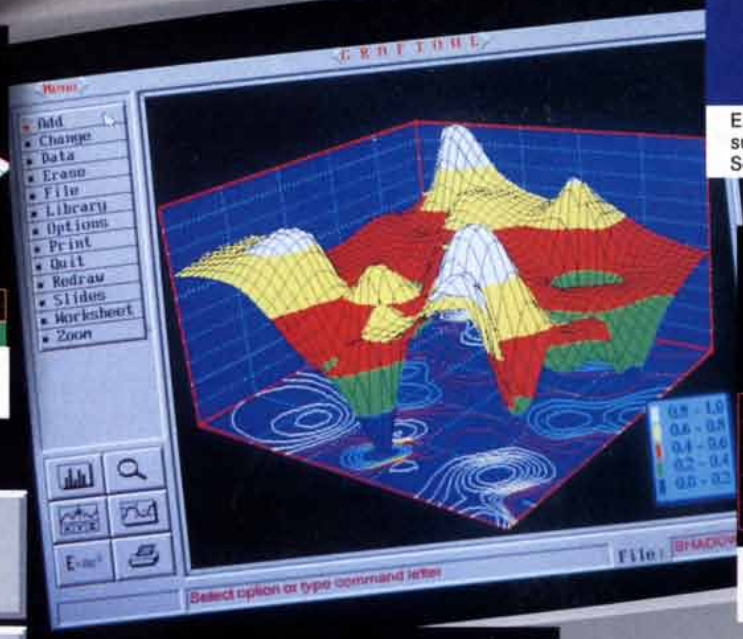
VERTICAL-CAVITY SURFACE-EMITTING LASERS: DESIGN, GROWTH, FABRICATION, CHARACTERIZATION. J. L. Jewell, J. P. Harbison, A. Scherer, Y. H. Lee and L. T. Florez in *IEEE Journal of Quantum Electronics,* Vol. 27, No. 6, pages 1332–1346; June 1991.

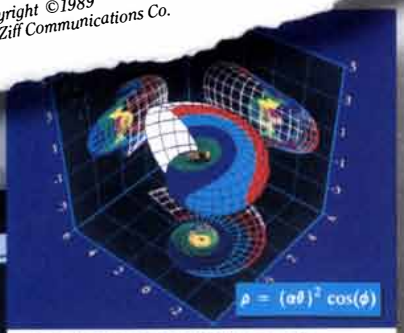
PC MAGAZINE
EDITORS' CHOICE
March 26, 1991
Graftool, Version 3.3

GRAFTOOL "is the best choice..." and "...delivers on all accounts."



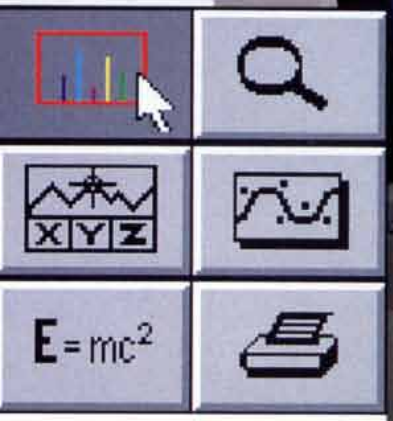
Generate contours, surfaces, and cross-sections from random X, Y, Z points.



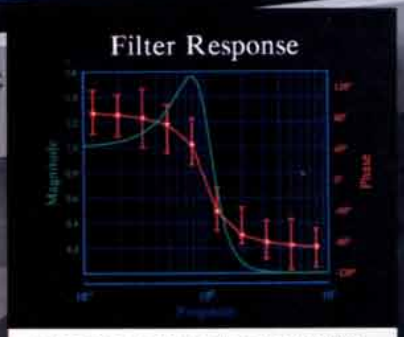


Easily create complex curves & surfaces with the graphical Formula-Solver.™

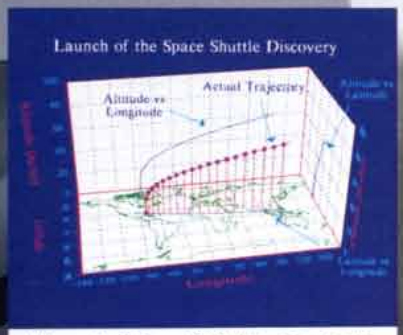
Choose from 26 graph types for business, technical, & scientific data.



NEW push-button "mouse-menus" allow easy access to frequently used features.



Automate customized graph creation with pre-recorded macro & batch files.



Integrate 2-D and 3-D data on a single graph to add meaning and impact.

Presentation graphics and analysis for scientific users.

Your technical data requires more muscle and sophistication than basic business graphics and plotting packages can provide. It's simply a matter of using the right tool for the job. In addition to publication-quality graphics, you need powerful analysis tools and capabilities such as Graftool's ■ Intelligent Data Cursor™ to read out data points on curves & surfaces ■ Linear & non-linear curve-fitting ■ Unlimited zoom & rotation ■ Multiple axes in linear, log, or probability scales ■ Advanced data handling, allowing over 268,000,000 data points ■ Powerful scientific spreadsheet which can directly read your Lotus or ASCII files ■ Unrestricted placement of graphs & text ■ Direct compatibility with Microsoft Word & WordPerfect.

With Graftool, all this power and flexibility purr quietly under the hood, while pop-up menus and push-buttons bring an ease of use previously unheard of in scientific software. Just "point & process" with increased productivity and greater understanding.

GRAFTOOL — the right tool for your technical solutions.

GRAFTOOL $495.

■ Interactive demo available
■ Academic discounts.

3D 3-D VISIONS

Call Toll Free for Details

1-800-729-4723

2780 Skypark Drive, Torrance, CA 90505

Cultured Cells for the Treatment of Disease

The successful growth of human skin cells in culture has made it possible to restore epidermis after severe burns and other forms of damage

by Howard Green

Biologists working on basic problems are always pleased if their research leads to practical application. In 1974 I had no intention of studying the treatment of any disease, when some experiments carried out in my laboratory on a cultured mouse tumor suggested the possibility of a new way of healing human burns. These studies led to a novel method of growing human epidermal cells in culture and, in 1979, to their use in generating epidermis for third-degree burn victims. The application of these cultures was then extended to the treatment of other diseases.

Although it may seem unlikely that the study of a mouse tumor could lead to a treatment for burned humans, the two subjects were linked in a way that could not have been foreseen. My graduate student James G. Rheinwald and I were studying a particular mouse tumor, called a teratoma, which contains pluripotent cells that are able to give rise to multiple somatic cell types. Through experiment, we arrived at conditions that permitted the proliferation of colonies of an unusual cell type. On examination, it became clear that these colonies consisted of squamous epithelial cells, or keratinocytes. This broad category includes cells of the epidermis, the outer layer of the skin, as well as those that form the surfaces of the oral cavity, the pharynx, the esophagus, the cornea and the vagina.

This finding was striking. At that time, there were no known conditions that would support substantial multiplication of cultured human keratinocytes. Could the environment that supported the growth of the mouse teratomal keratinocytes be equally successful for human epidermal keratinocytes? When we tested the latter under the same culture conditions, we were delighted to see that they grew rapidly and that they could be repeatedly subcultured—and therefore grown in large amounts.

HOWARD GREEN is the George Higginson Professor of Cellular Physiology in the department of cellular and molecular physiology at Harvard Medical School. He has devoted most of his scientific life to studies of differentiation, genetics and growth using cultured mammalian cells.

Tissue culture is not a new field. Although it began more than 100 years ago, it did not have any practical consequences or much influence on scientific research until the 1940s. At that time, use of the technique of dissociating cells from one another allowed greater proliferation than could be obtained from tissue fragments. But even today most of the estimated 200 cell types of which the human body is composed cannot multiply extensively under any culture conditions.

Of the few cell types readily cultivable from normal tissue, most are derived from connective tissue. These include fibroblasts, the myoblastic precursors of muscle cells, the endothelial cells that line the heart, lymphatic and blood vessels, and the mesothelial cells that line the peritoneal and pleural cavities, the two large body cavities. In addition to keratinocytes, a few other epithelial cell types, such as urothelial cells and mammary epithelial cells, can now be successfully cultivated.

Because human cells are stable in culture and do not lose their ability to differentiate, it seemed possible that our cultures could be used as grafts for the generation of epidermis. The epidermis is a particularly suitable tissue for regeneration from cultures because more than 90 percent of it is composed of one cell type—keratinocytes. These cells are named for their most abundant proteins, the keratins, a class of filament proteins. Other types of cells also reside in the epidermis, but they are not responsible for its structure. They include the melanocytes, or pigment-forming cells, *T* lymphocytes of the immune system [see "The Immunologic Function of Skin," by Richard L. Edelson and Joseph M. Fink; SCIENTIFIC AMERICAN, June 1985] and special macrophages called Langerhans cells, as well as neuroendocrine, or Merkel, cells.

The epidermis is built of two parts, each consisting of multiple cell layers. The inner part contains living cells; the outer, cellular skeletons, or corneocytes. Although the corneocytes have no metabolism, they confer most of the skin's resistance to chemical and physical attack. In their normally dehydrated state, they are also responsible for the skin's impermeability to water. In the living part of the epidermis, the keratinocytes prepare to form corneocytes by becoming increasingly specialized through a process of terminal differentiation, followed by programmed cell death. Indeed, the keratinocyte is the only cell type that carries out its functions better when it is dead than when it is alive.

In the basal layer of the living part of the epidermis, cells are generated by multiplication. Between its birth and its conversion to a corneocyte, each keratinocyte grows in size and remodels its cytoplasm, so that by the time it is ready to be converted into a corneocyte, it is structurally very different from its younger self.

The proliferating keratinocytes are mainly small cells lying on the basement membrane that separates the epidermis from the underlying connective tissue. The cells of this basal layer normally divide at a low rate, the progeny serving only to replace those cells undergoing terminal differentiation. It may require three to four weeks for a cell generated by a division in the basal lay-

er to move through the living and cornified layers and be shed from the outer surface as a corneocyte.

Underlying the epidermis is a specialized connective tissue, the dermis. Its principal cells, the fibroblasts, make the collagen and elastin that serve as extracellular structural elements. The dermis also contains hair follicles, sweat glands and sebaceous glands, as well as an orderly array of capillary blood vessels. The fluid that nourishes the epidermis originates within the capillaries but becomes laden with fibroblast products as it flows or diffuses through the dermis before it crosses the basement membrane and reaches the epidermis.

These fibroblast products are essential for the optimal proliferation of keratinocytes. By including fibroblasts in our cultures, we were able to provide their products to the keratinocytes. For this purpose, we employed an established line of fibroblasts developed in my laboratory years earlier and called 3T3 cells. After lethally irradiating the 3T3 cells to prevent their own growth, we introduced them in sufficient number to cover the bottom of the culture vessel before adding the epidermal cells. Although it is clear that products of the 3T3 cells are essential for the support of keratinocyte proliferation, the nature of these products remains unknown.

We also discovered that some previously known substances could further improve keratinocyte proliferation. For this reason, we later included in our culture medium epidermal growth factor, cholera toxin—which promotes the growth of the cells, while doing them no damage—and a mixture of growth factors discovered by other investigators.

Once the culture conditions were established, my colleague Yann Barrandon and I studied the proliferation of cultured keratinocytes in considerable detail. When cells of the epidermis are dissociated from one another with the enzyme trypsin, which breaks down junctions between the cells, and placed into a culture vessel, most are unable to form colonies because they have already begun the process of terminal differentiation. Even if the culture conditions are good, only 1 to 10 percent of the cells of an epidermis will proliferate and form colonies. When these colonies are small, the cells multiply exponentially, with a doubling time of about 17 hours.

If, at this point, the cells are again separated and transferred to a fresh culture, they will form colonies—this time, with an efficiency of about 70 percent. In humans, such proliferative zeal would be expressed in the epidermis only under conditions of wound heal-

HUMAN EPIDERMAL CELLS proliferate rapidly in culture. These flasks contain stained keratinocyte colonies grown for seven (*far left*), nine, 11 and 14 days (*far right*). In actual practice, the starting number of cells is greater, so growth is complete by seven to nine days. The confluent epithelium is then detached and used as a graft.

ing and then only for very brief periods. Therefore, we would probably be safe in concluding that our culture conditions support epidermal proliferation better than do conditions in the epidermis itself.

Even populations of growing keratinocytes are not homogeneous, and it is possible to isolate clonal subtypes with greatly differing growth potential. Under poor culture conditions, clones possessing a great deal of growth potential, called holoclones, degrade into clones that have restricted growth potential, called paraclones. For this reason, it is vital to use culture conditions that preserve holoclones.

After less than a week's growth, an epidermal colony becomes a monolayer of about 600 cells. The cells in the central region now begin to stratify and to differentiate terminally. As a result, new layers are added between the basal cells and the culture medium. But the inability of the cells in the upper layers to proliferate means that the entire colony can no longer grow exponentially, and as a result the average growth rate is reduced. To expand the number of cells as rapidly as possible, it is necessary to dissociate the cells before the colonies have begun to stratify, transfer the cells to other vessels and allow colony formation to begin anew.

Because of the separation of proliferating cells from terminally differentiating cells, each large colony acquires a polarity resembling that of the normal epidermis. The basal cells lie on the surface of the culture vessel, as if they were on the basement membrane of the epidermis, and the terminally differentiating cells are suprabasal, just as they are in the epidermis. The epidermis, however, is nourished from the basal side, which is closest to the blood vessels of the dermis, whereas the cultured cells are nourished from the differentiated side, closest to the overlying medium.

In the epidermis the keratinocytes of these different layers are held together by desmosomes, reinforced junctions between adjacent cells. In the cultures, desmosomes form between the cells of each growing colony. As the radius of the colonies increases, at a rate of up to two millimeters a day, the cells of neighboring colonies begin to make contact, and desmosomes form at these points. As the colonies continue to expand, they sweep the supporting 3T3 cells off the vessel surface. When the colonies become confluent, they form a coherent sheet of pure epithelium whose cells are all firmly linked to one another by desmosomes.

The cultured epithelium adheres strongly to the vessel surface but may be simply and gently detached as a stratified sheet by the action of a protein-digesting enzyme, Dispase. This enzyme dissolves the attachments between the basal cells and the vessel without disturbing the desmosomal junctions between adjacent cells.

At this stage, however, the cultured epithelium is only a facsimile of the epidermis. The newly grown tissue is stratified and has polarity, but its cellular organization is irregular, possibly because the cells have grown too rapidly to align correctly. In addition, the normal process of terminal differentiation has not been carried out precisely—for instance, few cornified cells exist on the outer surface, and there are hardly any of the granular cells that normally lie at the outer limit of the living part of the epidermis. The cultured cells also fail to synthesize certain keratins that normally appear during the differentiation of the outer cell layers in the living part of the epidermis. Obviously many changes would have to take place for the cultured epithelium to generate a properly organized and differentiated epidermis.

It remained to graft cultures in order to determine whether the necessary structural and biochemical changes could take place. In our first experiments we grafted human cultures onto athymic mice, which are unable to reject them. The result was a histologically normal human epidermis.

This success gave us sufficient confidence to explore the utility of cultured grafts to treat humans. In 1979 and 1980 Nicholas E. O'Connor and John Mulliken first tested cultured grafts on badly burned adults in Boston at the Peter Bent Brigham Hospital, now Brigham and Women's. Skin biopsies from these patients were brought to my laboratory, the cells were dissociated with trypsin and cultures were grown to confluence in small petri dishes. The epithelium was then detached with Dispase, washed free of extraneous protein, attached to a backing of gauze, brought to the hospital and applied to the patients. About half of the cultures engrafted and generated epidermis.

G. Gregory Gallico III and O'Connor enlarged the scale of this work when they treated children under the care of John P. Remensnyder at the Shriners Burns Institute of Boston. Some patients were burned over 90 percent of their bodies. In a remarkable series of trials, as much as half of the epidermis of these children was regenerated from cultures. Since then, patients have received cultured grafts in Leiden, Genoa, Stockholm and Mexico City as well as in other centers in the U.S., Europe and Canada.

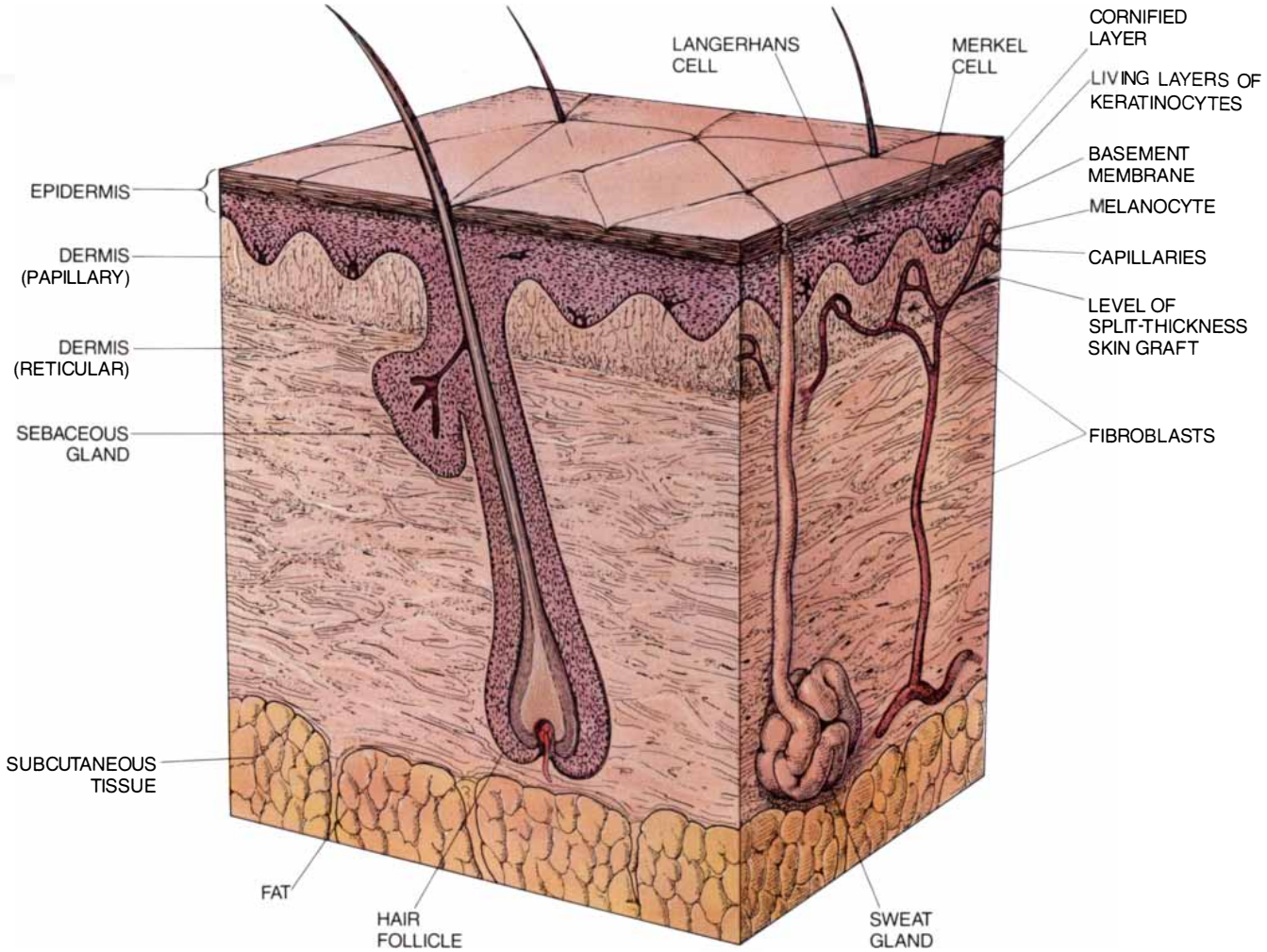
What I have described so far is the generation of epithelium over third-degree burn wounds; in such wounds not only is the epidermis destroyed but also the hair follicles and sweat glands, which ordinarily contain keratinocytes able to regenerate epidermis. Unless it is very small, such a wound cannot heal spontaneously. New keratinocytes must be provided, and they must generate progeny for the lifetime of the patient. For this reason, the epidermal cells of other humans will not do. They may function temporarily, but they will soon be rejected by the patient's immune system. Therefore, the cultured keratinocytes used to generate epidermis for third-degree wounds must come from the patient himself.

The amount of unburned epidermis that needs to be taken for cultivation is small. Starting with a three-square-centimeter biopsy, the area of epithelium can be expanded more than 5,000-fold within three to four weeks—yielding nearly enough epithelium to cover the body surface of an adult human, or 1.7 square meters.

Work on this scale requires numerous laboratory personnel to handle a large number of cultures simultaneously. For instance, to restore epidermis as quickly as possible, it is desirable to cover the entire front or back surface of the body in a single operation. For an average adult, this would require up to 350 grafts of 25 square-centimeters. The logistics of such a process have been worked out by Olaniyi Kehinde, first in my laboratory and later at BioSurface Technology, Inc., in Cambridge, Mass.

Long before cultured epithelium became available, skin grafting was an accepted technique. Since the method was developed in France a century ago, conventional grafts, which are called split-thickness grafts, have been transferred directly from one part of the body to another. Such a graft is about 0.3 millimeter thick and contains the entire epidermis of the surface from which it was cut, together with a somewhat greater thickness of dermis. Epidermis is restored at the donor site by keratinocytes of the hair follicles and sweat glands, whose deeper parts are not removed with the graft. These keratinocytes migrate to the surface of the wound and proliferate until the surface is covered with a new epidermis.

When a split-thickness graft is transferred to another location, it must unite with the wound bed. This union takes place between two connective tissues: the dermis of the graft and the dermis or deeper connective tissue of the



SKIN includes the epidermis and dermis. Dead keratinocytes, or corneocytes, form the outer part of the epidermis, and living keratinocytes form the inner part. The underlying dermis is composed of fibroblasts and their products, which are organized differently in the papillary and reticular layers.

wound bed. The primary adhesion may result from macromolecules present in plasma or tissue fluid, but the junction can have little strength until it has been traversed by fibroblasts and reinforced by their principal products, the collagens. These proteins are deposited in the form of coarse fibers, which are frequently visible as scars.

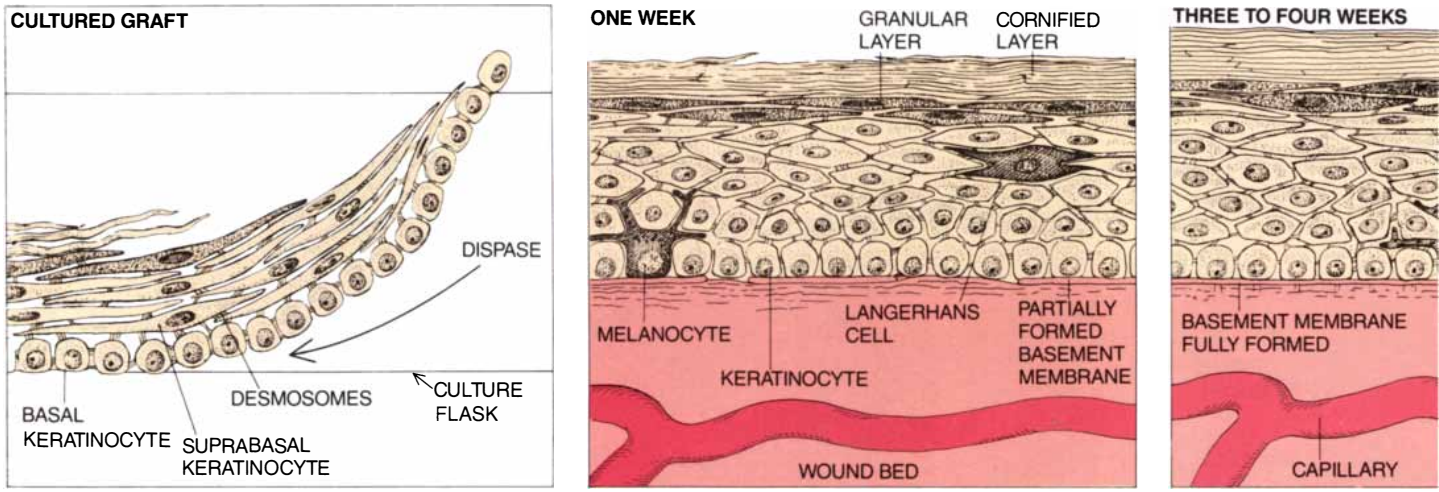
In this union, the epidermis is passive. It has not been greatly disturbed, since it is still attached to the basement membrane and the underlying dermis that was removed with it from the donor site. It seems remarkable that the epidermis survives the process. Until the capillary circulation is extended into the dermis of the graft, nutrient fluid reaching the epidermis must travel a long distance from the capillaries of the wound bed. Passage of the fluid may be somewhat obstructed by coagulated blood, inflammation and, later, by scar.

The union of a cultured epithelial graft with a wound bed is quite different. There is virtually no connective tissue in the cultured graft. Although some fibroblasts may exist among the trypsinized epidermal cells used to start the cultures, their multiplication is suppressed by the irradiated 3T3 cells, and few remain by the time the cultured keratinocytes have become confluent. The graft therefore presents to the wound bed an essentially pure epithelial surface. The direct contact between the two should make it easier for the epithelial cells to be nourished, because no intervening grafted connective tissue separates them from the capillaries of the wound bed. But attachment and anchoring of the epithelium will require structures different from those required by a conventional skin graft.

Conventional grafts are usually cut into a meshlike form so that they can cover a greater area of the wound. This practice, however, leaves interstices to be filled in. The epidermal cells that are to cover these gaps will, like the cultures, have to form an anchoring system. But they will also have to migrate a considerable distance across the wound bed and undergo more intense local proliferation than the cultured cells.

The normalization of epidermal and connective tissue structure that follows the grafting of cultured epithelium has been studied in detail by Carolyn C. Compton and her colleagues at the Shriners Burns Institute. Using biopsy samples taken from regenerated skin over a period of five years after grafting of cultures onto deeply burned surfaces, they found that the different elements of normal structure return at different rates. Within a week, the granular and cornified layers of the epidermis reappear and the organization of the cells becomes more regular. Soon the other normally resident cell types—melanocytes and Langerhans cells—make their appearance.

Melanocytes that are cultured alone do not grow as well as keratinocytes. But Michele DeLuca of the Cancer In-



ENGRAFTMENT OF CULTURED EPITHELIUM is followed by regeneration of the skin. Within a week, the epidermis contains granular and cornified layers, and soon after, it becomes populated by melanocytes, or pigment-producing cells, as well as Langerhans cells of the immune system. The basement membrane is complete by three to four weeks.

stitute in Genoa has shown that the growth rate of melanocytes improves when they are cultured in the presence of keratinocytes. Therefore, the cultured grafts do contain melanocytes. After grafting, the melanocytes multiply and migrate to occupy their normal positions, dispersed in the basal layer of the regenerated epidermis.

The Langerhans cells, which originate from bone marrow, cannot multiply under the culture conditions employed, and probably none are present in the grafts. Nevertheless, these blood-borne cells soon reappear in the epidermis regenerated from the cultured grafts because they have the ability to invade the epidermis from the underlying dermis and must routinely do so in order to maintain their numbers in the epidermis. The epidermis generated from cultures seems to provide all the necessary cues for melanocyte migration and proliferation and for the invasion of Langerhans cells. Among these cues, chemotactic factors are likely to be important.

Within a week after grafting of cultures, basement membrane begins to appear in local patches, and it becomes complete after three to four weeks. Anchoring fibrils, which normally secure the basement membrane to the dermis, also can be detected within a week. They become larger and more numerous as time passes, and within one to two years they become similar in number, size and branching to those of normal skin. In about a year the boundary between epidermis and connective tissue becomes convoluted by the formation of interlocking dermal and epidermal papillae, or projections, as in normal skin.

Most remarkable of all, within two to five years after grafting, the underlying connective tissue acquires a structure very similar to that of normal dermis. The connective tissue near the epithelium differentiates into a fine-fibered papillary dermis. Further away, the dermis becomes coarse-fibered and reticular, or netlike. In the reticular dermis, elastic fibers appear, and in the papillary dermis the blood vessels adopt a typical arched pattern. In summary, over the five-year period so far studied, not only the epidermis but also the underlying dermis has regenerated.

The epidermis appears to be a self-assembling structure. Each keratinocyte has all the necessary information to organize an epidermis, since a single cell isolated in culture and grown into an epithelium will, when grafted, generate an epidermis. The other cell types, when they repopulate the epidermis, receive instructions from the keratinocytes.

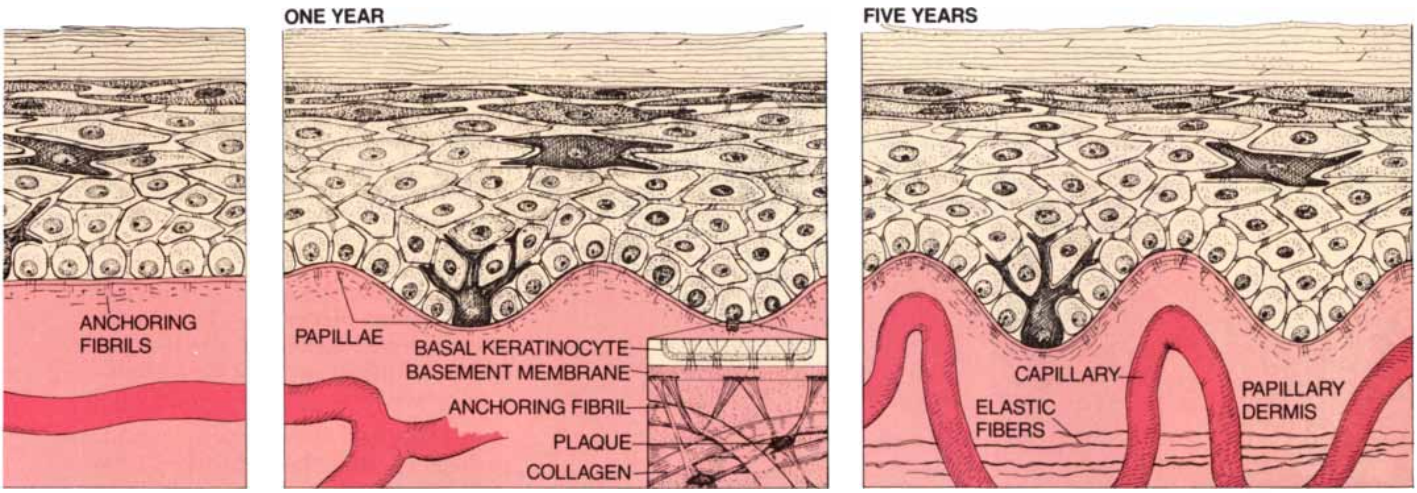
The keratinocytes, however, do not seem to take instruction. For example, the keratinocytes of the palms and soles have site specificity: their histological structure is distinctive, and they make keratins of a type not made elsewhere. When these cells are grown in culture and transplanted to other body surfaces, they keep their original site specificity. This is in contrast to the behavior of epithelial cells in early development, when they are known to be instructed by the adjacent embryonic connective tissue as to the type of differentiation they are to undergo. Evidently epidermal cells learn their lesson well, for they do not seem to forget it after birth.

The successful use of cultured grafts for treating burn victims depends on overcoming a major difficulty. The surface to be grafted is usually highly contaminated with microorganisms, and cultured grafts are more sensitive to this condition than are ordinary skin grafts. Cultures therefore require more meticulous preparation of the wounds and nearly complete elimination of microorganisms.

Among the methods of achieving this goal is the preliminary use of frozen skin grafts obtained from cadavers. When such grafts—called allografts, in contrast to autografts prepared from the epidermis of the future recipient—are applied to wound beds, they can be made to engraft with high frequency. The underlying tissue, now closed over, is soon freed of microorganisms by the defense mechanisms of the patient. The foreign grafts would soon be destroyed by the patient's immune system, but before or at that time they are excised and replaced by the patient's own skin.

In 1986 Charles B. Cuono, Joseph McGuire and their colleagues at Yale University modified this procedure. Rather than removing the entire engrafted allograft, they removed only the epidermal part and grafted the cultures to a bed of residual allodermis. This method achieves over 85 percent success in the engraftment of the cultures and has been adopted at a number of burn centers. In addition to providing a sterile bed for grafting, it may be that the residual allodermis confers some beneficial properties associated with its normal differentiated function, although all allogeneic cells in it will soon be destroyed, and most of the extracellular products will also be eliminated.

After they were successfully used in the treatment of large third-degree burns, cultures were applied to the treatment of other conditions in which it is necessary to regenerate epidermis. Some children, for example, are born with a condition called giant hairy ne-



After a year or so, the anchoring fibrils that attach the basement membrane to the dermis are fully developed, and the boundary between epidermis and underlying connective tissue is convoluted by epidermal and dermal papillae, or projections. Within two to five years, the papillary and reticular dermis regenerates.

vus. In this disorder, the abnormal melanocytes are unsightly and subject to malignant change. Usually giant nevi are excised deeply, and the wounds are covered with split-thickness grafts of unaffected skin. Sometimes, however, the nevi cover so much body surface that it is difficult to find enough unaffected skin to provide the necessary grafts. Since very little skin is sufficient to provide a large number of cultured grafts, these have been used by Gallico of the Massachusetts General Hospital to restore epidermis. Because the surgical wounds are clean, the cultured grafts take with high frequency.

Cultures for restoring epidermis have also been used after removal of scars, particularly those so extensive as to limit movement. Keloids, a condition of excessive scar production, have been treated by excision followed by the grafting of cultures by O'Connor.

Cultured oral keratinocytes have been used to regenerate epithelium in the mouth, and cultured urethral keratinocytes have been used to repair congenital penile defects. In England cultured epidermal keratinocytes have been used to generate an epithelial lining for the chronically infected mastoid cavities associated with middle ear diseases, thereby curing chronic and troublesome discharge from this region.

Cultured grafts have also been used to treat chronic skin ulcers. Irene Leigh and her colleagues at the London Hospital began to study the use of this technique to cure leg ulcers, a common condition in elderly people that is often associated with poor circulation and diabetes. These ulcers may be resistant to treatment and persist for years. In their first trial, a single application of a cultured graft was sufficient to cure a large fraction of such ulcers. This experience was confirmed and extended by a study carried out at the Boston University Hospital by Tania Phillips and Barbara Gilchrest. A similar study by Robert Teepe, now at the Ziekenhuis Leyenburg in The Hague, and his colleagues in Leiden showed that repeated applications of cultured grafts were even more effective. All three groups also found that a cultured graft produced prompt relief of the pain that often accompanies such lesions.

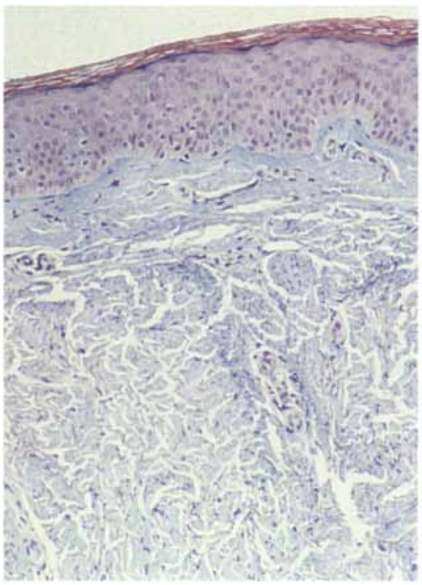
These studies showed that cultured allografts were as effective as autografts. At first, this finding encouraged the idea that, in contrast to skin grafts, cultured epidermal cells were able to survive in a foreign recipient. But when chromosomal or DNA markers were used to distinguish cells of donor and recipient, it became clear that within a short time after application of cultured allografts, no cells from the donor could be identified in the epidermis covering the healed ulcer.

How then were such ulcers cured? Recent studies have provided some clues. We now know that keratinocytes are able to synthesize and secrete a bewildering variety of polypeptide growth factors. One of these, TGF(alpha), is a polypeptide that, like epidermal growth factor, promotes proliferation and migration of keratinocytes. This polypeptide may therefore be considered an autocrine growth factor, stimulating the growth of the very cell type that secretes it. Given the numerous polypeptides now known to be secreted by keratinocytes, there may be more than one with an autocrine effect.

The role of cultured allografts in promoting healing of chronic ulcers may be most easily explained in the following way. Deep within an ulcer, patches of keratinocytes belonging to hair follicles or sweat glands may remain. Since the condition of the ulcer bed is unfavorable to keratinocyte migration and proliferation, the cells do not extend and cover the wound, but they are not destroyed. When suitably stimulated by the autocrine growth factors, they begin to migrate and proliferate. A similar response in the more numerous keratinocytes bordering the wound would shrink its perimeter.

The chemically mediated action that seems responsible for the healing of chronic ulcers is likely to be useful in any type of wound in which keratinocytes survive. For example, in deep second-degree burns there is often a healing problem, even though the deepest hair follicles and sweat glands still supply keratinocytes. The application of cultured allografts has been shown to promote the healing of such wounds, as well as wounds created by the removal of conventional skin grafts. Of course, whenever a skin defect arises from underlying causes, subsequent healing may not be permanent: the healed surface will remain subject to recurrent breakdown unless the underlying causes can be corrected.

Cultures for treating third-degree wounds must be made-to-order because the biopsy must come from the patient. Allogeneic cultures can be prepared from any donor. A practical source of epidermal cells with good growth potential is the skin removed at circumcision of healthy newborn infants. The cultures derived from this epidermis can be exhaustively tested for possible microbial contamination and can be confidently certified as healthy. The



SKIN REGENERATED from cultured epithelium (*far left*) is shown two and one half months after grafting onto allodermis. In contrast to meshed split-thickness grafts, the boundaries between the grafts are invisible; however, the epidermis is pinker and more fragile than normal. With time, the skin becomes paler and more robust. The tissue section (*near left*) shows skin regenerated from a culture five years after engraftment. It reveals the same features as normal skin: a mature epidermis, a convoluted basement membrane and a well-differentiated papillary and reticular dermis.

cells can be stored at the temperature of liquid nitrogen and used as desired to make cultures for grafting. It is now possible to prepare even the final graftable cultures in the cryopreserved state. Such grafts will probably be stored in hospitals, so that they will be instantly available when required for the treatment of wounds.

Learning how to cultivate keratinocytes has enabled us to use them in the treatment of disease. Although this experience cannot be automatically extended to other cell types and other applications, one must think about the possibilities. For instance, might it be possible to use cultured endothelial cells to line vascular prostheses; cultured urothelial cells to repair urinary tracts; cultured pancreatic islet cells to treat diabetes; cultured liver cells to provide hepatic functions or cultured myoblasts to treat muscle disease?

To become practical, all these applications will require substantial advances in knowledge. The poor cultivability of liver cells and pancreatic islet cells is still a serious limitation. While endothelial cells derived from umbilical cord grow well, those taken from adult humans do not. And although human myoblasts can be grown well, it is not easy to use them to generate muscle, because of the complex cellular organization of muscle and difficulties in introducing the myoblasts at the right location.

Also in the future lies the use of cultured cells for gene therapy. Cells synthesizing a product lacking in the body, say, a hormone or an enzyme, would be engrafted into a human and would supply the missing product. It is already possible to introduce a transcriptionally active exogenous gene into cultured cells, which then make the product.

Frequently, however, genes expressed quite well in cultivated cells seem not to be expressed after the cells engraft, perhaps because the transduced gene in a differentiated tissue is subject to some kind of repression not exerted while the cells are in cultivation. Much remains to be learned, not only about how the vectors can be improved to maximize expression but also about what cells can be cultivated and grafted so as to generate an adequate bulk of tissue whose product, derived from an exogenous gene, can reach the right location in the right amount.

More than 500 patients throughout the world have received cultured keratinocytes for the treatment of burns, ulcers or other conditions. This is a very small number in comparison with the number of patients burned or afflicted with chronic ulcers and other conditions, so there is a great need for cultured keratinocytes. As more cell types become cultivable, practical applications will probably be found for some of them as well. Although the necessary components of a medium suitable for established cell lines and fibroblasts have been known for many years, such a medium is not able to support growth of the vast majority of normal cell types. Improving the cultivability of any cell type is very difficult because no general recipe applies. We need to learn more about the biology of each cell type and how to apply this knowledge to its cultivation. At times, something more may be necessary: this may be called insight, intuition or just plain luck.

FURTHER READING

SERIAL CULTIVATION OF STRAINS OF HUMAN EPIDERMAL KERATINOCYTES: THE FORMATION OF KERATINIZING COLONIES FROM SINGLE CELLS. James G. Rheinwald and Howard Green in *Cell,* Vol. 6, No. 3, pages 331–344; November 1975.

GROWTH OF CULTURED HUMAN EPIDERMAL CELLS INTO MULTIPLE EPITHELIA SUITABLE FOR GRAFTING. Howard Green et al. in *Proceedings of the National Academy of Sciences,* Vol. 76, No. 11, pages 5665–5668; November 1979.

GRAFTING OF BURNS WITH CULTURED EPITHELIUM PREPARED FROM AUTOLOGOUS EPIDERMAL CELLS. Nicholas E. O'Connor et al. in *Lancet,* Vol. 1, No. 8211, pages 75–78; January 10, 1981.

PERMANENT COVERAGE OF LARGE BURN WOUNDS WITH AUTOLOGOUS CULTURED HUMAN EPITHELIUM. G. Gregory Gallico III et al. in *New England Journal of Medicine,* Vol. 311, pages 448–451; August 16, 1984.

RECONSTITUTION OF STRUCTURE AND CELL FUNCTION IN HUMAN SKIN GRAFTS DERIVED FROM CRYOPRESERVED ALLOGENEIC DERMIS AND AUTOLOGOUS CULTURED KERATINOCYTES. Robert C. Langdon et al. in *Journal of Investigative Dermatology,* Vol. 91, No. 5, pages 478–485; November 1988.

SKIN REGENERATED FROM CULTURED EPITHELIAL AUTOGRAFTS ON FULL-THICKNESS BURN WOUNDS FROM 6 DAYS TO 5 YEARS AFTER GRAFTING: A LIGHT, ELECTRON MICROSCOPIC AND IMMUNOHISTOCHEMICAL STUDY. Carolyn C. Compton et al. in *Laboratory Investigation,* Vol. 60, No. 5, pages 600–612; May 1989.

SOUTH GERMAN STATE LOTTERY

YOU CAN BE RICH - GET YOUR SHARE OF 440 MILLION $US
1,500,000 tickets sold - 746,308 winning numbers!

... so, nearly 1 out of every 2 ticket wins!

Your greatest opportunity to be a winner.

Unlike any other lottery in the world, the West German Government offers a lottery that produces almost one winner for every two numbers entered. Incredible as it may seem, this ancient lottery, dating back to the 17th century, has *ticket sales limited to 1,500,000 - with exactly 746,308 guaranteed winners.* Yes, that's right, almost every second ticket produces a cash prize. *And there's a pool of over $ 440 million U.S. dollars in prize money that must be awarded to winners.* What a great opportunity!

You can be a lottery winner. Here's how.

Our next lottery begins November 16, 1991 and runs for 26 weeks. Every Saturday during the 26 week period there will be government regulated drawings held in Munich. The winning numbers are drawn and announced under strict government supervision. Your personal ticket will keep you in the lottery each week until you are a winner or for the duration of the lottery. Each month you'll receive a statement of your account, plus the official government list of winning numbers and the monthly drawing schedule. You'll know where you stand at all times!

No West German taxes or deductions.

Your cash prize will be free of West German taxes. *All payouts are 100%, with no deductions of any kind.* The choice is yours...we will pay you in German Marks, U.S. Dollars, or most any currency you select. And this will be done in strict confidence. Winners are not announced publicly. Your individualized ticket number is registered with us in your name prior to mailing, allowing for complete safety and anonymity. You will be automatically notified of each win, promptly and in writing, no matter what amount you win.

Greater investment means greater reward.

Check the order coupon below and decide the amount you wish to invest. A larger investment means that the cash prize will be correspondingly larger. For example, full ticket = full prize, half ticket = half prize, etc. Simply enclose your payment with the order coupon. You may pay with your personal check, money order, travellers check, or bank draft (in U.S. Dollars). And if it's more convenient, we accept major credit cards (a nominal overseas handling fee is charged and payment must be for all six classes).

Multiply your chances of being a winner.

The lottery consists of 6 classes, one class per month with four weekly draws per class, and 2 extra draws in the sixth final and main class. This means that each ticket participates in 26 weekly drawings, unless, of course, it is a winner. Since all tickets are included in the lottery prior to the first drawing, each ticket has an equal chance of winning. Therefore, if you order multiple tickets, the opportunity of winnings becomes such that it is almost impossible to lose.

Summary of Jack-Pots

6 times	$	**2.94 Million**	US
2 times	$	**1.76 Million**	US
8 times	$	**1.17 Million**	US
20 times	$	588,000.00	US
6 times	$	294,000.00	US
22 times	$	147,000.00	US
40 times	$	73,000.00	US
66 times	$	58,000.00	US
82 times	$	47,000.00	US
94 times	$	35,200,00	US
102 times	$	23,500.00	US

and 745,860 prizes up to $ 14,700 US
all prizes are based on full tickets.

World renowned South German Lottery.

As L. E. Schetelig, we are a State Accredited Lottery Agent for the South German State Lottery. We represent thousands of clients like yourself throughout the world. Let us be of service to you. Perhaps you can become a winner of the South German State Lottery. Act today! The next lottery begins

November 16, 1991

For rush-orders Call Toll Free 1-800 628-6974 (from US only). Check time difference (6 to 9 hr.)

Yes, I would like to become a millionaire overnight and order !

All classes (1st to 6th class) 90. Lottery beginning Nov.16, 1991 to May 9, 1992 of the South German Lottery.

Please fill in number of tickets you want to order per month

Payment for all 6 classes one time payment			Monthly payment per class = 1 installment every month		
DM	or	US$*		DM	or	US$*
987.00	•	580.60	Full ticket	164.50	•	96.80
507.00	•	298.25	Half ticket	84.50	•	49.75
267.00	•	157.10	Quart. ticket	44.50	•	26.20

*US$ prices are subject to rate of exchange. Rate as of Aug. 91

Elect the one time payment for all classes or pay monthly, air mail postage and winning list after each class. No additional charges. Make check payable to "L.E. Schetelig"

On mail-order credit-card payment a handling fee of DM 35.00 (US$ 20.60) will be added.

All prizes are calculated at DM 1.70 = US$ 1.00 for your convenience. (Aug. 1991)

Mail Coupon to:

L. E. Schetelig, State Accredited Lottery Agent
P.O.Box 10 34 07
D-3500 Kassel, W.- Germany

90 / 141 Mr. ☐ Mrs. ☐ Ms. ☐ **Answer in English** ☒

First Name — **Please print in clear letters.**

Last Name

Street / P.O.Box

City / ZIP

Country

Payment also accepted by credit cards

Only one time payment accepted for credit card orders.

Charge to my VISA ☐ MasterCard ☐ American Express ☐

Expiry Month ____
Expiry Year ____

Signature _____ **(Only valid with signature)**

Genes, Peoples and Languages

The family tree relating human populations corresponds to another relating the languages of the world. Both trees imply a series of migrations; the biological evidence indicates a homeland in Africa

by Luigi Luca Cavalli-Sforza

More than 40 years ago, when I was studying bacterial genetics in the laboratory of Sir Ronald A. Fisher of the University of Cambridge, the place was saturated with mathematical theorizing. Thus, it is not surprising that I started thinking about a project so ambitious it seemed almost crazy: the reconstruction of where human populations originated and the paths by which they spread throughout the world. I reasoned that the task could be accomplished by measuring how closely living populations are related to one another and by deducing from this information a comprehensive family tree.

The goal is at hand. An exhaustive analysis of human genetic data gathered over the past 50 years, and of new data obtained with recently developed techniques, has enabled my colleagues and me to map the worldwide distribution of hundreds of genes. From this map, we have inferred the lines of descent of the populations of the world. Our tree agrees with another, smaller tree based on fundamentally different genetic data. Moreover, our reconstruction finds striking parallels in a recent classification of languages. Genes, peoples and languages have thus diverged in tandem, through a series of migrations that apparently began in Africa and spread through Asia to Europe, the New World and the Pacific.

The concept of a family tree is crucial to placing such events in their chronological sequence. If other factors are equal, the longer the time since two populations split apart, the greater the genetic difference—or distance—between them should be. Such an analysis could then be applied to a more complex history involving three or more populations.

Human populations are sometimes known as ethnic groups, or "races," if one likes, although racist misuse of the term has made it rather odious. They are hard to define in a way that is both rigorous and useful because human beings group themselves in a bewildering array of sets, some of them overlapping, all of them in a state of flux. Languages, however, gave us a little help.

For much of its history, the human species was organized by tribes, or groups of fairly closely related people. Tribal affiliations continue to be of cardinal importance in traditional societies. In addition, there is often a one-to-one correspondence between language and tribe. Thus, languages offer a rough guide to tribes, and tribal membership, when available, provides a rough classification of populations.

Because the situation is far more complicated in metropolitan societies, we reduced our practical problems by focusing our study on aboriginal populations: those that occupied their present territories before the great migratory waves that followed the voyages of discovery in the Renaissance. Distances between these aboriginal groups cannot be abstracted from the presence or absence of a single inherited trait, or the gene that expresses it, because each group carries practically all the extant human genes. What does vary is the frequency with which the genes appear.

A good example is furnished by the vast set of data for the Rh factor, a human blood antigen that comes in two forms, positive and negative. The character is inherited in a simple way, and it has been studied in thousands of populations for reasons of public health. Physicians must identify pregnant women who are Rh negative and whose fetus is Rh positive and administer an immunologic treatment immediately after delivery. The treatment prevents the woman's body from making antibodies that might injure children she conceives later on. Rh-neg-

Ethnicity and Language

ative genes are frequent in Europe, infrequent in Africa and West Asia, and virtually absent in East Asia and among the aboriginal populations of America and Australia [*see map on next page*].

One can estimate degrees of relatedness by subtracting the percentage of Rh-negative individuals among, say, the English (16 percent) from that among the Basques (25 percent) to find a difference of nine percentage points. But between the English and the East Asians it comes to 16 points—a greater distance that perhaps implies a more ancient separation. There is thus nothing formidable in the concept of genetic distance.

In reality, geneticists use formulas slightly more complicated than simple subtraction so that distances can tell as much as possible about evolutionary history. Should the fragments of a single population become utterly isolated from one another, for example, they will differentiate even in the absence of mutations and natural selection [see "The Genetics of Human Populations," by L. L. Cavalli-Sforza; SCIENTIFIC AMERICAN, September 1974]. Chance alone causes their respective gene frequencies to change, in a process called drift.

When other matters are equal, genetic distance increases simply and regularly over time. The longer two populations are separated, the greater their genetic distance should be. Distance might therefore serve as a clock by which to date evolutionary history. Statistical considerations show, however, that one cannot expect a single gene like Rh to provide an accurate chronology. It is essential to use averages of many genes in the calculus of genetic distances and, ideally, to retest conclusions with different sets of genes. Fortunately, thousands of genes are known, although only a small fraction has been tested in many populations.

There are many principles on which one can reconstruct trees from genetic distances. An example is furnished by a tree linking 15 populations that Anthony W. F. Edwards, now at Cambridge, and I published 27 years ago. The genealogy derives from distances calculated from genetic information then available, according to Edwards's formula for the "minimum genetic path." Essentially, this concept describes the tree having the smallest total branch length. When the tree is projected onto a map of the world so that its branching points match the populations' current homelands, the resulting pattern roughly coincides with anthropological reconstructions of ancient migrations [*see top illustration on page 107*].

Unfortunately, there is no strong evidence that the minimum genetic path provides the best way of fitting a tree to the data. Other tree-building methods may be more satisfactory for relating the length of branches to the passage of time and finding a datable "root" for the tree [*see bottom illustration on page 107*]. When possible, a root relates the populations to an outgroup—say, to chimpanzees, which are believed to have diverged from the line leading to humans between five and seven million years ago. If one assumes that the rate of evolutionary change is constant along all branches, one can equate their lengths to the time elapsed since they diverged. Such rooted trees may also be subject to biases, however, if some branches have undergone more rapid evolutionary change than others.

Mathematical techniques of population genetics can minimize biases by accurately predicting rates of evolution. The evolutionary model we used is the simplest. It predicts that the branches will evolve equally fast, provided that drift has been the major cause of change and that the various populations have been the same size, on average. Independent evidence confirms the former assumption; a judicious selection of populations makes the latter quite probable. Constant rates of evolution are likely when populations are large and live in territories spanning continents and for periods stretching back to original settlement.

With my colleagues Paolo Menozzi and Alberto Piazza of the universities of Parma and Turin, respectively, I designed a common analytical framework to study the history and geography of human genes. In a 12-year project we studied the body of genetic information that has accumulated over the past 50 years—more than 100 different inherited traits from about 3,000 samples taken from 1,800 populations. Most samples included hundreds or thousands of individuals. This set of data, which we call the classical set, is derived indirectly from the proteins that the genes express.

In addition, we recently developed an entirely new second set: molecular data studied directly in the coded sequences of DNA carried in the cell nucleus. Most molecular data we used were gath-

LUIGI LUCA CAVALLI-SFORZA has been professor of genetics at Stanford University since 1971. Born in Genoa in 1922, he earned an M.D. from the University of Pavia in 1944. He studied bacterial genetics in Italy and, from 1948 to 1950, at the laboratory of Sir Ronald A. Fisher of the University of Cambridge. He switched to human population genetics in 1952. Since then, he has studied consanguinity; genetic drift and the means of predicting it through demographic observations; the reciprocal relations between biological and cultural evolution; the cultural significance of names and surnames; and the reconstruction of human evolution. He has conducted fieldwork among African Pygmies and applied molecular techniques for the analysis of genes and for the permanent storage of genetic material taken from aboriginal populations.

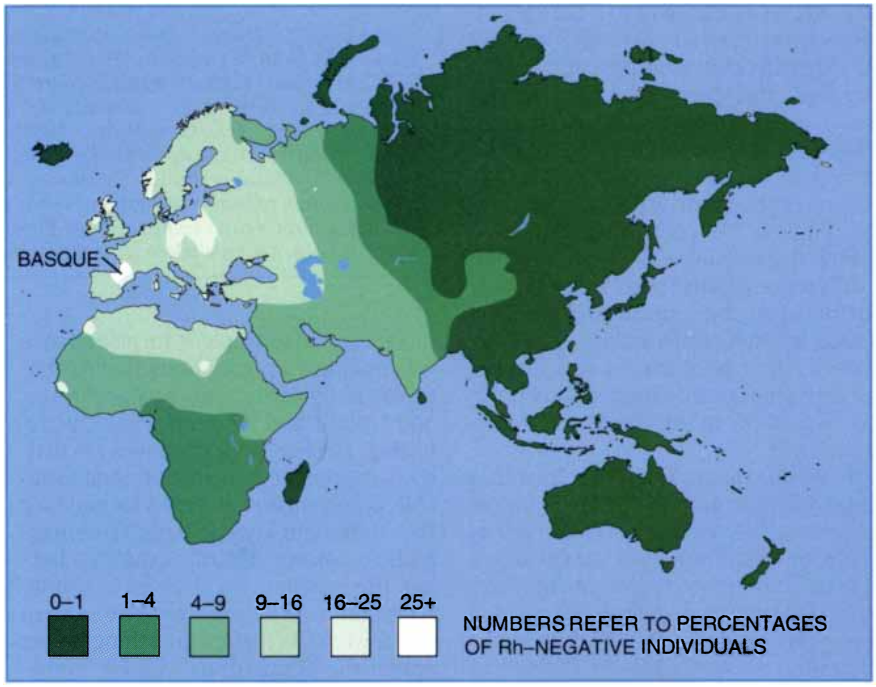
ered in a seven-year study undertaken in a collaborative effort between my laboratory at Stanford University and that of Kenneth K. and Judith R. Kidd of the genetics department at Yale University. Although such data are in many respects of higher quality than those based on gene products, so far they cover only one one-hundredth as many populations. In every comparison we have made thus far, however, the molecular data agree excellently with the classical data.

Our first result supports a conclusion that has emerged from studies of human physical and cultural remains: an African origin of our species. We found that the genetic distances between Africans and non-Africans exceed those found in other intercontinental comparisons. This result is exactly what one would expect if the African separation was the first and oldest in the human family tree.

The genetic distance between Africans and non-Africans is roughly twice that between Australians and Asians, and the latter is more than twice that between Europeans and Asians. The corresponding times of separation suggested by paleoanthropology are in similar ratios: 100,000 years for the separation between Africans and Asians, about 50,000 years for that between Asians and Australians, and 35,000 to 40,000 years for that between Asians and Europeans. In these cases, at least, our distances serve as a fair clock.

A different and quite elegant clock, it turns out, had been devised by other workers studying a kind of genetic data that differed fundamentally from our own. Their most interesting findings became available to us only when our analysis was nearing completion, but they confirmed our findings in all essential points. The set covers the relatively small number of genes encoded in the DNA of the mitochondria, cellular organelles that metabolize energy. We at Stanford had also initially studied such genes, but the late Allan C. Wilson and his colleagues at the University of California at Berkeley developed methodologies having higher resolution. (Here I can do but partial justice to Wilson's many contributions to molecular evolution. He died of acute leukemia in July, at the age of 56.)

Mitochondrial genes differ from those in the nucleus in fundamental ways. Nuclear genes derive about equal contributions from the father and the mother, but those in mitochondria are passed to offspring almost exclusively by the mother. This simple mode of inheritance makes mitochondrial genes very



GENE MAP shows the Rh-negative factor to be most common among the Basques and less common further west. Such data suggest that the Basques preserve vestiges of an early European population that later mixed with newcomers from Asia.

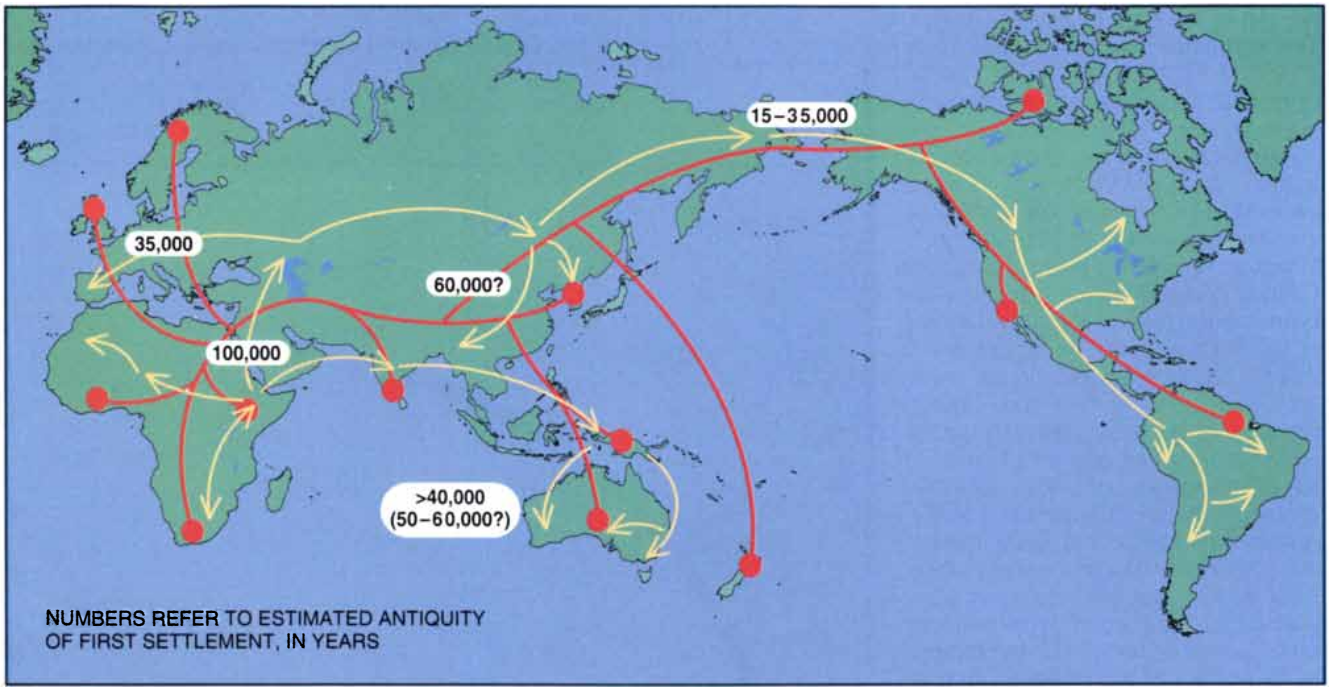
convenient for the estimation of genetic distances. They also have higher mutation rates than nuclear genes, so that one may in part alter the statistical determination of genetic distances, calculating them not from the frequencies of genes but from mutations in the genes themselves.

The mitochondrial clock is based on the number of mutations that have accumulated rather than on the changes in gene frequencies. Whereas we hypothesized that our gene frequencies had drifted apart at constant rates, the Wilson group hypothesized that their mitochondrial genes had mutated at constant rates. Because of the nature of the data, it is easier to provide a root for a mitochondrial tree than for a nuclear one. One need merely compare one's tree with an external group—the Wilson group used chimpanzees—that is known to have diverged at a given date or range of dates.

From such distances, the Wilson group derived a tree of descent that showed more differentiation in Africa than anywhere else. That finding indicated that human mitochondrial DNA had been evolving for the longest time in Africa—that is, it can be traced to a single African woman. In addition, the workers were able to date the branching points of the tree by comparing DNA from humans and chimpanzees, whose lineages were known to have diverged about five million years ago. Their tree thus calibrated, the Wilson investigators were able to estimate the dates of its later branches. Most important, they were able to estimate that the African woman had lived 150,000 to 200,000 years ago. They have therefore confirmed our conclusions by completely independent means.

Recently the workers have brought their estimated date forward somewhat, but their African woman still precedes the date we assign to the divergence of African and Asian populations. In fact, she *should* be more ancient: the two dates refer to different events—the birth of an individual woman and the splitting apart of a population to which she belonged. The news media confused matters by giving wide circulation to the label "Eve" for this woman. In fact, we have no evidence that there ever was a time when only a single woman lived on the earth. Many other women might have lived at the same time, but their mitochondrial lineages simply went extinct.

Some of these conclusions remain controversial. Although paleoanthropologists agree that the genus *Homo* originated in Africa about 2.5 million years ago and that fossil evidence of anatomically modern *H. sapiens* appears only around 100,000 years ago, in Africa or near it, not all of them accept the "out of Africa" theory. One group maintains that modern humans emerged at a much earlier time and in many Old World populations all at once [see "The Emergence of Modern



GENES AND STONES tell similar stories. The earliest genetic tree (*red*) is projected onto a map; the ends have been placed in current homelands (*red dots*). More recent genetic work implies two routes of migration from Africa to Asia (*yellow*); details of the routes are speculative. Archaeological estimates of first-settlement dates appear beside migration routes.

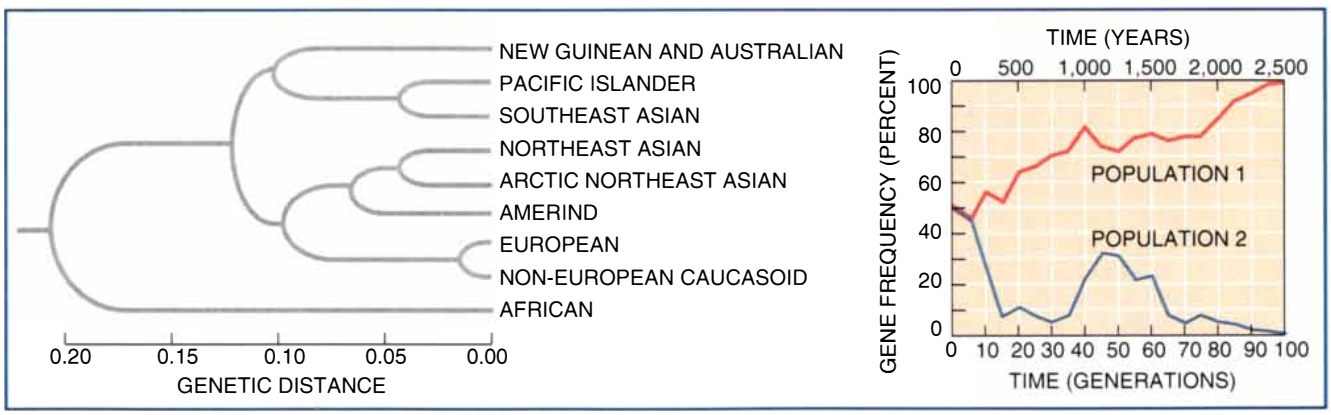
Humans," by Christopher B. Stringer; SCIENTIFIC AMERICAN, December 1990].

Not only have we traced the earliest modern humans to Africa, we have also recovered evidence of a series of migratory waves. This pattern tells much about the origin of existing populations. Moreover, the interplay of our work with that of linguists and archaeologists promises to uncover still more detailed information.

In general, migratory processes reflect changes that may be viewed at once as pressures and opportunities. At many times, humans and their hominid progenitors have been able to increase in numbers greatly and thus to expand geographically. Such demographic success must have stemmed in the main from cultural developments, which for the prehistoric period must be inferred from the archaeological record. This record—bones and stone implements, for the most part—shows that Africa was indeed the original homeland of hominids. From there, migrations must have proceeded from Africa to Asia via the isthmus of Suez and, later, from Asia to Europe. These regions were settled by hominids by perhaps a million years ago.

The next stage is harder to recover because it depends on the time at which one imagines modern humans emerged from the hominid stock. In any case, it is clear that this emergence had already occurred when humans expanded from Asia to the Americas—an event that had to await a time when the Bering Strait was dry and the climate mild enough to permit passage by land. The settlement of Australia and the Pacific islands must also have been accomplished only recently, after the mastery of marine navigation.

Australia appears to have been settled by migrants from Southeast Asia at least 40,000 years ago and perhaps 10,000 to 20,000 years earlier than that. Archaeologists are divided, however, on the first entry into the Americas. So far the first fully convincing signs of humans in Alaska are dated to about 15,000 years ago. There seem to be ear-



CHANGE OVER TIME produces genetic differentiation, such as that reflected in this ethnic family tree (*left*). Drift, the mechanism of change, can be modeled by computer (*right*). When two halves of a population are first separated, they carry a gene at equal frequencies, but time and chance can eventually push them in opposite directions.

lier dates for sites in South America. The estimates therefore range from 15,000 to 35,000 years ago. Our nuclear genetic data suggest the settlement began around 30,000 years ago.

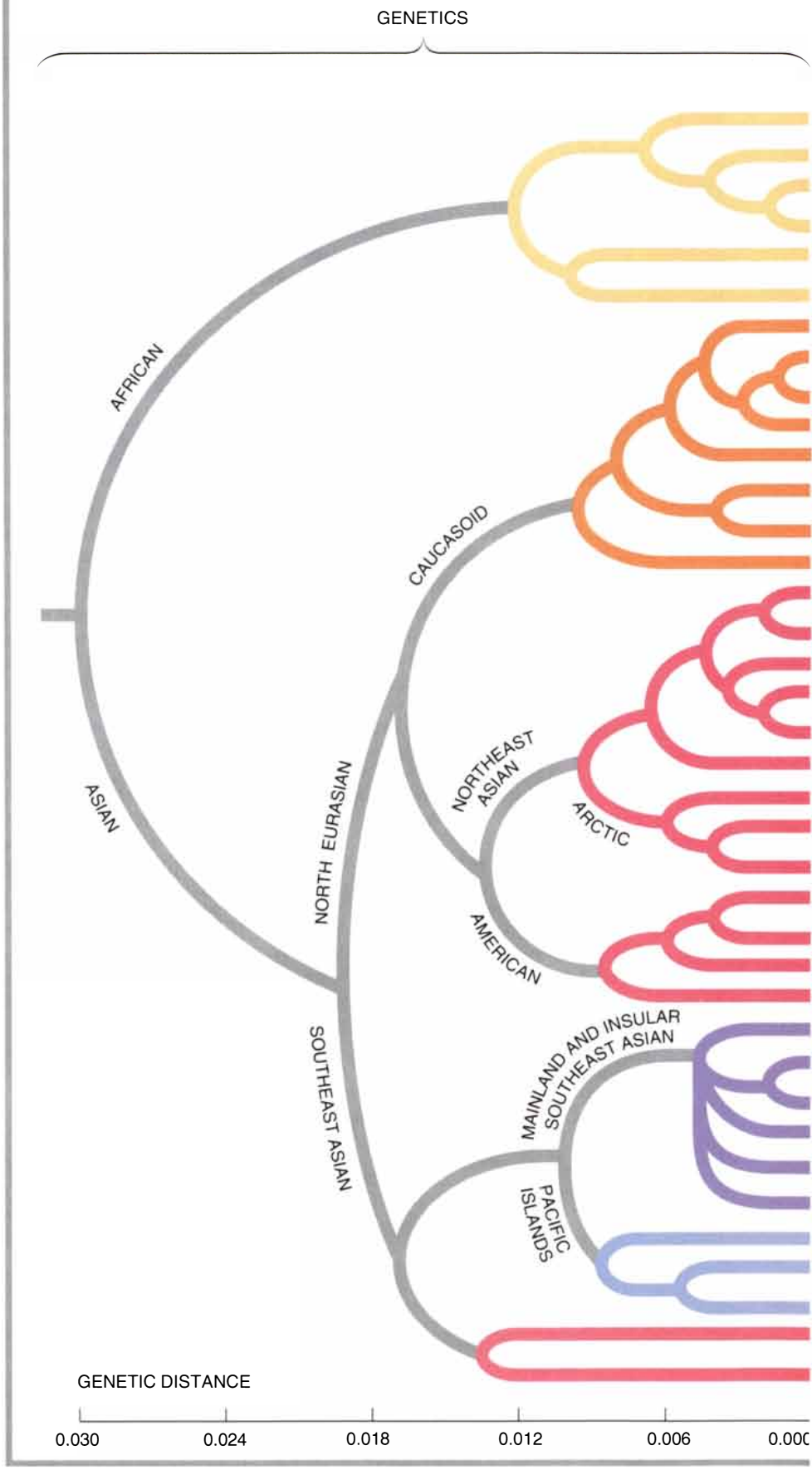
Europe has been swept by many migratory waves, but vestiges of the earliest settlement remain. A key to the puzzle was suggested in 1954 by Arthur E. Mourant, then at the Medical Research Council Population Genetics Laboratory in London, one of the first students of "gene geography." He hypothesized that the Basques (a population in northern Spain and southwestern France) were the oldest inhabitants of Europe and that they had conserved some of the pristine genetic constitution despite contacts with later immigrants. The theory is supported by data on the Rh-negative gene: the Basques have a higher frequency for the gene than any other population in the world. Work on many other genes confirms the hypothesis, as do studies of the profound linguistic differences between the Basque language and those of neighboring peoples.

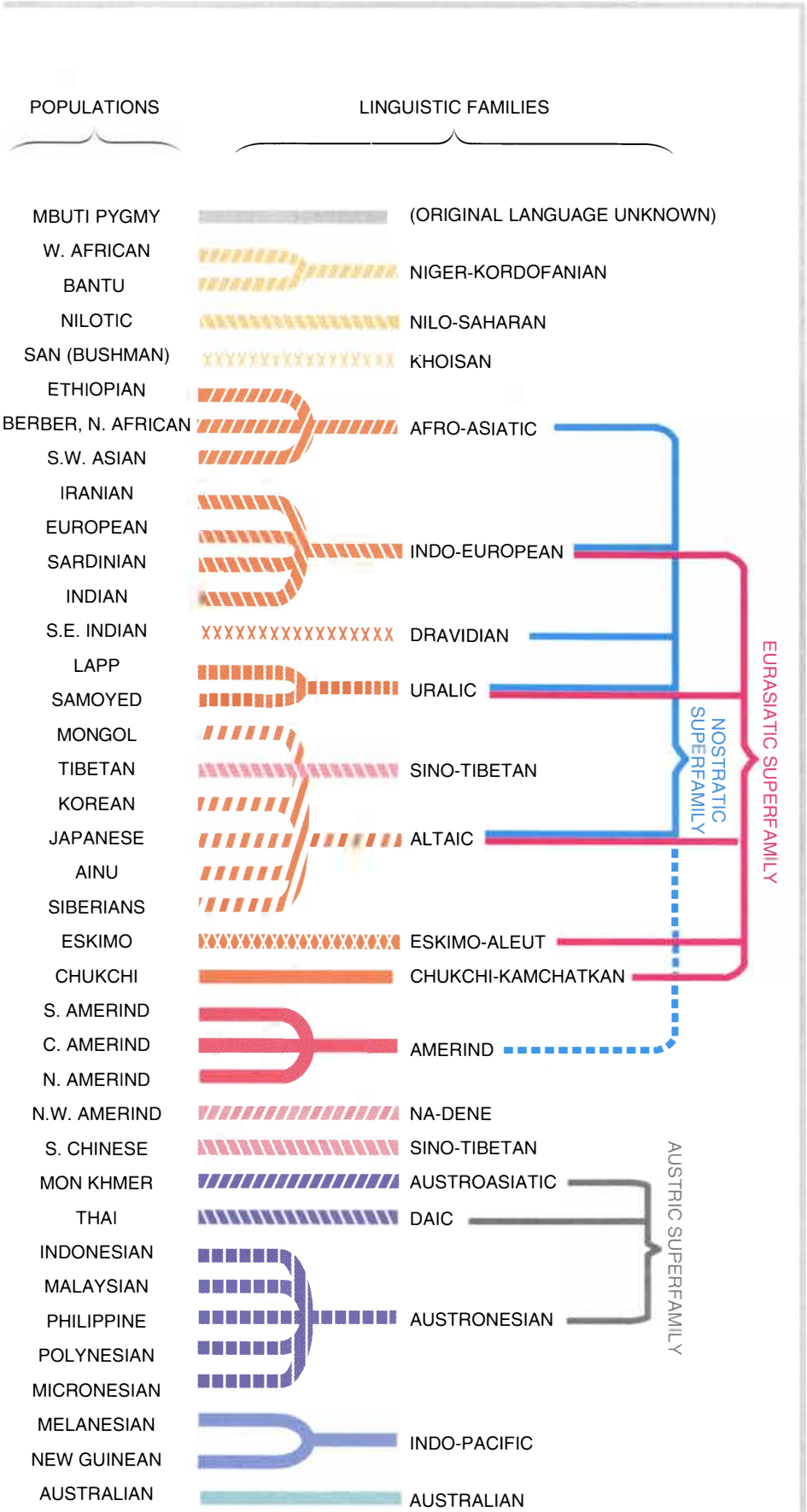
Indeed, a recent analysis of how genes vary from one part of Europe to the next suggested a model of how all of Europe was settled. This model holds that early Neolithic farmers brought their genes, culture and Indo-European languages from the Middle East to Europe in a process of slow expansion [see "The Origins of Indo-European Languages," by Colin Renfrew; SCIENTIFIC AMERICAN, October 1989]. Because the Basques' forebears lived at the far end of that migratory path, they presumably underwent the least genetic admixture with the farmers.

It should be noted that one can only hope to recover a pattern of settlement that reflects successful migrations. There might have been failures, too. In the Americas, to take a very recent example, the Vikings established short-lived settlements, but the genetic contribution they might have made to the local gene pool is unknown.

Our third major finding was that the distribution of genes correlates surprisingly well with that of languages. We concluded that, in certain cases, a language or family of languages can serve to identify a genetic population. A striking example is provided by the nearly 400 languages in the Bantu family of central and southern Africa, which are related to one another and correspond closely to tribal boundaries and the genetic affiliations among tribes. The reason this should be so had been advanced on linguistic grounds in the 1950s by Joseph H. Greenberg of Stanford.

Correlation of Peoples and Languages





Linguistic classification from Merritt Ruhlen, A GUIDE TO THE WORLD'S LANGUAGES

Greenberg's hypothesis, which has since won wide support, was that the Bantu tongues are descended from one tongue or a handful of closely related dialects spoken by early farmers in eastern Nigeria and Cameroon. As the farmers expanded into central and southern Africa beginning at least 3,000 years ago, their languages diverged but not so much as to obscure their common origin. Because the explanation applies to the genes of these populations, Bantu—originally a linguistic category—can now be extended to designate a group of populations having both a linguistic and a genetic basis.

In 1988 my colleagues and I published a genetic tree of evolutionary origins of 42 world populations, together with their respective linguistic affiliations. The tree demonstrates that the genetic clustering of world populations closely matches that of languages. With very few exceptions, the linguistic families seem to have a relatively recent origin in our genetic tree. Moreover, recent attempts by two groups of linguists to generate higher linguistic categories ("superfamilies") gave results that were also in line with the totally independent genetic evidence. It was exciting to discover that we had confirmed a conjecture made by no less a pioneer than Charles Darwin, who stated, in chapter 14 of *On the Origin of Species by Means of Natural Selection,* that if the tree of genetic evolution were known, it would enable scholars to predict that of linguistic evolution.

Why should genetic and linguistic evolution correspond so closely? The answer lies not in genetic determinism but in history: genes do not control language; rather the circumstances of birth determine the languages to which one is exposed. Linguistic differences may generate or reinforce genetic barriers between populations, but they are unlikely to be the leading cause of the correlation. Human evolution is punctuated by the splitting of populations into parts, some of which settle elsewhere. Each fragment evolves linguistic and genetic patterns that bear the marks of shared branching points. Hence, some correlation is inevitable.

One may object that complete separations, such as those established when a splinter group migrates to a new continent, must be rare. But one does not need oceans or mountain ranges to separate populations: simple distance will do the job, as genetic studies in many species prove. Because migratory interchange is normally greater at short distances, one expects and finds a higher degree of genetic difference the farther apart two subgroups are placed. It

is just the same for languages. When there are no special barriers, both genetic and linguistic variations tend to be continuous, and discontinuity tends to appear in both when there are some barriers to free intermigration.

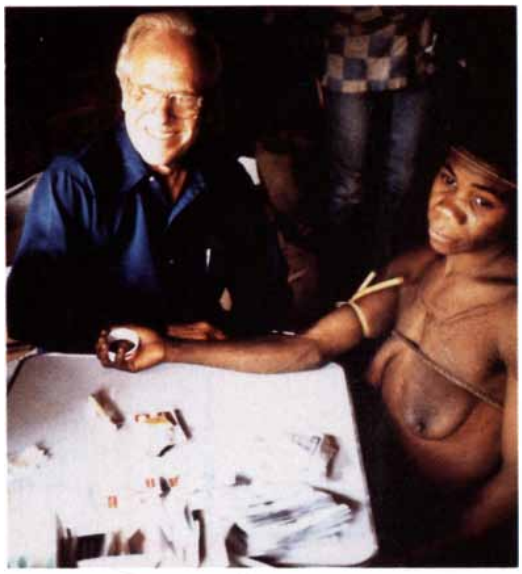
Two kinds of exceptions should be noted to our rule of the correspondence between genes and languages: those in which there is a replacement of language and those in which there is a replacement of genes. The former occurs when people give up their ancestral language for a new one, perhaps that of immigrants, conquerors or a newly risen cultural elite. Such replacements do not always occur, however, and they are less likely when the new language derives from a different family. Basque is an extreme case of a relic language that has evidently survived through thousands of years of continuous linguistic turnover in neighboring regions.

Gene replacement, usually partial, occurs when one population mixes with another. The mixing can be perfectly gradual, affecting the relative frequency of all genes in like proportion. This gradualism sharply distinguishes genes from languages, which in principle are either replaced or not. A language retains its ancestral integrity even if it adopts vast numbers of words from another linguistic family or subfamily. Linguists agree, for example, that English remains a member of the Germanic subfamily despite its borrowings from the French, the Greek and the Latin. What matters is that the structure and basic vocabulary retain family traits.

This difference means that when a tiny minority imposes its language on a conquered majority, language replacement is close to complete, but gene replacement is in proportion to the demographic ratio. Hungarians, for example, speak a language from the Urals (which divide Europe from Asia) imposed by the Magyar conquerors of the Middle Ages but carry a European genetic pattern. Only with some difficulty can one detect any traces of the Magyar genes in the modern population.

Large-scale gene replacement is perhaps rarer. But at least one likely example can be seen in our pair of complementary trees: the Lapps, or Saame, of northern Scandinavia. Their language also belongs to the Uralic family, but their genetic pattern suggests a mixture between Mongoloid peoples of Siberia and Scandinavians, who are responsible for the majority of their genes. Genetic admixture is evident also in the Lapps' hair and skin, which vary from extreme light to dark. A situation not unlike that of Lapps applies for Ethiopians, a genetic mixture of Africans and Caucasoids from Arabia, with a predominance of the former.

Even a modest trickle of genes can produce great effects if it continues long enough. A classic example is that of African Americans, who today derive on average 30 percent of their gene pool from people of European ancestry. This is the mixture that would have resulted had 5 percent of all black unions been with Europeans in each generation since the institution of American slavery and had all the progeny been classified as black. Another 1,000 years of such flow would leave but little of the original African genome.



GENETIC SAMPLE is taken from a member of the Aka tribe of African Pygmies by the author.

It is perhaps surprising that so much of the expected correlation between languages and genes remains, despite the blurring caused by gene or language replacements. In part, this may reflect our concentration on aboriginal populations. In any case, other analyses now confirm the existence of this correlation at a microgeographic level, sometimes in a dramatic way. Perhaps the most striking example is furnished by the close agreement between our analysis of genetic patterns in Native Americans and Greenberg's recent classification of New World languages into three major families. The two studies proceeded independently, using quite different kinds of data, yet each strongly implies that there were a handful of discrete migrations into the Americas.

The ultimate explanation of this correlation of genes and culture must be sought in the two mechanisms of transmission: horizontal and vertical. Genes, always transmitted from parents to children, describe a vertical path through the generations. Culture can also pass vertically from generation to generation, but unlike genes, it can also be transmitted horizontally, between unrelated individuals. High fashion, for example, is generally transmitted from Paris to the rest of the world every season (although Italy now appears to be taking the lead). In the modern world horizontal transmission is becoming increasingly important. But traditional societies are so called precisely because they retain their cultures—and usually their languages—from one generation to the next. Their predominantly vertical transmission of culture most probably makes them much more conservative.

Gene and language replacements are more than annoying exceptions to our rule. Each exception operates according to rules of its own, which should explain much about the evolution of populations and languages and hence of the development of human culture. Studies of such replacements could thus complement the work we have done. Anthropological fieldwork must catch up with such tools and with the rapidly vanishing data. Priceless evidence is slipping through our fingers as aboriginal populations lose their identity. Growing interest in the Human Genome Project may, however, stimulate workers to gather evidence of human genetic diversity before it disappears.

FURTHER READING

CULTURAL TRANSMISSION AND EVOLUTION: A QUANTITATIVE APPROACH. L. L. Cavalli-Sforza and Marc W. Feldman. Princeton Univeristy Press, 1981.

DRIFT, ADMIXTURE AND SELECTION IN HUMAN EVOLUTION: A STUDY WITH DNA POLYMORPHISMS. A. M. Bowcock, J. R. Kidd, J. L. Mountain, J. M. Hebert, L. Carotenuto, K. K. Kidd and L. L. Cavalli-Sforza in *Proceedings of the National Academy of Sciences,* Vol. 88, No. 3, pages 839–843; February 1, 1991.

RECONSTRUCTION OF HUMAN EVOLUTION: BRINGING TOGETHER GENETIC, ARCHAEOLOGICAL AND LINGUISTIC DATA. L. Cavalli-Sforza, A. Piazza, P. Menozzi and J. L. Mountain in *Proceedings of the National Academy of Sciences,* Vol. 85, No. 16, pages 6002–6006; August 1988.

HISTORY AND GEOGRAPHY OF HUMAN GENES. L. L. Cavalli-Sforza, P. Menozzi and A. Piazza (in press).

We've found the perfect way to talk to Europe's decision-makers.

English with a German accent.

In Europe, as in the U.S., money talks. Only in Europe, it doesn't always speak the same language.

For the U.K. and a broader, pan-European market, the business language of choice is English.

But Europe's largest and most important business market, Germany, speaks an entirely different language.

So how do you talk sales in what will shortly be the world's largest integrated consumer marketplace?

The answer is "Europe Now".

The Wall Street Journal Europe has teamed up with Handelsblatt, Germany's only national business daily, for a media package that bridges the communication gap.

The Journal Europe provides strong European presence while Handelsblatt puts the emphasis on Germany.

By buying them together as "Europe Now", you gain a highly efficient and economical alternative to separate buys of Europe's leading business publications.

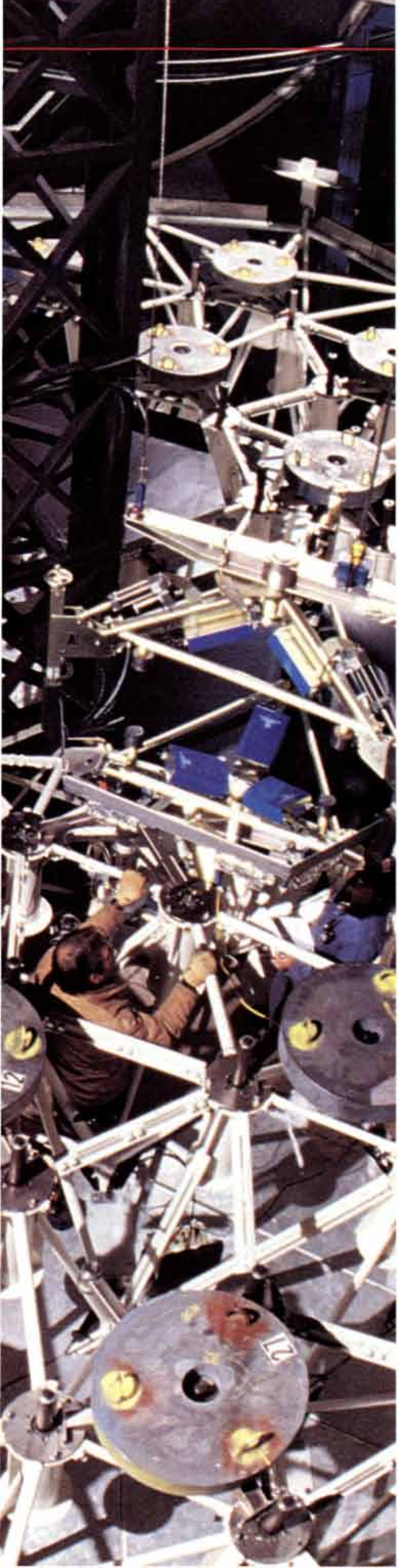
So if you've got something to say to Europe's most influential business people, talk to Guy M. Holroyd (Handelsblatt) at 1-212-754-0850 or Dave Manchee (The Wall Street Journal Europe) at 1-212-808-6690 first. They'll provide the details in a language anyone can understand. Numbers that speak for themselves.

Europe Now... for Europe today.

MIRRORING THE COSMOS

by Corey S. Powell, *staff writer*

Photographs by
Roger Ressmeyer/Starlight



A new generation of telescopes that embrace huge light-collecting mirrors and cunning optical technology promises a tremendous leap in astronomers' ability to explore the universe.

The engineers working inside the dome of the Keck Telescope on August 8 were barely aware that Hurricane Fefa was bearing down on the instrument's 4,200-meter-high perch atop Mauna Kea. Their attention was concentrated instead on a hexagonal mirror that was being carefully lowered into position and precisely aligned with nine other mirror segments already in place.

Although nobody applauded, that 10th mirror represented a milestone: the Keck Telescope had just surpassed the light-gathering area of the famous Hale Telescope on Mount Palomar in California. When the 36th and final segment is installed early in 1992, the Keck's completed mirror will stretch 10 meters across, making it by far the largest telescope in the world—for a time.

The Keck is only the first of a family of optical behemoths built around radically new mirror designs that will dwarf the venerable Hale. Astronomers have built a procession of ever larger instruments since Galileo turned his crude spyglass skyward in 1610, but the Hale seemed to mark the pinnacle of telescope engineering when it was completed in 1947. Its five-meter mirror weighed 20 tons and required more than a year to cool before it could be ground. Grinding, interrupted by World War II, eventually required 11 years to complete.

The very magnitude of that effort cast doubt on the notion of building instruments significantly larger than the Hale. Doubling the dimensions of the Hale mirror would make it eight times as heavy and therefore prone to sagging. The telescope mount and observatory dome also would have to be much larger, more complex and therefore very costly. Moreover, atmospheric turbulence in and above the observatory dome can negate the better resolution that should theoretically accompany a larger telescope.

Yet astronomers have continued to raise increasingly provocative questions that can be answered only by pushing the technology of their telescopes. They hunger for more light so they can study distant and faint objects that may hold the clues to such grand mysteries as the origin (and ultimate fate) of the universe or that may reveal how galaxies like our own Milky Way formed. Some of the most fascinating objects in the cosmos—from energetic quasars and disrupted galaxies to stars in the process of being born and possible planets around distant suns—remain enigmatic simply because they cannot be observed with sufficient clarity. Or, as John Kor-

MIRROR SEGMENTS in the Keck Telescope are anchored to a motorized support structure and aligned with one another to within one ten-millionth of a meter. A full complement of 36 mirror segments will make the Keck the world's largest optical telescope.



KECK TELESCOPE uses a segmented-mirror design proposed by project scientist Jerry Nelson (*left*). In 1987 workers began assembling the telescope's dome. Two years later they installed the telescope structure; in October 1990 the first

mendy of Dominion Astrophysical Observatory in British Columbia puts it, "the science gets better very rapidly as the seeing gets better."

An obvious way to get a clearer view of the cosmos is to place a telescope in space, above the earth's obscuring atmosphere. But the orbiting *Hubble Space Telescope* has failed to live up to expectations. And in the past few years, telescope makers discovered ways to overcome many of the limitations of previous earthbound instruments.

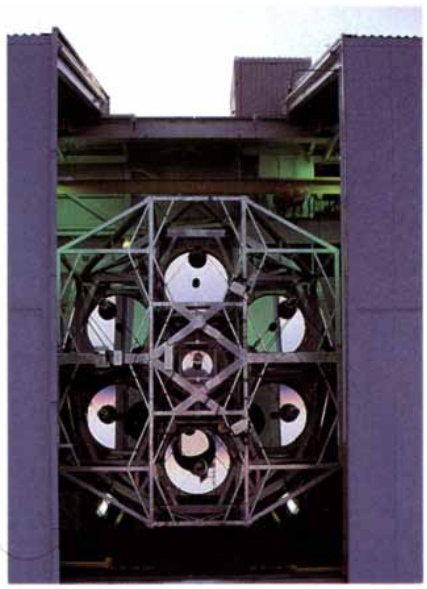
Recent breakthroughs in casting techniques permit opticians to fabricate huge mirrors that are relatively lightweight and inexpensive. Simultaneous advances in grinding technology make it possible to form these mirrors into an extremely high curvature. The steep, almost cereal bowl–shaped surfaces of these mirrors reflect the light to a focus very close to the mirror. Such so-called fast mirrors have a great advantage: they permit the construction of extremely stubby telescopes and far smaller (and hence less expensive and more stable) observatory structures.

A technique called adaptive optics promises unprecedentedly sharp views even from telescopes that are modest (one to four meters in size). In this approach, a computer-controlled correcting mirror precisely distorts its surface so as to cancel out distortion caused by the turbulent atmosphere. "Ground-based astronomy is getting more like space astronomy," reflects Alan N. Stockton of the University of Hawaii's Institute for Astronomy. "There are very complex systems in which everything must work perfectly or the whole telescope fails."

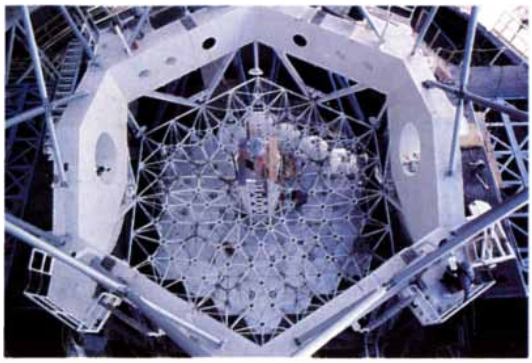
Jerry E. Nelson of the University of California at Berkeley, the project scientist for the Keck Telescope, realized that one way to circumvent the difficulties of casting and grinding a single huge mirror is to build it out of smaller pieces. The Keck Telescope will ultimately equal the area of a single 10-meter mirror. The benefits of segmented mirrors are plain to see: the Keck's mirrors are only 7.5 centimeters thick, yet each is small enough so that it remains rigid and does not sag.

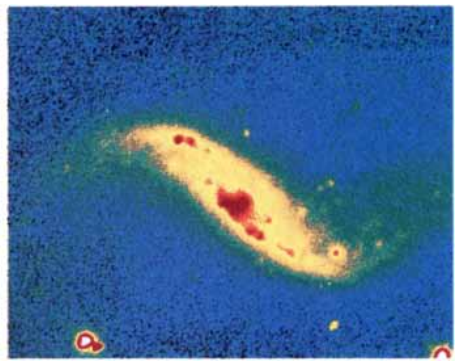
In choosing a segmented-mirror design, Nelson uncorked a few touchy technical challenges. Each Keck mirror is only one element of the overall reflecting surface, and so it must be ground to a very specific, complex shape. After considerable experimentation, Nelson's team discovered it could attain the desired curvature by bending each segment in just the right way during the





OPTICAL TELESCOPE DESIGN has evolved quite slowly. The benchmark five-meter Hale Telescope (*left*) was completed in 1947. It was surpassed in 1976 by the Soviet six-meter telescope (*center*), which uses a clever, simplified mount design but otherwise has proved disappointing. In 1979 the Multiple Mirror Telescope (*right*) broke with tradition by using several mirrors to form a single image, thereby helping to pave the way for Keck's segmented-mirror design.





mirror segments were put in place. A series of test images produced at the end of 1990, including this view of the spiral galaxy NGC 1232 (*right*), demonstrated that the Keck's intricate mirror-alignment system operates as intended.

grinding process, then allowing it to spring back, whereupon it would assume the desired form.

The Keck Telescope's 36 mirrors also must remain precisely aligned as the telescope turns and as the temperature varies. An intricate network of pressure sensors located beneath each mirror segment detects the exact position of that segment and relays the information to a central computer. Twice each second the computer sends commands to an array of motorized actuators that maintain the correct position of the mirror segments.

Many scientists and engineers wondered whether this approach would work. At "first light" in December 1990, with only nine segments in place, the Keck Telescope erased such doubts by making a series of commendably clear test images of nearby galaxies. Those results also impressed the W. M. Keck Foundation, which had funded $70 million of the $94.2-million cost of the telescope. This past April the foundation, which administers the fortune of William M. Keck, the founder of the Superior Oil Company, announced it would grant up to $74.6 million for the construction of a second, identical telescope (now known as Keck II) a short distance away. Keck II is scheduled for completion in 1996.

Despite the quality of its early images, the Keck Telescope's importance for astronomy will rest more on its light-gathering ability than on its fine resolution. Lennox L. Cowie, a cosmologist at the Institute for Astronomy, affectionately describes the instrument as a "light bucket," a giant collector of photons from extremely faint objects. This will make it ideal for spectroscopy, which requires more light than imaging does. Spectroscopy is one of astronomy's most potent tools because it reveals information about the composition, motion and physical state of distant cosmic bodies. Indeed, when the Keck begins its scientific work next year, two of the three detectors attached to it will be spectrometers.

The Keck Telescope's innovative segmented-mirror design does not spell the end of large, single mirrors, however. The Keck drew on lessons learned in the building and operation of the Multiple Mirror Telescope (MMT) at the Steward Observatory in Arizona. The MMT consists of six 1.8-meter mirrors, equivalent in area to a single 4.5-meter light collector. Ironically, just as the Keck is proving the validity of the multiple-mirror concept, Steward is readying plans to replace the MMT's six reflectors with a large one-piece mirror that exploits an equally radical approach.

Frederic H. Chaffee, Jr., the director of the MMT, explains that the proliferation of very large telescopes meant that a 4.5-meter collecting area "wouldn't be terribly competitive by the end of the century." But at the same time, financial constraints would not permit the construction of an entirely new telescope. So Chaffee turned to J. Roger P. Angel, his neighbor at Steward Observatory Mirror Laboratory at the University of Arizona, for a larger mirror.

Solid Alternatives

Angel's mirrors are made using a novel technique called spin casting. In the past, opticians created large, flat disks of glass and then ground them to the proper shape. At Steward, molten glass is poured into a circular mold inside a large, spinning oven. The centrifugal forces maintain a curved surface as the glass cools. Even though the curvature of the resulting blank is not precisely right for a reflecting telescope, spin casting drastically reduces the amount of grinding that must be performed. Innovative, computer-controlled polishing machines at Steward then create the steeply sloped surface shape needed for an extremely fast mirror.

The mirrors made at Steward use a honeycomb construction to reduce enormously the weight of the mirror while preserving its rigidity. In essence, a thin, curved sliver of glass sits atop a lightweight honeycomb glass base. Engineers used similar tricks when casting the five-meter Hale Telescope mirror, but Angel has pushed the technique much further, so that his mirrors are roughly three fourths air.

Although spin casting has received quite a bit of publicity, some scientists and engineers associated with other projects grumble that it has not yet been proved for extremely large mirrors. A persistent rumor has it that the honeycomb casting technique creates a kind of "quilting pattern" on the surface of the mirror that cannot be polished away. Angel simply laughs when asked about such problems. "Things are going very well. We're going from strength to strength," he chimes.

So far Steward has produced honeycomb mirrors up to 3.5 meters wide. In January 1992 Angel plans to cast the first of the laboratory's very large mirrors, a 6.5-meter disk for the MMT. Next the lab will begin work on a series of eight-meter mirrors for some mammoth telescope projects that are still on the drawing board. Depending on how the telescope technology race shapes up, these eight-meter reflectors may be, for a short time, the largest single-piece mirrors in the world.

Otherwise the victory laurels may belong to Schott Glassworks or to Corning Glass. These optics firms are pursuing a third major type of mirror design, the thin meniscus (constant thickness) mirror. Like Angel's mirrors, these will be spin cast. The thin meniscus mirrors, however, will not be supported by a honeycomb. They therefore will be far too frail to maintain their proper shape without some mechanical assistance. Instead they will rely on computer-controlled actuators, much like those that

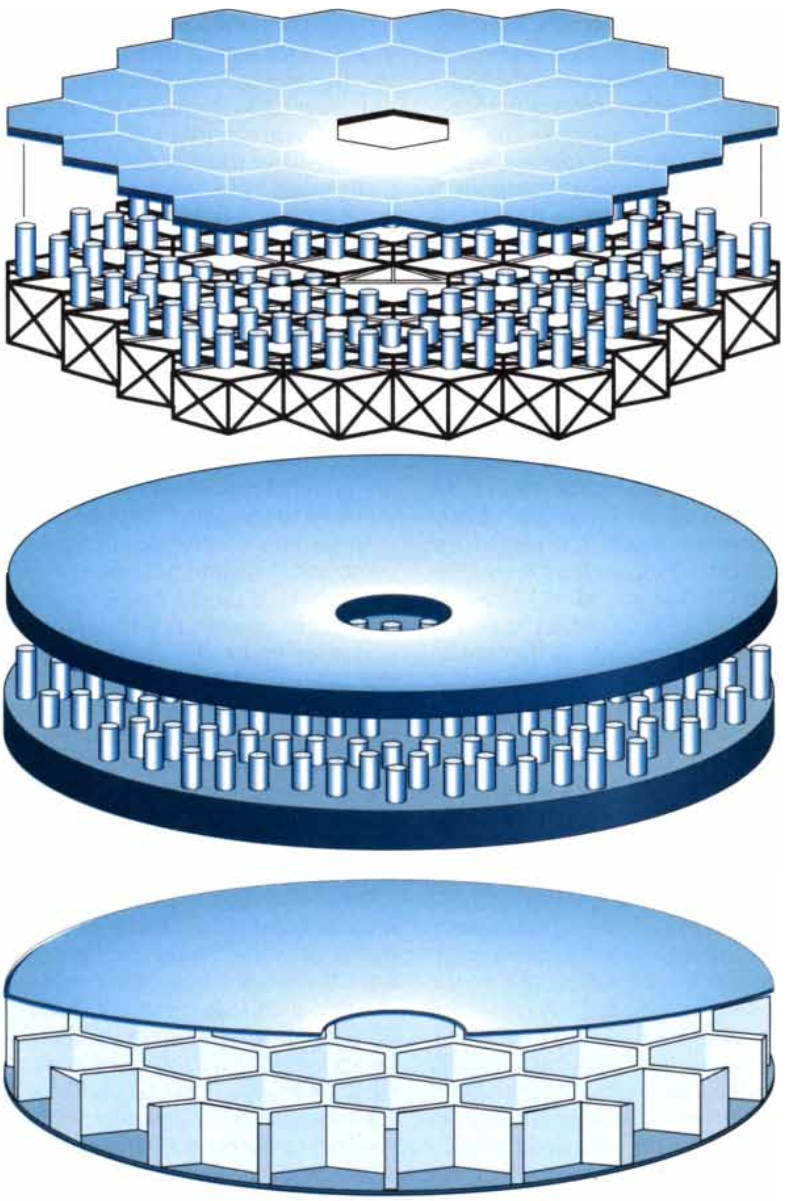
align the Keck mirrors, to flex the glass surface to the correct curvature.

At least two groups are developing giant telescopes that will incorporate thin meniscus mirrors. The European Southern Observatory (ESO) has begun building the Very Large Telescope (VLT), a $225-million array of four 8.2-meter telescopes and eight one-meter ones on Cerro Paranal in Chile. Each 8.2-meter mirror will be a scant 17.5 centimeters thick. Meanwhile the Japanese government has pledged $300 million to construct the 7.5-meter Japanese National Large Telescope on Mauna Kea.

The ESO's New Technology Telescope (NTT), a 3.6-meter device at La Silla Observatory in Chile, has already demonstrated the promise of thin meniscus mirrors. The 24-centimeter-thick NTT mirror is sufficiently flexible that engineers can slightly alter its shape using a network of 75 motorized controllers. When the telescope began operation in 1989, the mirror suffered from a deformation similar to the one still plaguing the *Hubble Space Telescope*. In the case of the NTT, however, engineers could twiddle the controls to bend the mirror to its correct shape. The telescope now produces images of superb quality, sometimes attaining resolution as fine as one third of an arc second. (One arc second is 1/3,600 of a degree; the disk of the full moon extends about 1,800 arc seconds across.)

Like Nelson and Angel, Massimo Tarenghi, the program scientist for the VLT, exudes confidence. "One-piece mirrors are necessary to achieve the highest resolution. That's not a dig at Keck, but a segmented mirror loses resolution," he maintains. Tarenghi raves that the VLT's active control system should also be able to compensate for some of the thermal effects that plague many telescopes. Angel, whose mirrors do not permit such adjustments, understandably takes the opposite point of view. "The goal of astronomers has always been rigid mirrors. Now the engineers

New Ways to Gather Starlight



SEGMENTED MIRRORS incorporate many small, thin mirrors that are easier to build and to support than a single large mirror. Motorized controllers keep the segments aligned so that they form a single image.

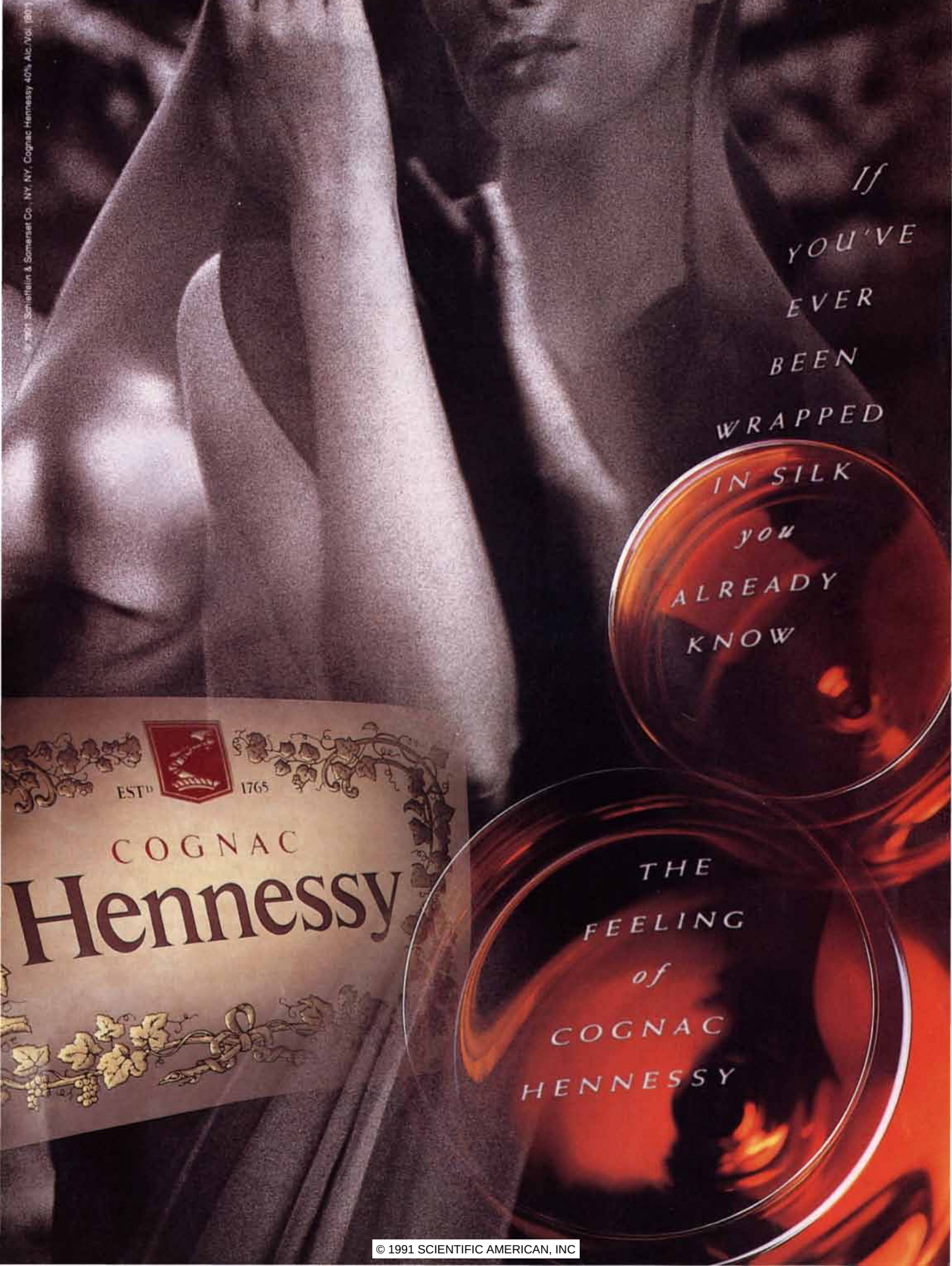
- ***Keck I and Keck II telescopes,*** 10-meter equivalents (California Institute of Technology, University of California and National Aeronautics and Space Administration)
- ***Spectroscopic Survey Telescope,*** 8-meter equivalent (Pennsylvania State University, University of Texas)
- ***German Large Telescope*** (proposed), 12-meter equivalent (German government)

MENISCUS MIRRORS are solid but too thin to support their own weight. Mechanical actuators adjust the mirrors so that they are always bent to the correct shape.

- ***Very Large Telescope,*** four 8.2-meter mirrors (European Southern Observatory)
- ***Japanese National Large Telescope,*** 7.5-meter mirror (Japanese government)

HONEYCOMB MIRRORS consist of a thin reflective surface atop a glass honeycomb structure. These mirrors are stiff but very light because most of their volume is air.

- ***Multiple Mirror Telescope conversion,*** 6.5-meter mirror (University of Arizona)
- ***Columbus Project*** (proposed), two 8.4-meter mirrors (Ohio State University, University of Arizona, Italian government)
- ***Magellan Project*** (proposed), 8-meter mirror (Carnegie Institution, Johns Hopkins University, University of Arizona)
- ***National Optical Astronomy Observatory*** telescopes (proposed), two 8-meter telescopes at separate sites (U.S., U.K., Canadian governments)



HONEYCOMB MIRRORS designed by Roger Angel of the Steward Observatory Mirror Laboratory offer one way of making reflectors that are large and rigid yet relatively lightweight. The mirrors are cast in a spinning oven that imparts a curved surface to the glass while it is molten, thereby simplifying the grinding process.

say that a flexible, distorting mirror is good. The logic just escapes me," he says bemusedly.

Most astronomers are content to sit out the technical debates and wait for the actual results. "With enough effort, I'm sure they'll all work," says Charles H. Townes of the University of California at Berkeley, a doyen of high-resolution astronomy. "It just remains to be seen which one works best." In fact, the most revolutionary development in optical astronomy lies not in the mirrors themselves but in new ways to manipulate the images they collect.

Fighting the Air

All telescopes, no matter how well designed, face a common enemy: the earth's turbulent atmosphere. Even at the best sites, atmospheric distortion prevents astronomers from observing details less than about one half of an arc second across, no matter how large their instrument. A properly designed telescope focuses all the light rays from a distant star to a single point.

Fortunately for poets but unfortunately for astronomers, air has the ability to bend light, causing stars to twinkle. Turbulence in the atmosphere acts like a series of tiny lenses placed in front of the telescope, causing a single star to come into focus in many places simultaneously. As the atmosphere continuously shifts, these points dance around, producing a blurry image. In principle, a flexible reflector incorporated within a telescope could be deliberately bent and deformed in a way that would exactly cancel out atmospheric effects.

That idea, called adaptive optics, dates back at least to 1953, when Horace W. Babcock, now at the Observatories of the Carnegie Institution of Washington, first proposed the scheme for outwitting atmospheric blurring. In the past few years, advances in atmospheric physics, mirror design and computer power have at long last brought adaptive optics to the threshold of reality. The technique holds the prospect of allowing earth-based telescopes to view some objects as if from the crystal clarity of outer space.

In the early 1980s astronomers began working on rudimentary adaptive optics systems to correct for image motion, the simplest kind of atmospheric distortion. René Racine of the University of Montreal has recently developed a relatively simple adaptive optics system along these lines, known as the high-resolution camera, or (as the astronomers prefer) HRCam. The device incorporates a detector that measures how significantly starlight has been bent by the atmosphere, along with a mirror that tips and tilts to compensate for the largest image motions and distortions.

Two years ago Racine installed HRCam on the 3.6-meter Canada-France-Hawaii Telescope on Mauna Kea for a series of optical experiments. The resulting improvements in resolution were so dramatic that researchers at Mauna Kea quickly clamored to use it for regular observations. Other prototype devices—most notably an infrared adaptive optics system developed by the ESO—are now in use elsewhere, but "HRCam is probably the only real adaptive optics system doing routine science," says Robert D. McClure of Dominion Astrophysical Observatory.

McClure has used the Canada-France-Hawaii Telescope to search for a class of stars known as Cepheid variables. Cepheids serve as celestial yardsticks because their period of variation reveals their absolute luminosity and hence their true distance. In this way, McClure and his co-worker John Tonry, along with Racine and Michael Pierce of Kitt Peak National Observatory, hope to determine an accurate distance to the Virgo Cluster, a rich galaxy cluster some 50 to 75 million light-years distant.

Finding the distance to the Virgo Cluster is one of the central problems of current cosmology. Knowing that distance will enable researchers to determine how fast the universe is expanding and thus to calculate the age of the cosmos. Two recent test images taken using HRCam show the first Cepheids even seen in a Virgo Cluster galaxy. McClure hopes to conduct a full-scale investigation next spring.

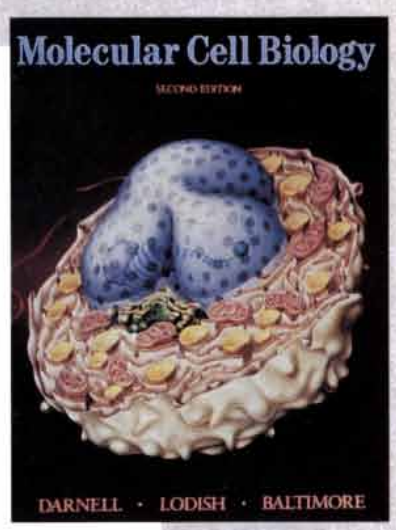
HRCam is also proving a boon for Kormendy in his search for evidence of black holes in the centers of nearby galaxies. He estimates that HRCam improved the resolution of his images from about 0.7 to 0.4 arc second, enough to have made "a very big difference" in being able to detect the presence of an extremely compact mass in galaxies millions of light-years away. Four of the 14 galaxies he examined, including the well-known Andromeda Galaxy (M31), do in fact show evidence of harboring massive black holes.

One shortcoming of many experimental adaptive optics systems is that they need a fairly bright point source to use for calibration. Unfortunately, not every astronomical object sits near a convenient bright reference star. Researchers are therefore hard at work trying to increase the sensitivity of their adaptive optics systems; the number of nearby stars that could serve as guides increases dramatically with increasing faintness.

A group at the Institute for Astronomy led by François Roddier is working on an unconventional system that uses sensors placed above and below the focal point of the telescope's mirror. These sensors record the way the atmosphere has bent and distorted the incoming light. This method has the advantage of needing very little light, and so it should work for fairly faint objects. A computer program translates the data from the sensors into electrical impulses that cause a thin correcting mirror to bend and flex in an exactly compensating manner.

Roddier's relatively uncomplicated system does not attempt to negate all atmospheric distortion, only enough so astronomers can clearly distinguish the true centers of stellar images. It also will probably work only at infrared wave-

KEEPING YOU CURRENT IN TODAY'S CHANGING WORLD



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!

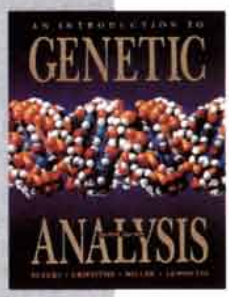
Order your copy today and save 20%!

1105 pages, 1095 illustrations, hardbound.

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	$43.99*	
Biochemistry, 3/e	Stryer	1843-X	$54.95	
Intro to Genetic Analysis, 4/e	Suzuki et al	1956-8	$45.95	

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 $54.95.

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.

4906

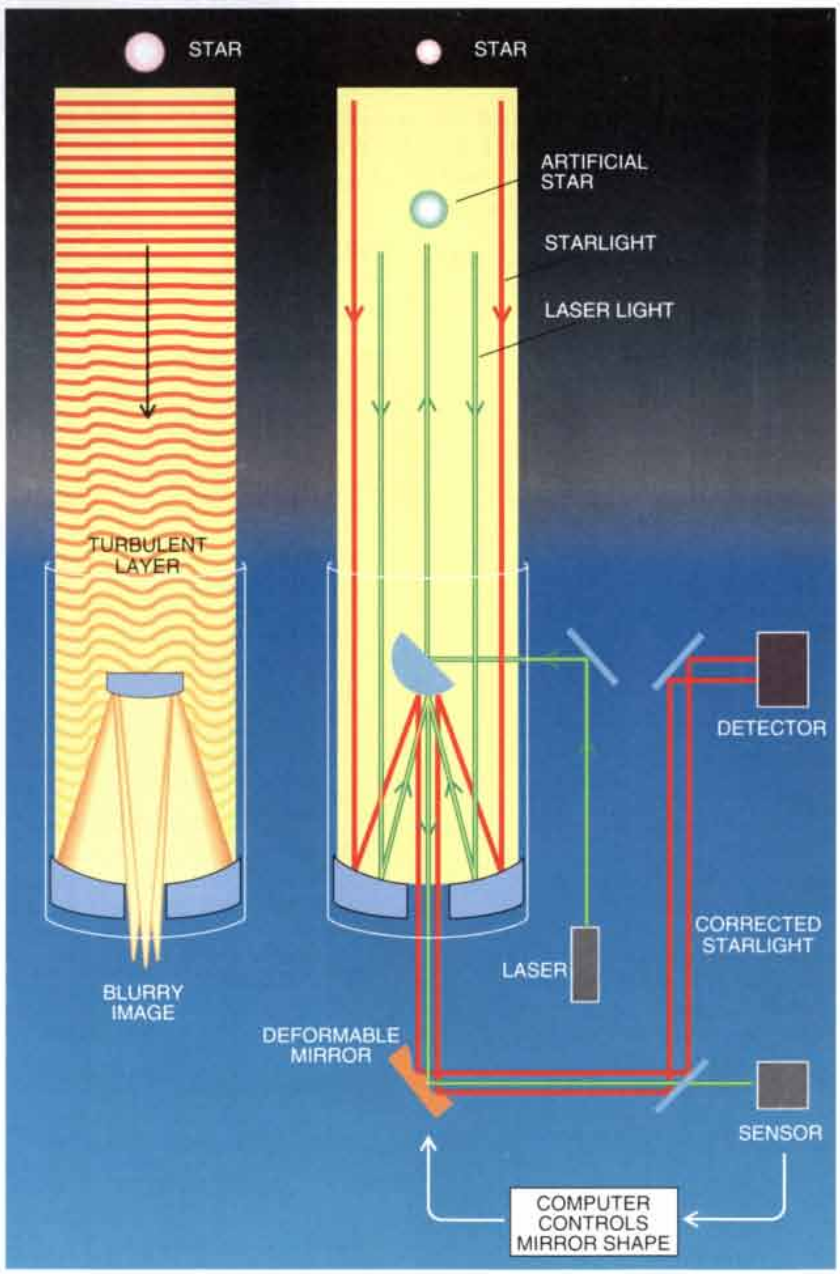
W. H. Freeman and Company
The book publishing arm of Scientific American

How Adaptive Optics Works

In principle, large telescopes can form sharper images than can small telescopes. In practice, however, the earth's atmosphere bends starlight on its way to the ground (*left*), limiting even the largest telescopes to a resolution of about 0.5 arc second, more than 10 times the theoretical limit. Sharper images not only show more detail, they also allow a telescope to detect fainter objects by squeezing the light into a smaller area.

Adaptive optics systems can negate atmospheric blurring, thereby permitting large telescopes to live up to their full potential. Adaptive devices incorporate sensors that measure the atmospheric distortion and a deformable mirror that bends in a way that cancels that distortion. The sensors generally require a lot of light to produce accurate corrections.

The most precise systems will use powerful lasers to project artificial stars onto the sky very near the target. In this way, it will be possible to ensure that there is always a sufficiently bright reference source present. When the deformable mirror corrects the artificial star, it also corrects the image of the nearby astronomical target. A detector then collects the images and separates the real object from the artificial star (*right*).



lengths, where images blur considerably less than they do in visible light. In return, this approach is inexpensive and astonishingly compact. A test setup, cobbled together out of spare lenses, a temporary mirror and a roughly wired-in computer, sits on a table that would fit comfortably (if not beautifully) in a dining room.

The very simplicity of Roddier's design has engendered some skepticism. "Many people told us not to pursue this approach," says J. Elon "Buzz" Graves of the Institute for Astronomy. Others still wonder just how useful such a simple system will be. Laird Thompson of the University of Illinois politely describes Roddier's system as "an evolutionary stage" in adaptive optics, one that will shortly be replaced by a more advanced technique that uses a laser-generated guide star.

Using lasers to create an artificial star in order to calibrate an adaptive optics system avoids the otherwise inevitable trade-off between sensitivity and resolution. "To clean up scientifically, astronomers need a bright star close by the object they want to study," explains Thompson, who has spent several years working on a civilian laser-guided system. A bright reference star provides enough light to perform extremely thorough image corrections. This past May the concept was validated when the U.S. Department of Defense declassified research from its decade-long effort to develop laser-guided adaptive optics systems for the Strategic Defense Initiative (SDI). For SDI, the goal was to destroy an incoming missile using a high-powered laser. The problem of preventing the laser beam from dispersing before it reaches its target involves many of the same issues as does correcting for the way starlight twists and turns on its way to the ground. In the SDI approach, a laser beam bounces off a layer of sodium atoms about 90 kilometers up in atmosphere, creating a pointlike reference beacon.

The SDI laser-guided system can produce far more sophisticated corrections than a simple device like HRCam, but at a price. "Too expensive for astronomy," says Graves, citing an oft-quoted cost of about $15 million. Thompson estimates that a version customized for visible-light astronomy could be built for "a few million dollars."

But most work in adaptive optics is now being done at infrared wavelengths, where atmospheric distortion is much less pronounced. Edward Kibblewhite of the University of Chicago, who champions the SDI-derived work, thinks a commercial system that operates in the infrared can be made for "a few hundred

thousand dollars"—a bargain compared with the cost of a new telescope.

Some astronomers believe the time for laser-guided adaptive optics still lies some distance in the future. Racine, for example, worries that "as it stands now, the SDI system is not suitable for astronomy," because light is scattered from the artificial star. Members of Roddier's group grumble that military adaptive optics is unnecessarily complicated for astronomical work. Yet early tests run by Robert Q. Fugate at the U.S. Air Force's Starfire Optical Range look promising. And even Racine is quick to note that "when it is available and optimized, laser-guided adaptive optics will bring about a real revolution."

Despite the tremendous promise, adaptive optics has limitations. The intricate optical systems used to correct the image tend to lose some of the light captured by the telescope. Moreover, mating adaptive optics to the biggest telescopes will not be easy. Large mirrors observe a wider swath of turbulent air and therefore require more complicated corrections to unscramble the starlight. The very fast, short-focal-length mirrors now in vogue complicate the job of sensing image distortions.

Another restriction is the small field of view over which the corrections can be made. Here an unfortunate compromise emerges: the better the resolution, the smaller the area over which adaptive optics works. The relatively simple HRCam can produce a corrected image that spans about two arc minutes (120 arc seconds). Roddier's system, which also does not strive for the very highest resolution, should be able to yield corrections over an area with a radius of 10 to 20 arc seconds, according to Daniel McKenna of the Institute of Astronomy. But the very high resolution SDI-derived systems will be able to produce a clear, visible-light image only four or five arc seconds wide.

The Ultimate View

The differences are significant because many particularly interesting objects (galaxies, structures around quasars and gas clouds where stars are being born, to name but a few) extend several seconds to many minutes across. McClure, for example, remarks that his studies of the Virgo Cluster would not be possible using a tightly focused laser-guided system. In all likelihood, several different kinds of adaptive optics systems will be used depending on the site, telescope, type of observation desired and amount of money available.

If adaptive optics seems like a dream come true, then another technique,

Earthbound Telescopes Take on *Hubble*

Before its much delayed launch in the spring of 1990, the *Hubble Space Telescope* was regarded as the instrument that would revolutionize astronomy. From its vantage point above the earth's distorting atmosphere, *Hubble* was expected to help astronomers solve some of the most tantalizing questions about the origin of galaxies and the age of the universe.

Hubble did not work quite as planned, however. In addition to its well-known focusing problem, the telescope has been plagued by a series of ailments. Thermal expansion of solar panels has produced a case of the jitters. The failure of two gyroscopes threatens to disrupt the telescope's pointing system. Most recently, a power supply to the Goddard High Resolution Spectrograph, the instrument least compromised by *Hubble*'s faulty optics, began acting up. And now the burst of progress in ground-based telescopes is taking even more of the luster off *Hubble*'s mission.

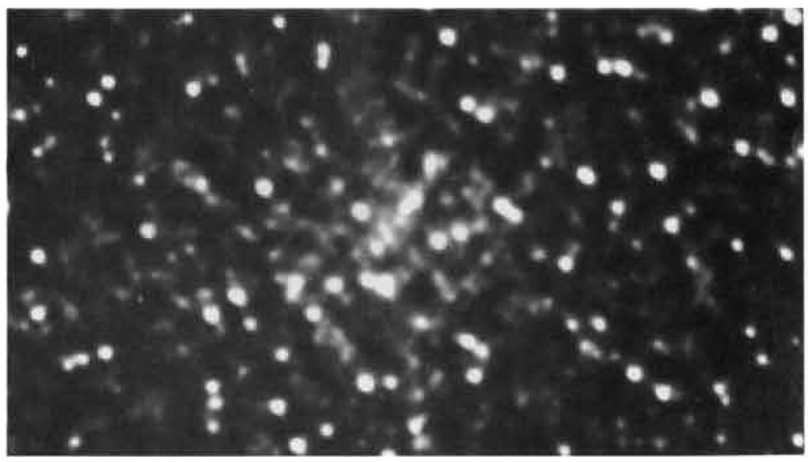
Hubble's single greatest virtue was supposed to be the clarity of its images. When it is working properly, *Hubble* should be able to resolve details as small as 0.06 arc second, some 10 times better than was possible from the earth's surface just a few years ago.

But recent developments in adaptive optics, which can correct for the blurring caused by the atmosphere, should soon make it possible to achieve resolution of about one tenth of an arc second from the ground. Moreover, *Hubble*'s 2.4-meter mirror is dwarfed by the new telescopic giants; the Keck, for example, will collect 17 times as much light as does *Hubble*.

Still, astronomers are loath to write off the space telescope. Shrinivas R. Kulkarni of the California Institute of Technology points out that terrestrial telescopes using simple adaptive optics systems will experience a problem similar to the one plaguing *Hubble*. Some uncorrected light mixes with the image, smearing out some of the fainter details. *Hubble* will receive a set of corrective lenses in 1994 to fix its vision, but ground-based astronomers will have to wait for improvements in adaptive optics technology. *Hubble* also has another advantage: it will produce sharp images across its entire field of view, not just in one corrected section.

Finally, some unfortunate aspects of the atmosphere cannot be overcome by either adaptive optics or clever telescope design. Although pleasantly transparent to visible light, the earth's air strongly absorbs most ultraviolet rays; it also emits and absorbs certain infrared wavelengths. No other instrument can rival *Hubble*'s views at these wavelengths.

On the other hand, *Hubble* cost roughly $1.5 billion, compared with about $100 million for Keck I and a few million dollars to develop and build an advanced, laser-guided adaptive optics system. Given the competition for federal astronomy funding, it is no wonder that many researchers quietly agree with Laird Thompson of the University of Illinois when he muses that "with hindsight, astronomers would not invest the money in *Hubble*."



***HUBBLE* RESOLVED individual stars in the core of the globular cluster M15. But ground-based telescopes are catching up in clarity.**

NEW TECHNOLOGY TELESCOPE in Chile (*left*) incorporates superb optics and a flexible primary mirror whose shape is adjustable; the instrument sits in an octagonal "dome" specially designed to reduce atmospheric distortion. In its first year of operation, the telescope has produced some exceptionally fine images, such as this one of supernova 1987A (*right*).

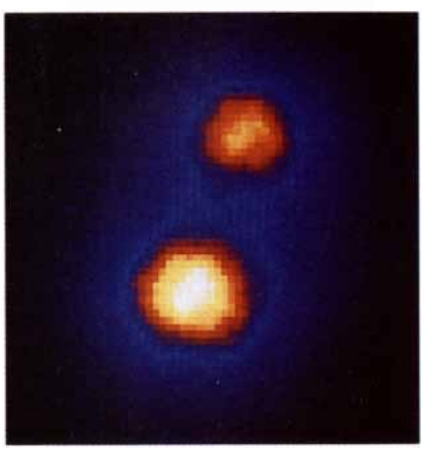
called optical interferometry, must be pie in the sky. Adaptive optics should soon yield images having a resolution of about one tenth of an arc second; experimental interferometry systems can already resolve stars only a few thousandths of an arc second apart. Interferometry, which has been exploited by radio astronomers for years, involves combining images from two or more widely spaced telescopes. In this way, it is possible to simulate the resolving power that would result from a single telescope as large as the distance between the linked instruments [see "The Very-Long-Baseline Array," by Kenneth I. Kellermann and A. Richard Thompson; SCIENTIFIC AMERICAN, January 1988].

Interferometry requires that signals from the telescopes be combined in phase, that is, that the waves of the incoming radiation must all be carefully aligned. Because of the extremely short wavelength of light, optical interferometry requires that the positions of the various telescopes and the shape of the wave fronts must remain highly stable. (Green light, for example, has a wavelength of about 0.5 micron, or millionths of a meter; radio waves, in comparison, are millimeters to meters in length.) As with adaptive optics, optical interferometry becomes particularly challenging for large telescopes because they look through wide, variable patches of the atmosphere, making it exceedingly difficult to keep light striking one part of the mirror in phase with that falling on another area.

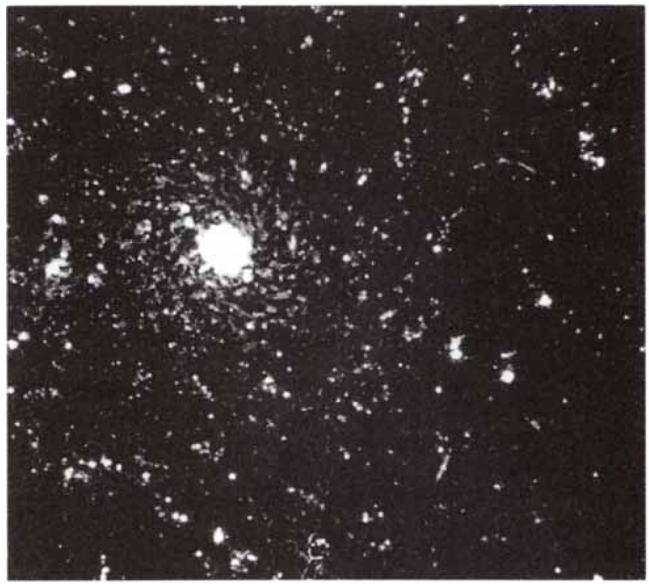
Despite these difficulties, groups at Mount Wilson Observatory in California, the University of Wyoming at Laramie, the University of Cambridge and a handful of other locations are hard at work developing prototype optical interferometry systems. For the time being, most of these devices use small telescopes; the Cambridge design, for example, will consist of four 50-centimeter telescopes spaced 100 meters apart. A related approach, being pursued by Gerry Neugebauer of the Hale Telescope, involves placing a mask over the telescope's large mirror. Holes in the mask simulate several small, closely spaced individual telescopes.

Small telescopes can gather only modest amounts of light, so interferometry will initially be limited to relatively bright objects. On the positive side, many enticing, fine details that astronomers would love to look at—the surface features of nearby stars, for example—occur on relatively bright objects. Neugebauer's group and a team from Cambridge working on the William Herschel Telescope in the Canary Islands have succeeded in producing crude maps of the surfaces of red giant stars.

Furthermore, large-telescope optical interferometry may not be far off. When the Keck Foundation announced the planned construction of Keck II 85 meters away from the Keck I, one of its stated goals was optical interferometry. Nelson envisions starting with a simple adaptive optics system to make the Keck's images sufficiently stable to permit interferometry. Eventually he hopes



LASER-GUIDED ADAPTIVE OPTICS demonstrates its capabilities in these test images of a star in Ursa Major. Both views are from the same 1.5-meter telescope. The uncorrected image shows just a faint blur (*left*); the adaptive optics system yields a drastically improved image that clearly reveals the star to be double (*right*).



CANADA-FRANCE-HAWAII TELESCOPE on Mauna Kea (*left*) has been the test-bed for a prototype adaptive optics system. The telescope can now capture images of individual variable stars in galaxies in the Virgo Cluster (*right*), a project intended for the ill-fated space telescope. Studying these stars will help establish the true size and age of the universe.

to switch to a more elaborate, laser-guided device and to add a set of four auxiliary telescopes.

The ESO's Very Large Telescope program includes an even more ambitious interferometry program. The group of eight one-meter telescopes and the four 8.2-meter behemoths will ultimately be combined to form a 190-meter-wide interferometric array. Jacques M. Beckers, who directs the interferometry program for the VLT, predicts wondrous results from this system: resolution of two thousandths of an arc second over a field of view of eight arc seconds (an astonishingly broad area for interferometry). "Astronomers will be able to image the star Antares and take a good look at the disk around Beta Pictoris," he says, sounding somewhat giddy. "We're seeing totally new optical astronomy ahead of us, with resolution and light gathering far beyond the *Hubble*."

Beckers envisions that the VLT will produce its first interferometric results in 1997, a rather ambitious goal given that the first of the 8.2-meter instruments in the array will not be completed until 1996. Nelson projects a similar timetable for the Keck I–Keck II system. Nevertheless, most astronomers would agree with Stockton when he says, "I've got a feeling that optical interferometry won't really be available for 10 to 15 years." Peter J. Warner, part of the Cambridge interferometry group, questions the wisdom of attempting interferometry on a large telescope this early in the game. "It will be fantastic if it works," he agrees, "but I would not want to be involved, because it will take a very long time."

One way to reduce the difficulty of both adaptive optics and interferometry is to use longer wavelengths of light. As Angel puts it, "infrared doesn't notice the little wiggles in the atmosphere" as much as visible light does.

Beyond Light

Seeing tends to be better in the infrared, and so adaptive optics systems are much easier to design for these longer wavelengths. "There's no doubt that adaptive optics systems can be put on the large, eight- to 10-meter, telescopes at infrared wavelengths," Thompson claims, "but it is not clear how long it will take, or whether anyone will try, to do it in visible light." Infrared images taken at some medium telescopes, such as the MMT, are already approaching the instruments' theoretical resolution limit, even without adaptive optics.

Infrared observations are also of great interest because they offer a window on subtle, relatively cool astronomical objects, such as planets, brown dwarfs and newborn, dust-shrouded stars. Radiation from distant galaxies and quasars gets tremendously stretched and weakened by the expansion of the universe. As a result, emissions that began as visible light reach the earth as longer-wavelength infrared rays. "Near infrared is the obvious place to do cosmology," comments Cowie, who recently found a mysterious class of faint galaxies that shine most prominently in those wavelengths.

More and more, optical astronomy is being done at infrared wavelengths. Optical telescopes can observe into the near infrared range with little modification, and the new-generation devices are being built with solid state infrared observations in mind. High-resolution infrared detectors, known as infrared arrays, are making it possible to produce more detailed images in infrared wavelengths. "People here consider wavelengths up to five microns as 'extended optical,'" reports Gerry Luppino of the Institute for Astronomy.

Such changes in thinking represent the fruits of a centuries-long effort to bring the entire cosmos within the range of human perception. Giant telescopes and adaptive optics are merely the latest weapons in this audacious struggle. And some highly promising techniques, such as optical interferometry, have yet to show their full potential. "I don't know what the limits are," Beckers says. "There is no end in sight."

FURTHER READING

ADAPTIVE OPTICS REVISITED. Horace W. Babcock in *Science,* Vol. 249, pages 253–257; July 20, 1990.

ADVANCED TECHNOLOGY OPTICAL TELESCOPES IV. Edited by L. D. Barr. Bellingham, Wash., SPIE, 1990.

THE NEW GROUND-BASED OPTICAL TELESCOPES. Buddy Martin, John M. Hill and Roger Angel in *Physics Today,* Vol. 44, No. 3, pages 22–30; March 1991.

WHEN BIG IS BEAUTIFUL. S. R. Kulkarni, M. Shao and C. A. Haniff in *Nature,* Vol. 352, No. 6334, pages 383–384; August 1, 1991.

THE DECADE OF DISCOVERY IN ASTRONOMY AND ASTROPHYSICS. National Research Council. National Academy Press, 1991.

Brain, Food

Drugs based on neuropeptides may soon treat eating disorders

The excess fat on some 40 million Americans is no longer dismissed as physical evidence of someone who simply cannot say no to food. Like ulcers and narcotics addiction before it, obesity is now being examined as a treatable biochemical imbalance. More than appearance is at stake. The condition—body weight of 20 percent or more above average—is a significant contributor to major illnesses, including cancer, diabetes, heart disease and stroke.

An effective treatment for obesity would improve health and life expectancy, reduce health care costs and likely earn its developer a fortune. To those ends, major drugmakers, including Abbott, Eli Lilly, Pfizer and Merck, along with biotechnology firms such as Interneuron Pharmaceuticals and academic institutions around the world, are craving insight into what is becoming known as the nonsupply side of fat.

These scientists are looking for ways to control appetite selectively without causing as many side effects as the stimulants now on the market. They are investigating long-acting proteins known as neuropeptides and short-lived neurotransmitters that are found in regions of the brain that affect mood and appetite as well as in the digestive tract. Some of the compounds stimulate appetite for particular nutrients such as fat and carbohydrate; others suppress those urges to eat.

The relationship of these compounds to physiology and behavior is compelling: studies show that obese people as well as anorexics and bulimics have disturbed levels of the neurochemical substances. As a result, drugmakers hope to treat obesity and other eating and body weight disorders with synthetic versions of these natural compounds or with drugs that either stimulate or block their function. "Most endocrinology groups are working on neuropeptides," declares Ferid Murad, vice president of pharmaceutical research and development at Abbott Laboratories in Abbott Park, Ill.

Murad admits that large companies like his have deliberately remained quiet about the potential of these molecules for treating eating disorders because "it is still early, we don't know much about how they work yet." But within a year, Abbott intends to begin human trials of an appetite suppressant modeled after a neuropeptide called CCK, for cholecystokinin.

"CCK was found in the brain and then in the gut," Murad recalls. "We said, 'It's a hormone. It's there. It must be interesting.' " It was: initial tests in animals convinced the researchers that the peptide could suppress appetite. But, like other neuropeptides, the CCK molecule is too big and too highly charged to be administered orally. Eventually researchers hope to synthesize smaller molecules that will not have to be injected. Merck and others have the same wishes for second-generation versions of appetite-stimulating CCK antagonists to treat anorexia.

NEUROBIOLOGIST SARAH LEIBOWITZ studies how certain brain chemicals affect the eating preferences of animals like this genetically obese rat. Photo: Robert Prochnow.

Some of the Brain Chemicals That Control Eating

INHIBITORS

- ***Cholecystokinin:*** Present in the brain as well as in the gut and oral cavity where taste is perceived; suppresses appetite for fat and carbohydrate.
- ***Cyclo [His-Pro]:*** A particularly stable molecule that fluctuates in response to food intake or deprivation and seems to regulate satiety.
- ***Corticotropin-Releasing Hormone:*** Controls release of steroid hormones and by itself suppresses appetite.
- ***Dopamine:*** This basic chemical messenger in the brain is frequently targeted by appetite-controlling drugs; suppresses appetite for fat and protein.
- ***Serotonin:*** Seems to tell the body when sufficient carbohydrate has been consumed; bulimics show a deficiency in this substance.

STIMULANTS

- ***Neuropeptide Y:*** One in a family of neuropeptides that includes pancreatic polypeptide; encourages appetite for carbohydrate and speeds its use by the body.
- ***Galanin:*** Its presence at the end of the day or active period stimulates a preferential appetite for fat.
- ***Norepinephrine:*** Related to adrenaline; increases appetite for carbohydrate during the first meal of the day.

Life is unpredictable. Thankfully, the Accord is not.

Will the stock market make a meteoric rise? Will bell-bottom pants make a comeback? Does the guy driving next to you know what he's doing?

Maybe.

At Honda, we concentrate on the third question. Which is why the Accord is a safe car to drive.

Settle in behind the wheel of the Accord and the first safety feature will be right in front of you. The driver's side airbag Supplemental Restraint System (SRS). It's standard on all 1992 Accords.

Once inside, take comfort in knowing that the Accord's body, front and rear, is specifically designed to shoulder the brunt of a collision. This helps keep the passenger compartment out of harm's way.

If life tosses you an unexpected turn, underfoot are anti-lock brakes, one of the safest, most effective ways to stop a car. In panic braking situations, even in adverse conditions, your wheels don't lock up. So you stay in control.

We could go on. The Accord's safety features are numerous. Rest assured, all are designed to give you peace of mind as you travel through life.

Which means that whatever tomorrow may bring, you're prepared in an Accord. A soothing fact that we know you'll be thankful for. HONDA

Another promising eating-associated neuropeptide has a name that is a mouthful: histidyl-proline diketopiperazine, otherwise known as cyclo[His-Pro]. Levels of the substance in the blood increase after food is consumed; no change occurs when water is ingested. "It's a very comfortable structure, not susceptible to enzyme attack," notes John F. Wilber, a professor of endocrinology at the University of Maryland School of Medicine. That stability allows the peptide to inhibit appetite for up to 12 hours in rodents, a charming attribute in the eyes of drugmakers.

Satiety can also be induced in humans by increased levels of the neurotransmitter serotonin, says Richard J. Wurtman, director of the Clinical Research Center at the Massachusetts Institute of Technology and co-founder of Interneuron Pharmaceuticals, a small neurological therapeutics company based in Lexington, Mass. This chemical messenger is particularly effective at suppressing appetite for carbohydrates, making it a natural target for drug therapy, he asserts. "A large number of very large people eat very large amounts of carbohydrate," Wurtman notes, possibly in an attempt to improve their moods by stimulating production of serotonin.

That theory guided his design of a drug called dexfenfluramine, which increases the brain's production of serotonin and also maintains high levels of it in the synapses, or gaps, between nerve cells. Interneuron plans to seek approval for the drug from the U.S. Food and Drug Adminstration within the next year or two.

Neuropeptides seem to have an even more profound and longer-lasting effect on body weight and food intake than do neurotransmitters, says Sarah F. Leibowitz, a neurobiologist at the Rockefeller University. Experiments in rats show that neuropeptide Y (NPY) begins stimulating appetite for carbohydrates within minutes after it appears in the brain, but this effect lasts up to 24 hours.

NPY kicks in early in the feeding cycle, perhaps to ensure that carbohydrate is available to provide quick energy during the early waking hours, Leibowitz postulates. At the end of the eating period, she points out, the taste for fat develops as activity of another neuropeptide, called galanin, increases. GAL stimulates fat intake.

Human preference for specific foods also varies according to the time of day. These patterns may be controlled by the interplay and rhythms of neuropeptides, Leibowitz suggests. "Generally in the morning we don't want fat, but our appetite for it will rise as we go into evening," she notes. The circadian aspect of neuropeptide function could have important implications for treating patients—why give a drug in the morning to suppress an appetite that will not surface until late afternoon?

Food preferences also shift as an individual matures. Particularly for females, the appetite for carbohydrate surges at puberty, when NPY content in the hypothalamus is highest. Appetite for fat increases in both sexes shortly after puberty, when GAL becomes activated.

Just because people with eating disorders have disturbed levels of particular neuropeptides does not necessarily mean these compounds cause the problems, Leibowitz declares. But the substances could help maintain unhealthy patterns. Administering minute amounts of NPY to test animals encourages overeating and therefore obesity.

In normal rats, levels of the compounds change in response to the availability of food: NPY appears after periods of food deprivation; levels decrease after feeding. The fluctuations in appetite and the different neurochemicals that affect it prompt Leibowitz to assert that "there won't be a single neuropeptide treatment for all people at all times. I think we are going to need a variety, multiple solutions to the problems of appetite and weight control."

Cocktail combinations of neuropeptides or drugs that act on them will still be only supportive therapy for the behavioral modification needed to cure an eating disorder. "Eating disorders must be approached from two directions—both brain and behavior are important," says Wylie W. Vale, president-elect of the National Endocrine Society and professor of peptide biology at the Salk Institute in La Jolla, Calif.

Even so, Vale adds, "the most far-reaching progress is being made at the molecular and biochemical level." Researchers have already found some of the genes that control production of eating-associated neuropeptides. In Leibowitz's laboratory the activity of these genes has recently been shown to be disturbed in animals that overeat and become obese.

"It's not like it's 100 years away," Leibowitz declares. "We're starting to get markers and predictors of tendencies toward obesity, anorexia, bulimia. It's exactly what we're trying to do with all disease—to predict tendencies to move in a certain direction, say, 'Okay, you've got a weakness'—and then help steer people into preventive programs." Hunger for this type of knowledge is not likely to be satiated any time soon. —*Deborah Erickson*

Sustainable Jobs

Green work will be different, but there might be more of it

The Worldwatch Institute has published a study that may turn on entrepreneurial light bulbs in certain corporate suites and raise blood pressure in others. According to Michael G. Renner, the author of the study, companies that threaten to close plants or relocate if they are forced to comply with environmental regulations are often guilty of job blackmail. Companies point fingers at green legislation, Renner says in "Jobs in a Sustainable Economy," but in most cases, laws to protect water supplies or air quality do not deserve the blame for reduced operations.

Most of the time, Renner contends, jobs are lost because automation has cut the amount of human labor many tasks require or simply because companies are running out of resources. "We need to step back from the controversy, short of taking any environmental measure, and ask, 'If we didn't do anything, what would happen?'"

Certain jobs would be lost if the world moved away from fossil fuels, excess packaging and goods designed to be rapidly obsolete, Renner acknowledges. The motor vehicle industry alone employs more than four million workers and contributes almost 7 percent of global economic output—but the dollars are deceiving, he says. Although sales revenues for West German auto companies more than doubled from 1977 to 1987, the number of jobs grew by less than one quarter.

"In the next generation, the creation of millions of jobs in pollution control would seem inevitable," Renner writes. He reports that analysis of the $100 billion spent on pollution control in the U.S. in 1988 showed that it supported or newly created almost three million direct or indirect jobs. By the end of this decade, America's pollution-control outlays are expected to exceed $171 billion a year. The number of jobs will increase, too, but not always where money is spent.

Even so, employment gains are likely to outweigh losses from environmental policies, Renner asserts. Renewable energy sources will be a key source of jobs because their capital requirements, with the exception of solar energy, are low and their labor components are greater than those of conventional fuels. Renner cites a recent Worldwatch analysis that figures 100 workers would be needed to generate

1,000 gigawatt-hours of electricity in a nuclear plant, versus 116 in a coal-fired plant, 248 in a solar-thermal facility or 542 on a wind farm. German studies drew similar results, he notes.

Some of the new jobs created by a sustainable economy will utilize skills already possessed by workers. For instance, the tasks involved in building automobiles and highways—such as engineering, trucking and concrete pouring—are similar to what would be required to build railroad systems or urban bicycle paths.

Recycling resources saves money and creates more jobs than other waste disposal technologies, such as landfills and incineration. A study conducted by the Center for the Biology of Natural Systems at Queens College concluded that reaching New York City's mandated rate of recycling 25 percent of its waste by 1994 will create 1,400 jobs. Planting trees is labor intensive, as is the installation of weatherproofing necessary to conserve energy.

Governments might consider giving more active assistance to small firms whose innovative pursuits were hindered by lack of capital rather than corporate tradition, Renner suggests. Large companies could be made to realize that prevention is indeed often cheaper than the cure for toxic wastes and pollutants. "It's probably very important to set up the right framework within which these companies can operate," Renner says.

Then again, one can always get tough. Renner advocates leveling the playing field by eliminating subsidies and programs now in place that work against sustainability, such as building highways instead of public transportation. It might be useful to require companies to make available information about how their operations affect human health or the environment, Renner argues. A step in that direction was taken by the European Economic Community, which has agreed on a mildly worded statement that corporations should not be allowed to hide their activities behind claims of proprietary technology.

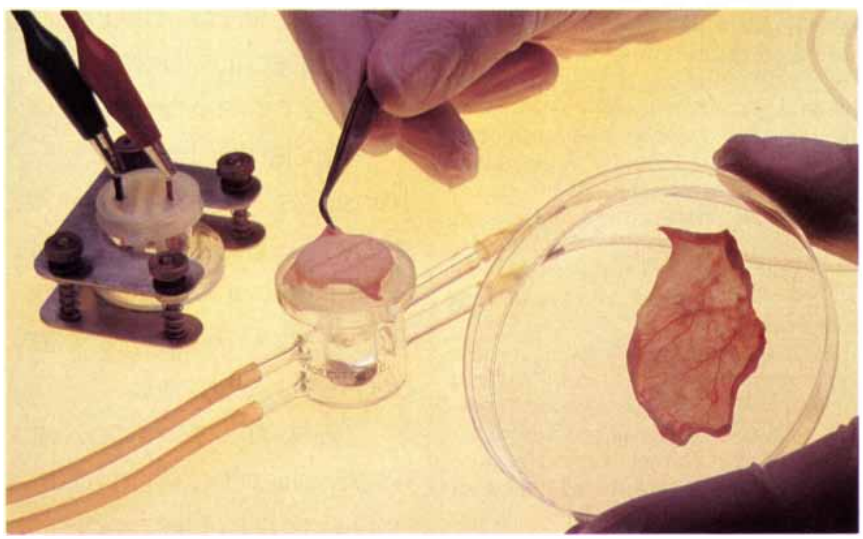
Environmental consciousness pays dividends in international competitiveness, Renner declares. He cites the case of Sweden and Germany, where stringent legislation has compelled companies to develop solutions that achieve government-ordered goals. Both have become leading developers of environmental technology and are exporting it to the rest of the world. It's not easy being green, but enlightened self-interest is clearly best served by moving ahead. *—Deborah Erickson*

Skinside Out

Drug delivery in reverse could be a diabetic's boon

Richard H. Guy, a professor of pharmaceutical chemistry at the University of California at San Francisco, has been seeking ways to deliver drugs through the skin. One technique he has studied is iontophoresis, the use of tiny electric currents to push ions across membranes. Guy soon realized that while certain compounds were going through the skin as planned, others were coming out. The observation gave him an idea: maybe transdermal drug delivery could work in reverse as a diagnostic method.

So Guy quickly filed a patent on the technique for extracting biological molecules, as well as illicit drugs. He found a licensee to develop the technology, Cygnus Therapeutic Systems. The drug delivery company in Redwood City, Calif., hopes iontophoresis will one day replace the self-inflicted



PHARMACEUTICAL CHEMIST RICHARD GUY **(at left)** ***is developing a noninvasive glucose monitor with Russell Potts of Cygnus Therapeutic Systems'. The experimental device shown here draws metabolites through the skin. Photo: Matthew Mulbry.***

Walking

Hearing your customers and having the ability to respond efficiently and thoughtfully to their needs is key to the present business climate. The companies of Bell Atlantic know this, so they've created some tools to help you do just that. Features

the line between

like 'voice mail,' which allows clients to leave detailed messages 24 hours a day. And features that can forward your calls both in and out of the office to follow you wherever you are. Improve your bottom line and customer communication. Let Bell Atlantic

success

enhance your current phone system or create a new one for you. To get the phone system that's right for you, call your representative at Bell of Pennsylvania, Diamond State Telephone, New Jersey Bell or The C&P Telephone companies.

and failure.

It doesn't cost to call, but it might cost not to. 1-800-843-2255.

Bell Atlantic
We're More Than Just Talk.®

pinprick that enables millions of diabetics to detect the critical levels of glucose in their blood.

The Cygnus researchers are now perfecting their methods on a tiny device called a diffusion cell, which mimics the skin's relation to the interior of the body. A solution with a known amount of glucose and other metabolites is pumped through a chamber with a volume of approximately a cubic inch. The skin clamped over this receptacle is real, left over from face-lifts and tummy tucks or taken from hairless mice. Self-adhesive gel attaches electrodes to the skin surface; glucose accumulates in this barrier material when electric current is passed between the electrodes.

Even though a tiny electric current speeds the transfer of ions across the skin, concentrating measurable levels of glucose is difficult. Because glucose does not carry a charge, it must be swept along by highly mobile ions, such as sodium and chloride. Other ions also compete for the charge. The Cygnus researchers hope to improve the odds by using a gel that excludes the most prevalent ions.

When tested on mouse skin, the amount of glucose that comes out "seems nicely linear" with the circulating level, says Russell O. Potts, the firm's associate director for biophysical research. "The Achilles' heel in commercializing the monitor may be the amount of time it takes to collect material," he admits. "We can do it in 30 to 60 minutes, but we're shooting for a 10-fold reduction in time," he notes. Even then it will be hard to best the almost instantaneous results from the conventional stick tests.

To get a faster reading without bloodshed, the researchers hope they can also extract breakdown products of glucose. These smaller metabolites have a negative charge that may help them move through the skin and into a monitor more rapidly. Calculating the ratio of these fragments to glucose may help compensate for the method's inherent slowness, Guy suggests, by providing extra numbers to read sooner.

"The logic is good," comments Ronald R. Burnette, an associate professor at the University of Wisconsin School of Pharmacy. But he cautions that using the metabolites to estimate glucose levels will be complicated because all sugars and carbohydrates have the same building blocks. "Whether they are able to separate out these pieces or not, the amounts they collect will still be very small. Detecting extremely small amounts of things can be extremely expensive." Another problem is limiting the amount of skin that must be covered by a sensor to capture glucose effectively. "Ideally, we'd like something like a dot on the end of an eraser, but we can certainly go for a patch or a cuff," Potts offers.

The potential value of a noninvasive glucose monitor has not escaped the big players in the diabetic market. Eli Lilly, one of the world's leading suppliers of insulin, and Becton Dickinson, a major supplier of syringes, have both nibbled on the concept. Alza Corporation in Palo Alto, Calif., which pioneered transdermal skin patches, has explored the approach as well. "We dabble in the area of asking how do we sense things and control our system in response to that," confesses J. Bradley Phipps, Alza's manager of materials research for electrotransport.

If transdermal sensing can be made to work, the benefits for diabetics could be more far-reaching than eliminating an uncomfortable test. "If you could pull out and sense glucose and deliver insulin, you'd have a nice closed loop," Phipps says. If transdermal monitoring does become used extensively, prices for the technology could drop enough to encourage truly novel diagnostic applications, he observes. "Suddenly they might notice, 'Hey, this group of people with the flu or a certain virus took twice as long to recover, and they're all low in potassium.'" For now, though, Cygnus is sweet on a simple sugar. —*Deborah Erickson*

Plastic for the '90s

Some clever chemistry churns out a new polymer

In the 1960s and 1970s new plastics seemed to reach the market almost every week. They came with family names like polyethylene, polyester, acetal and acrylic, earned public recognition under trademarks such as Plexiglas and Kevlar and endeared themselves to engineers as stronger, lighter alternatives to metals and other traditional materials. Molding, extruding and stamping transformed these resins into everything from leisure suits to popcorn poppers, car parts and computer housings.

Then the procession of novel polymers stopped. Alloys, blends and composites prevailed. Fine-tuning of properties became the game. But Amoco Corporation in Naperville, Ill., has demonstrated that there is still some room for innovation in polymer chemistry. Earlier this year it introduced the first totally new resin in nearly a decade.

Made by combining aliphatic diamines and terephthalic acid, the plastic, known technically as polyphthalamide, is tradenamed Amodel.

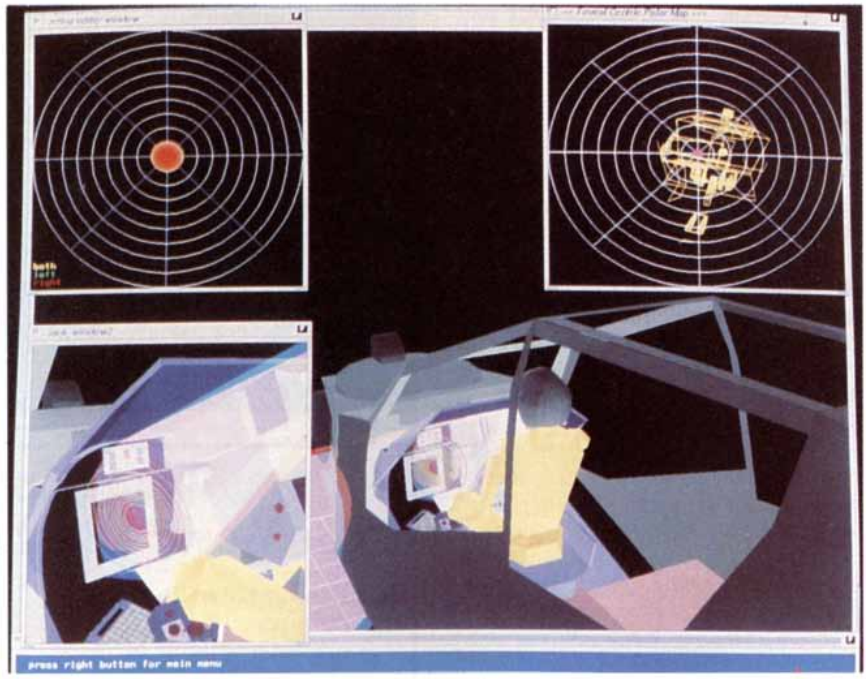
The semicrystalline resin is "looking good and capable," says Edward T. Shea, Amoco's business development manager for performance products. The new thermoplastic has a balance of properties that promises to make it welcome in everything from irons and electric outboard motors to oil-drilling apparatus and packages for integrated circuits and light-emitting diodes. Several hundred evaluations are taking place in the automotive industry alone, Shea notes.

Amodel's price of just under $3 a pound is closer to the $2.25 average cost of traditional engineering plastics such as nylon than the $4.75 or more commanded by high-performance polymers. "We're going after applications where we know people are looking for high-performance materials but don't want to pay the price," Shea says.

The secret to producing polyphthalamide is a quick synthesis that breaks one of the cardinal rules of polymer manufacturing. Polymer chemists usually seek to avoid turbulence in reaction vessels, but Amoco's scientists opted for it deliberately. After the ingredients are combined with water and heated under pressure, the mixture is injected at high velocity into a heated reaction zone, where it begins to polymerize. The rapid transit through this step ensures that the substance does not remain long enough for heat to degrade it. The roiling fluid continuously disperses particles like a rough sea giving off mist. This vapor is sprayed into another section of the reactor, where a drop in pressure causes it to condense.

When it is molded into parts, Amodel can resist temperatures up to 580 degrees Fahrenheit. Modifying the material with other substances such as glass, carbon and rubber enhances its properties dramatically, Shea notes. Glass and carbon fibers make the polymer strong and stiff. Adding Amodel fibers to rubber yields a material suitable for automotive hoses and fan belts. A super-tough variety resistant to wear and chemicals as well as heat is being introduced to the oil fields of Texas and Alaska, as spacers for drilling pipes.

With visions of a broad market, Amoco has begun building a large-scale manufacturing plant in Augusta, Ga., that Shea says will be "big enough" to meet demand. "Maybe we'll build a second plant in a few years, perhaps overseas," he offers with the 1990s-style understatement. —*Deborah Erickson*



***SOFTWARE MANNEQUINS** are placed in a computer mock-up of a helicopter to test whether a cockpit display is legible. Source: NASA Ames Research Center.*

Human Spec Sheet

Picturing the way the world looks to its users

What size, type and color are needed so that the words "fuel low" are perceptible to a jet fighter pilot whose aircraft begins to shake vigorously as it nears the speed of sound? At how many decibels does a voice become unintelligible through a blizzard of static?

The multidisciplinary field of human factors, an attempt by social scientists, psychologists and engineers to quantify human performance, has tried to supply answers to these and similar questions. But design engineers who wish to tap into this storehouse of data must often rely on graphs and charts that communicate the concept but not the experience.

The Armstrong Laboratory at Wright-Patterson Air Force Base in Dayton, Ohio, and other software developers are trying to fill this gap by expanding computer-aided design beyond the mere simulation of an airplane's structure or electrical connections. The laboratory's Fitts Human Engineering Division is assembling a data base that shows the designer how a display is perceived through the eyes of a pilot.

"Usually computer-aided design is structured to see things from the engineer's perspective," says Thomas B. Sheridan, a professor at the Massachusetts Institute of Technology who was formerly president of the Human Factors Society. "The whole perspective of human factors is to be able to see things from the user's point of view."

Aerospace designers have assumed a leading role in compiling human-factors data for obvious safety-related reasons. "You want to make sure that when a pilot ejects from a fighter plane that he doesn't lose his kneecaps," says Kenneth R. Boff, who began the Wright-Patterson project. Besides aviation engineers, the system is intended for use by designers of an automobile or even a videocassette recorder.

The data base, called Computer-Aided Systems Human Engineering, contains a range of information on human factors, from memory ability to spatial and temporal perception. "These are spec sheets for human performance," Boff says. To use the system, a personal computer extracts information from a compact disc that contains graphics and text that was originally published in a printed volume called *Engineering Data Compendium: Human Perception and Performance*. This work, co-edited by Boff, combines information from leading studies on human factors.

The data base can simulate actual conditions on the computer screen. For example, legibility can be assessed while making text vibrate furiously. Or an image that depicts the blur of myopia can

be inspected—and then cured. "You can take an eyeball, pop it on the software test bench and accommodate it with different lenses," Boff says.

The Wright-Patterson system is one of a number of software tools developed in recent years to help understand how well people use machines that are still just files in an engineer's computer. The University of Pennsylvania has been perfecting a software mannequin called Jack that a designer can place in a simulated cockpit to test whether the operator can reach or see a control panel. A window appears on the display that shows what the controls look like from Jack's vantage point. "These tools can assess whether a design is okay before building a full-scale model," says Norman I. Badler, a professor of computer and information science who is the creator of Jack.

The U.S. Army and the NASA Ames Research Center have jointly used Jack to simulate display legibility in a helicopter cockpit. The research team supplemented Jack with two vision models, one developed by the David Sarnoff Research Center, the other by the Lighthouse, a nonprofit agency for the blind or visually impaired.

The David Sarnoff model predicts to what extent two similar letters—a capital O and Q, for example—can be distinguished at different angles from the pilot's direct line of sight. Concentric circles in the upper left window of the illustration represent the distance (in degrees) from the pilot's visual axis. A colored area in the middle denotes where the pilot can readily discriminate the two letters. The discrimination data are projected onto the cockpit control panel using the Lighthouse model (indicated by contours in lower left and main windows). The Lighthouse model also shows a wire frame of the control panel on a simulated retina (shown in upper right window).

The next step for human-factors specialists is to measure mental work load more precisely. Currently simple questionnaires are used to evaluate whether a pilot has trouble coping with a set of control tasks.

Studies at the Armstrong Laboratory measure brain waves to judge whether manipulating cockpit controls will overload a pilot. New simulations may also draw on the laboratory's investigations into virtual reality, three-dimensional simulations of sight, sound and touch. The designer could then reproduce operator performance—reacting in the heat of a dogfight, driving through a snowstorm or cursing while plodding through the programming of a VCR. *—Gary Stix*

No Accounting for Disaster

In the heyday of the 19th century, small, local fire insurance companies dotted the American landscape. Often they would collect premiums in peace for years until the fateful day someone kicked over a lantern and burned down the town. Then, overwhelmed, the insurer would declare bankruptcy and decamp.

Although regulators now keep close watch on insurance companies—which have grown into multibillion-dollar transnational entities—not that much may have changed. Some natural and artificial disasters may still be too big for the insurers to handle. Events such as earthquakes, floods, nuclear accidents and chemical manufacturing mishaps have the potential to generate claims beyond any company's resources.

In addition, major disasters violate the actuarial assumptions—well-defined probability and statistical independence—that are the basis of most insurance, says economist Ralph A. Winter of the University of Toronto. Roughly the same number of drivers have accidents every year, for example, and the fact that one driver has an accident does not generally affect the likelihood that another driver will also do so. But a Bhopal or a Richter 8.5 earthquake arrives less predictably. The chances are that if one policyholder is filing a claim, many others will be waiting right behind.

Traditionally, insurers maintain actuarial data on the frequency of risks and their average cost. If 1 percent of accident insurance policyholders file claims every year, and the average claim is $5,000, then the annual "expected loss" on each policy (and so, excluding overhead, the premium) should be $50.

But major disasters present a string of years during which no one files a claim, followed by a year during which everyone does. The concept of "expected loss" is not as intuitively clear. Furthermore, without a steady trickle of claims, estimating probabilities is difficult. Techniques such as probabilistic risk assessment (PRA) provide guidelines that are not always applied. Elisabeth M. Pate-Cornell of Stanford University reports, for example, that industrial insurers displayed virtually no interest in PRA until after taking a bath with claims against Union Carbide resulting from the Bhopal accident.

How can insurers and their customers deal with outsize risks? One problem that economists have long identified is the "moral hazard" of insurance—the fact that people have less incentive to avoid a risk once they are insured against it. Disasters put a new twist on moral hazard. Say that a power company might face $10 billion in claims if one of its nuclear reactors melts down, but it only has a net worth of $1 billion. Spending money, either on safety measures or insurance premiums, to bring the expected loss from an accident down to $1.1 billion would be worthless—the company would still go bankrupt. That, points out Steven Shavell of Harvard University, is why nuclear power is so heavily regulated.

A similar moral hazard afflicts insurers facing inordinate payouts in a disaster, whether to a single company or to myriad correlated claimants. Ironically, Douglas B. Bernheim of Princeton University suggests that an insurer in such a situation would face a far smaller risk—merely its net worth, not the cumulative values of its policies. So it would be able to charge lower premiums. (Bernheim, recently arrived from Stanford, reports that earthquake insurance does not conform to his prediction. It is priced well above the level required to cover average rates of property damage. He gambled and won by not buying.)

"The industry is not in a position to insure these kinds of events"
—Howard Kunreuther

In the face of such conundrums, some economists conclude that the private-sector insurance market is simply unsuited to risks that put the very survival of companies at stake. "The industry is not in a position to insure these kinds of events and feel comfortable," says Howard Kunreuther of the Wharton School. He advocates a three-tiered system for insuring against major risks such as environmental disasters: a deductible that policyholders would fund by themselves, insurance and then a government-funded cap on liability to avoid the risk of bankruptcy.

Some plans for insuring against major risks, such as the Price-Anderson Act for nuclear power and Superfund for the chemical industry, already take roughly the form Kunreuther suggests. He contends that even though the government is ultimately responsible for taking care of claims, the plans avoid moral hazard because firms still have substantial amounts of money at risk. Ideally, he says, a plan would be structured so that additional damages always result in some additional liability for the company at fault.

Other economists, however, see such schemes as a devious, misguided form of subsidy. By eliminating risk for claims over a certain amount, the government essentially relieves companies of having to purchase insurance or carry reserves against those losses. Price-Anderson, for example, limits utility companies' liability for claims from nuclear accidents to $7 billion. Geoffrey S. Rothwell of Stanford estimates that the resulting subsidy to nuclear power is more than $22 million per reactor a year, or nearly $2.5 billion a year for the industry as a whole.

Furthermore, according to Shavell, Price-Anderson may have a perverse economic effect on safety because it acts specifically on liability for accidents. "It leads to more need for safety regulations," he says. The same subsidy enacted as a tax break, for example, would not have the same perverse effect, he contends.

Meanwhile the market for insurance against slightly smaller, more common disasters may offer some interesting insights. Carl Shapiro of the University of California at Berkeley, who edited a series of papers on risk and liability in the summer 1991 issue of the *Journal of Economic Perspectives,* reports that the market seems to be working well in pricing insurance for airlines (risks even appear down since deregulation). And Bernheim says that even the erratic behavior of satellite insurance premiums is pretty much in line with what economic models would predict. Perhaps that is the most perverse incentive of all: if major disasters occurred more often, they would be easier and more predictable to insure. *—Paul Wallich*

The city that was true to its nature

Not far from the streets of Philadelphia, nature is alive and well in a 1,200-acre freshwater tidal marsh fringed with woodlands. Canadian geese use it as a migratory stop. Muskrats and red bellied turtles call it home. Years ago, people from a nearby refinery donated the first portion of land to what has become the Tinicum National Environmental Center. Today it is administered by the U.S. Fish and Wildlife Service, with support from the refinery and others.

Do people create places where nature's creatures can live as nature intended?

People Do.

Chevron

For more information, write: People Do-T, P.O. Box 7753, San Francisco, CA 94120.

Painting in Color without Pigments

Curiosity and the thrill of discovery motivate the amateur scientist to try to understand and perhaps control nature. But unlike the professional, the amateur must conduct research with few supplies and modest equipment. The technique I invented for painting without pigments began as an amateur project more than 10 years ago.

Those readers who have developed their own photographs have probably noticed that black-and-white prints are sometimes discolored with browns and yellows. I became interested in figuring out what produced those colors, in particular because I knew that the chemicals used in the darkroom would not serve as pigments.

For the past four decades, THE AMATEUR SCIENTIST has found a place in SCIENTIFIC AMERICAN. The department has taught readers how to build a laser, an X-ray machine and even an electron microscope. SCIENTIFIC AMERICAN plans to continue this tradition by publishing amateur projects that are related to feature articles in the magazine. This month we present the work of Dominic Man-Kit Lam, co-author of "Chromoskedasic Painting," on page 80. Alexandra J. Baran is a chemist who has assisted Lam with many ventures during their five-year acquaintance. THE AMATEUR SCIENTIST will alternate with MATHEMATICAL RECREATIONS, which will next appear in the December issue.

After several years of part-time experimentation, I invented a technique for painting a rainbow of colors on black-and-white photographic papers by applying colorless chemicals. The method utilizes commonly available photographic supplies, and I am delighted to share my methods with the readers of SCIENTIFIC AMERICAN.

My colleague Bryant W. Rossiter of the Eastman Kodak Research Laboratories coined the term "chromoskedasic" to describe my technique. The term "chromoskedasic" means color by light scattering, and my technique exploits this phenomenon. Black-and-white photographic papers contain silver salts that form tiny particles when exposed to light and chemicals. Silver particles that are roughly the same size will scatter certain wavelengths of light and absorb others, producing a specific color. In chromoskedasic painting, the size of the particle is the variable through which color is controlled [see "Chromoskedasic Painting," by Dominic Man-Kit Lam and Bryant W. Rossiter; page 80].

Chromoskedasic techniques require no special skills, although a degree of artistic talent will obviously enhance the results. An eye for art also helps in deciding which images are worth keeping and which are best left in the darkroom.

I highly recommend that readers assemble all necessary materials and dilute all needed solutions before removing the photographic paper from its protective box. For the paintings described below, I used Kodak papers and photographic solutions. Other brands of materials should work just as well.

Readers should heed the warnings on the various chemicals. (You can obtain additional hazard information, free of charge, from a photographic chemical supplier by requesting "Material Safety Data Sheets" for specific products.) When mixing photographic chemicals, wear safety glasses and old clothes that cover you completely. Some of the chemicals are strong acids that can be very harmful if ingested or splashed on the skin or eye. Never pour water into a concentrated acid. Dilution should always be performed by carefully adding the acid to the water. All the chemicals mentioned in this article should be diluted to the desired concentration with water.

Over the years, I have made hundreds of paintings and have taught the technique to my daughter and to friends. In the following paragraphs, I describe three projects in order of increasing sophistication. I suggest that readers first adhere to my instructions and, before moving on to the next project, experiment a bit to get a feel for the process.

I used only the most basic chromoskedasic methods to create *Among the Coral Reefs,* the abstract painting at the top of the opposite page. I removed a piece of Kodak Polycontrast 3RCF black-and-white photographic paper (60 by 40 centimeters) from its protective covering. After exposing the paper to dim red light for five minutes, I poured

AUTUMN IN NEW ENGLAND

about half a cup of a 10 percent solution of Kodak Ektamatic S30 Stabilizer onto the paper. I swirled the stabilizer on the paper in an attractive pattern. The regions exposed to the stabilizer would eventually emerge as light yellow.

On other areas of the paper, I immediately poured a small amount of 10 percent solution of Kodak S2 Activator, followed by a 50 percent solution of Dektol. These chemicals produced reds, oranges, yellows, greens, blues and grays. I exposed the paper to light from a fluorescent lamp for five minutes. Finally, I bathed the paper in a 50 percent solution of Kodak Rapid Fixer, then rinsed the paper with water and allowed it to dry.

With a little more effort, readers should be able to create images like *To Run,* shown at the right below. Using an artist's brush and black ink, I painted a running horse on rice paper. I then took a photograph of the horse using Kodak black-and-white film and developed the negative. I used a photographic enlarger to transfer the negative image onto a piece of Kodabrome paper (20 by 25 centimeters).

Immediately after, under red light, I poured a 10 percent solution of stabilizer onto the paper and swirled it around. I did the same with a 10 percent solution of S2 Activator on the same part of the paper; the stabilizer and activator combined to make oranges. To produce a deep red-orange color, I applied a 50 percent solution of Dektol. Using a rag, I spread a 50 percent solution of fixer over the entire paper and then submerged the paper in the fixer for 20 minutes. Finally, I washed the paper in a water bath for 30 minutes and let it dry.

To create *Autumn in New England,* the impressionist painting shown on the opposite page, I used both chromoskedasic techniques and ordinary enamel paints. Under red light, I cut a 100-by-300-centimeter piece of black-and-white, polycontrast photographic paper from a large roll. Using a brush dipped in full-strength S30 Stabilizer, I drew the peak of a snow-covered mountain. I then turned on the fluorescent room lights, exposing the paper for five minutes.

To add ivory-colored details on the mountain, I brushed on, under red light, a 20 percent solution of stabilizer. I painted other areas with the diluted stabilizer to make the outline of a lake. To produce brown tones, I worked these areas over with a 50 percent solution of S2 Activator. After dipping a fountain pen in the activator, I outlined trees in the foreground of the picture. I shaded the trunks brown by applying stabilizer and then activator.

AMONG THE CORAL REEFS

TO RUN

I exposed the material to fluorescent light for a total of about 30 minutes and then applied a 50 percent solution of fixer with a spray bottle over the entire piece of photographic paper. After 20 minutes, I washed the entire paper with a garden hose and allowed it to air-dry. I then used enamel paints to add white mist on the mountains, blue water in the lake and yellow and red leaves on the trees.

I am sure readers will find many interesting effects by experimenting with such chromoskedasic techniques. I cannot even begin to imagine the various works of art that might be created using this medium. Will the new medium provide a palette for the creative mind of a new da Vinci or Picasso?

FURTHER READING

NEBLETTE'S HANDBOOK OF PHOTOGRAPHY AND REPROGRAPHY: MATERIALS, PROCESSES AND SYSTEMS. C. B. Neblette. Van Nostrand Reinhold, 1977.

BOOK REVIEWS by Philip Morrison

Kaleidoscope Order

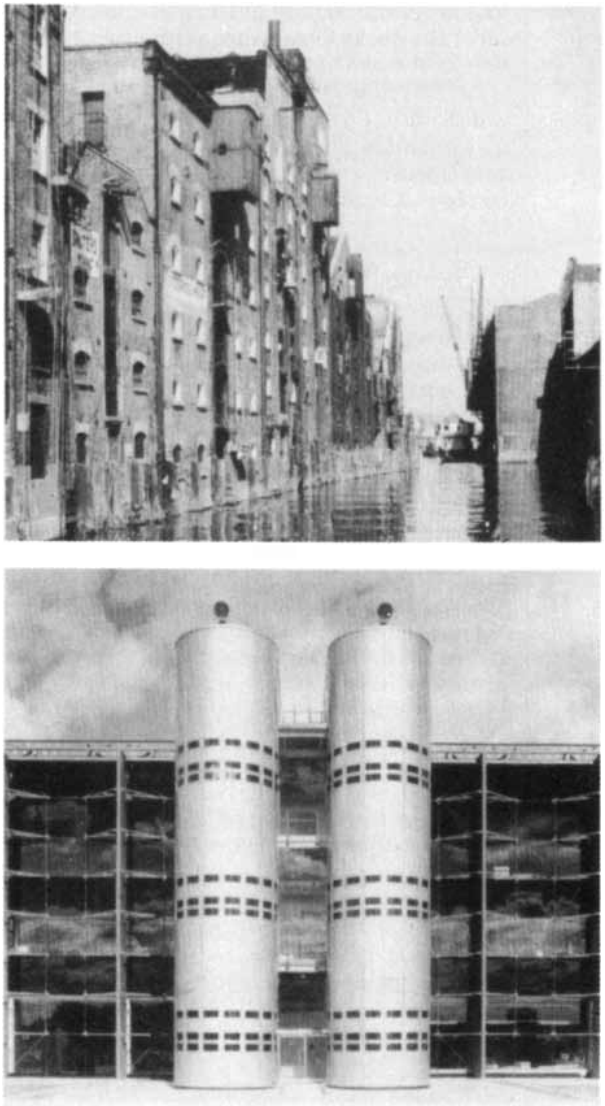
LONDON DOCKLANDS, by Stephanie Williams. A VNR Architecture Guide. Van Nostrand Reinhold, 1990 (paperbound, $24.95).

From Marblehead to Bombay, a dock is a solid finger, small or large, sent outward from shore to lengthen the interface between land and water. In a few ports, the geometric game is inverted: fingers of water instead run into the land, narrow channels that usually end in sizable basins where ships may deliver or take on their burdens.

London's Docklands are the most famous of these inversions, or wet docks. By decree of Elizabeth I, the mushrooming sea trade of her realm was assigned about a quarter of a mile of "legal quays." These ran along the smooth riverbank eastward from the Tower of London, just at the head of the Pool of London, where by law all dutiable goods had to be landed. By the 1790s, the entire Pool was daily clogged by 1,000 ships, served by 3,500 barges and lighters. Rich-laden ships spent days to make good a few dangerous yards of river; they queued as long as six weeks wating to be unloaded. To the insufferable delay was added a lively threat of well-organized pilferage.

Father Thames is no ample harbor, but a modest river; a fringe of docks sent out into the river would have choked passage still more. The West India merchants therefore proposed a set of wet docks to be let into seven areas of the boggy flatlands along both sides of the Thames as it meanders eastward to Greenwich. In 1802 the first high-security dock was opened, its warehouses and basin enclosed and guarded like a fortress. The West India Dock had room for 600 sailing ships. The spoil heap from the requisite earthworks that formed the first wet docks is still there, the Mudchute. Not far off is the pioneer shipyard where the gigantic iron steamer *Great Eastern* was built and launched after an epic struggle against gravity. The rural population was displaced; industrial London had found its sites too in the few peninsulas bounded by the deep bends of the Thames.

The 19th century generated both the amenities, the middle-class residences, the parks and the museums westward in London, and the murky, teeming, often miserable workplaces and dwellings of Docklands. The docks covered about 10 square riverside miles adjoining the single square mile that was and is financial London, from Lloyd's to the Bank of England. Docklands takes its full place in the prose of Dickens and the engravings of Doré. A little closer to the City than the West India Dock lies the old district of Limehouse, around the basin built under the direction of Thomas Telford to link the Thames to the national canal system. Legendary until World War II, Limehouse became far-famed for its aura of lawlessness and debauchery, its opium dens and brothels. To remind us of the complexity of London life, the book shows us as well the superb tower of the 18th-century church of St. Anne's Limehouse, with its unaltered organ installed in 1851 "after winning the organ prize at the Great Exhibition."



THE OLD* (top)*: towering warehouses drew forth Dickens's graphic portraits of poverty. The new* (bottom)*:* Financial Times *Print Works, "the finest building in Docklands."

Victorian London extended the docks eastward at breathtaking scale, with still more surrounding heavy industry. The largest gasworks in Europe, lighting and fueling the metropolis with coal gas, was built there, to be torn down only in 1985. (Its great heaps of ash and waste, wonderfully named the Beckton Alps, are now the site of a dry ski slope.) The Royal Docks are the easternmost in Docklands, some five miles from the financial district. Once they formed the largest enclosed dock system in the world, busy up to the mid-20th century, especially in the grain trade. Today no ships come; the tall warehouses are derelict. "Nowhere is the landscape in London more desolate," although its backdrop is still theatrical, "gigantic, abandoned, stalking cranes…and monumental grain silos." Now the action is on the single runway of the new London City Airport that stretches out alongside the still waters of the Royal Albert Dock.

The *Luftwaffe* had already burned out much of Docklands, then a sprawling economic target, before the docks ended for good in the 1960s.

Jobs and housing both faded away step by step as London—indeed, much of Britain—gave up first on heavy manufacturing, then on light. New container and specialized bulk ports, new railways, trucklines, even the turn from coal to North Sea gas and oil, all contributed to the change, which has its parallels worldwide.

Docklands was born out of the crisis of the exponentially rising sea trade of a mechanizing Britain. Within 30 years after the start of the West India Docks, Docklands had been given its form through wild expansive growth. A rush of speculation and investment left a cityscape of grim if sturdy warehouses, walled and guarded docks, diverse industries, dormitory housing and, finally, the crowded dwellings of the marginals and the working poor. Hit hard by war, then by powerful economic currents, Docklands idled and at last emptied, its jobs down one full order of magnitude; the docks closed down from the time of the mid-1960s.

But there the land lay, flatlands (and water) with a developed area the size of Manhattan south of Central Park, convenient to the ever busier City of London. A new boom entered Docklands after the government funded the London Docklands Development Corporation in the summer of 1981. LDDC acted decisively to draw private investment. Once again "the market has ruled." After one decade, a strew of glass-walled office buildings houses service industries and eager commerce. Home ownership flowers, some of the 13,000 new dwellings kept low in price.

This small book is a district-by-district guide to the unfinished history of this place as of 1990. Almost every double spread bears a page of building photographs or drawings (all black-and-white) and a knowing and apposite history of the structure shown, with an architectural critique, by no means always circumspect. Docklands holds "one of the worst collections of late 20th century building to be seen anywhere in the world…you want to cringe…with the hopeless crassness… of it all." Gems are here as well. Two essential new pumping stations, a large one and a small, are structures rich in unexpected color. "One of the most splendid buildings," a secure concrete control house for storm water pumps, is turned by its ornamented roof and columns and visible fan into a symbolic appreciation of hydraulics, with a fragrance of China but without pretense or cliché. There are a couple of exemplary groups of owner-planned and owner-built terrace houses, designed and occupied mainly by building tradesmen.

Ongoing is the largest building project in Europe, Canary Wharf, two dozen modern commercial office buildings, along with their less pinstriped kin, the shops, hotels and pool, that make up a corporate city of transatlantic pedigree and scale. It will employ 40,000 when it is finished in 1997, business cycle permitting. Industry is here, too, but now it is mostly informational, a dramatic printworks for the *Financial Times*—"the finest building in Docklands"—and a couple of trim microwave- and computer-crammed centers, "anonymous" and "secure." Soon, as many people will work in Docklands as now work in the city of Bristol, some 200,000, perhaps half of them living in Docklands as well. Although the flowing Thames will never run twice the same way, in the Docklands narrative are hints of kaleidoscopic order among these fragments bright and dark.

Sacred Soccer, Ritual Roads

CHACO & HOHOKAM: PREHISTORIC REGIONAL SYSTEMS IN THE AMERICAN SOUTHWEST, edited by Patricia L. Crown and W. James Judge. School of American Research Press, 1991. Distributed by the University of Washington Press ($35; paperbound, $15.95).

This volume reports a protracted, subtle and fascinating argument carried over a dozen chapters. There are also a prologue and a consensus (visibly a bit uncertain) among the 12 authors, all expert archaeologists of the Southwest. They seek for the first time to compare long unrelated interpretations for two of the best-known prehistoric cultures, the Chaco and the Hohokam. More or less contemporary, directly adjoining and yet distinct, these two ancient civilizations are full of wonder and of puzzles.

Stretching from a point just northwest of the Four Corners in Utah, far across the grasslands and piñon-juniper scrub plateau to the edge of the Rio Grande Valley in New Mexico, certain ruined settlements of the old people share a remarkable common feature. At roughly 70 of the far-flung masonry ruins lies some fragment of a prehistoric road. The careful construction of these roads is recognizably that of the "formal, engineered" roads that still link a score of large masonry structures in isolated Chaco Canyon far up along the San Juan River. There, huge, symmetric, often multistory pueblos form the geographic and probably the

cultural core of this pattern. Dating by tree rings is on the whole reliable, especially in the canyon itself, where masonry construction began about A.D. 900. The last big log to be cut (and man-hauled from the distant mountains) for those buildings can be dated by its rings to A.D. 1132. The place dwindled steadily thereafter for a century or two. We call this world Chaco, and of course we know much more of it than these remarkable roads.

On the other side of the arid desert of southern Arizona, from the Verde Valley southeastward to some 50 miles beyond Tucson, wherever the land bears tall saguaro cactus instead of trees, the archaeologists have for some time identified sites of another long-lived and distinctive culture. As maize fed the skillful dry farmers of Chaco, these Arizonan people too, "First Masters of the American Desert," had maize for their staple. Although they dwelt where as a rule rainfall cannot support a crop, they lived by systematic irrigation. Their admirable weirs, inlets and canal grids took life-giving water from natural springs and watercourses to spread it frugally over the terraces the villagers farmed. We call them Hohokam; much has been said about the 1,500 years of Hohokam growth and eventual decline.

The Hohokam canals were manifestly "starkly functional," whereas the roads of the Chaco people, similarly taxing as public works projects built and maintained over centuries, seem beyond clear necessity. No pack animals, no wheeled carts, traveled those roads. Backpacking maize over the distances we see cannot usually pay off in energy currency. How can we explain the costly Chacoan roads?

Moreover, the archaeologists point out no fewer than 163 Hohokam settlements over the entire region that share another remarkable piece of public architecture, a large oval depression or two, usually surrounded by earthern embankments. Most of these oval sites can be dated roughly by a variety of not always concordant methods, say, by distinctive potsherds found elsewhere as datable intrusions. (Tree rings do not work in the saguaro desert.) The earliest Hohokam oval was built about A.D. 750, the last around A.D. 1100. The largest of about 200 ovals is found at a famous village ruin, Snaketown, in the Salt-Gila River Basin near Phoenix.

What are those ovals, no more economic than the Chaco roads may be? The answer was already given in the 1930s; strongly challenged since, it is now more or less consensual. They are the ballcourts for the sacred ball game

we know well from the Mayan cities of the Yucatán. They are functionally comparable to their Mesoamerican prototypes yet different enough in form to suggest that it was the ball game and not the court construction itself that diffused from Mexico.

"Canals were the spine of Hohokam subsistence and settlement; roads were filaments in a larger yet looser web at the center of which was Chaco Canyon." So argues one Arizonan archaeologist. But ballcourts? Views are shifting. Whereas once common features of the settlements of both Hohokam and Chaco were seen as the actions of rather independent villagers who knew a little about each other, a more complex centralized system is now proposed for both regions. Chaco Canyon is a core area poor in resources that served as a center for equalization of subsistence in a region of varied fortunes; exchange was ritually maintained. Perhaps the road users were often pilgrims of labor or of succor. The ballcourts served Hohokam too as "places for ritual games." Both centers then drew in and instructed the outliers, for purposes of peacekeeping, planning, ritual and resource exchange. Eventually the stress of prolonged drought (with rare flooding) overcame the benefits of complex organization, and the regional systems declined quickly below the roughly 30,000 people they may each have served at their zenith.

If you would learn about computer simulations of ancient water use or of dragging timbers out of the mountains, about turquoise, macaws and mosaic mirrors, consult the book and its imposing list of 500 studies!

Vedas to Vegas

NET OF MAGIC: WONDERS AND DECEPTIONS IN INDIA, by Lee Siegel. University of Chicago Press, 1991 ($60; paperbound, $19.95).

It is the fullness of India that most astonishes the traveler: a dozen tongues, a hundred gods, a thousand costumes, a million pilgrims. In spite of such manifold evidences of reality, the philosopher-priests of the land have since Vedic times contrived to ascribe all this wonder to illusion. It is *maya,* the net of the world itself, in which Indra, who rules the gods of all creation, has entangled every human eye and mind. It is not high Indra but human beings who affix the claims of a less potent magic to his. The kernel of the paradox that besets all the magic by the most cunning performers is concisely put by our author: " 'I'm writing a book on magic,' I explain, and I'm asked, 'Real magic?' By *real magic,* people mean miracles...and supernatural powers.... *Real magic,* in other words, refers to the magic that is not real, while the magic that is real, that can actually be done, is *not real magic.*"

Our guide to India and the magic done there, both real and imaginary, is himself a scholar-magician, Sanskritist and conjurer, professor of religion at the University of Hawaii. He is both a learned academic and a prepared and prying American visitor passing among the magicians of India. Convincingly authentic where that counts, his book is also a funny parody and a wry satire, aimed as much at gaudy and deceptive Las Vegas as at pretentious Calcutta.

The structure of the book allows a reader to part fact from fiction (so one is led to believe, and an external inquiry or two did confirm the claim). Field notes, traveler's chronicles both of today and of ancient Moghul courts, instructive and amusing fictional models, outrageous confidence men and even one or two celebrated godmen appear. The catchy entertainment manages as well to be a serious eye-opener for any reader intrigued by the high question: What ought we to believe?

An insider filled with empathy for his charming, raffish characters, Professor Siegel sees his work mainly as an examination of the aesthetics of wonder. How the magic works is a small, almost banal part of the entire skein of myth that is implicit behind each performance. Siegel does not systematically disclose the cunning simplicities of the conjurers, but helpful reality does creep in. Those devices are worldwide. Indian magic is now infiltrated by the gimmicks and illusions of the West. Yet so familiar a piece of American stagecraft as the performer's ingestion of a batch of needles and thread, to be removed at once from the mouth as a string of threaded needles, is a very old Indian trick, taken westward decades ago without attribution.

A few hints can be given with advantage, although certainly it is not really urgent for the author to explain just how poor street entertainers are able to pick mangoes from trees that they grow in a few minutes, step by concealed step, from road dust. (By the way, photographs taken on Delhi streets a few years ago show the "mango trees" scarcely two feet high, as old engravings indeed confirm. That does not prevent many an eloquent eyewitness from recounting how the full tree towered above them.) General arguments alone eliminate all chance that

this is a form of instant horticulture.

But the scientific reader thirsts for a few details, to illustrate the well-known abstract categories of magic, its misdirection, falsified assumptions, variable means, and the like. The author gives one startlingly simple example. No normal show, it was the specially commissioned performance of ritual self-mutilation by a group of devotees of a certain religious leader in Kerala. Their high purpose is to take on themselves the pain of their host and patron. They in fact do slash their tongues, pierce themselves with needles and nails, slice their belly skin with daggers. Such painful wounds will heal. But "when the ecstasy was so intense that it seemed to shake the earth," the old leader pushed a stiletto point into his eye. "A translucent liquid…dripped down the shining blade." One twist, and the eyeball itself came visibly out on the knifepoint! Reader, your anxiety is misplaced; according to the host, the old man had done this very feat for years. Long before, he had lost one eye for good, in some mishap during self-mutilation. Now, before every performance, he inserts a fresh sheep's or goat's eye into the empty socket.

The Indian Rope Trick runs as a motif throughout, with a long, delicious solo passage played among the old books and the new conjectures of professionals in India. Its several themes can be easily found one by one; ropes can be stiffened—the everyday street method merely encases a length of bicycle chain in a woven cotton tube; boys can be levitated; a bloody decapitation is neatly undone. But the whole astonishment remains beyond our testing, covered by 1,000 years of ever renewed myth. One of the book's dreamlike chapters discloses to us the mind of a magician who imagines that at last he has found the modern way: helium! (His calculations, alas, are off by a power of 10.) Less ambitious levitations usually follow one of two schools: an iron framework hidden under the costume of a real person aloft by a yard or two acts to cantilever the load to a visible central staff; or a feather-light mannequin gives human semblance to a figure floating in the air.

Few books tell so much so deeply about magic or about the grounds of belief itself. In the beginning, magic is mainly the word, but that word is of man. The magician's witty patter evokes unfulfilled desire, personal fear and even the sacred. The unvarying ritual, the abracadabra of the Indian magician, closes this unique book, itself a magic show: *Om! Gilli-gilli-gilli! Yantru-mantru-jalajala-tantru!*

ESSAY: ANIMAL THINKING by Donald R. Griffin

A cognitive revolution has rekindled scientific interest in human and animal thinking. Psychologists and ethologists have begun to recognize that internal cognitive processes such as learning and memory, problem solving, concept formation, expectation, intention and decision making have important effects on animal behavior. This, together with computer simulation of many types of human thinking, has replaced the behaviorism that used to recognize only observable behavior as worthy of scientific consideration. One major facet of behaviorism has, however, remained strongly entrenched: an implicit assumption that all animal cognition is unconscious.

It is difficult to believe that all animals pass their entire lives as unconscious "sleepwalkers." We know very little about what differences there may be between conscious and unconscious processing of information, and still less what neural processes lead to decisions about which of various possible actions will achieve desirable results or avoid unpleasant consequences. Yet as far as we know, the central nervous systems of animals operate on the same basic principles as human brains, and no uniquely human neuronal or synaptic mechanisms have been discovered. Much of the experimental analysis of brain function has therefore been based on studies of animals.

This residue of behaviorism is customarily justified by the claim that although animals may have conscious mental experiences, these are hopelessly inaccessible to scientific analysis, and thus hypotheses about them cannot be objectively tested. The same skeptical philosophical arguments have been applied to human consciousness. Yet we manage to understand the thoughts and emotions of our companions with reasonable accuracy, even though we cannot prove precisely what they are thinking or feeling. The crucial difference is generally held to lie in the unique human ability to report mental experiences, largely through our use of language. How, then, can we ever discover what animals might be thinking?

Cognitive scientists who study animal behavior have recently begun to analyze evidence that animals sometimes experience simple conscious thoughts about matters of crucial importance to them. Of course, there is an enormous difference in complexity between human and animal thinking. Animal thoughts may well be limited to simple perceptual consciousness, thinking about the concrete world—food, predators and social companions—without any introspective thinking about the thoughts themselves. Thinking about the likely results of alternative actions is much safer than trying them out in the real world, where mistakes are often fatal.

The versatility with which animals cope with the challenges they face often suggests that they are indeed thinking about what they are doing. Chimpanzees, for example, prepare sticks suitable for probing into termite nests, often being obliged to do so at some distance from the place where the probes will be used to obtain food. An occasional green-backed heron fishes with bait. The heron breaks a small twig into even smaller pieces, carries one to the water's edge, drops it in, watches it intently and seizes small fish attracted to the floating bait. But behaviorists still argue that no matter how skillful or ingenious its behavior may be, there is no way of telling whether the animal is conscious of what it is doing.

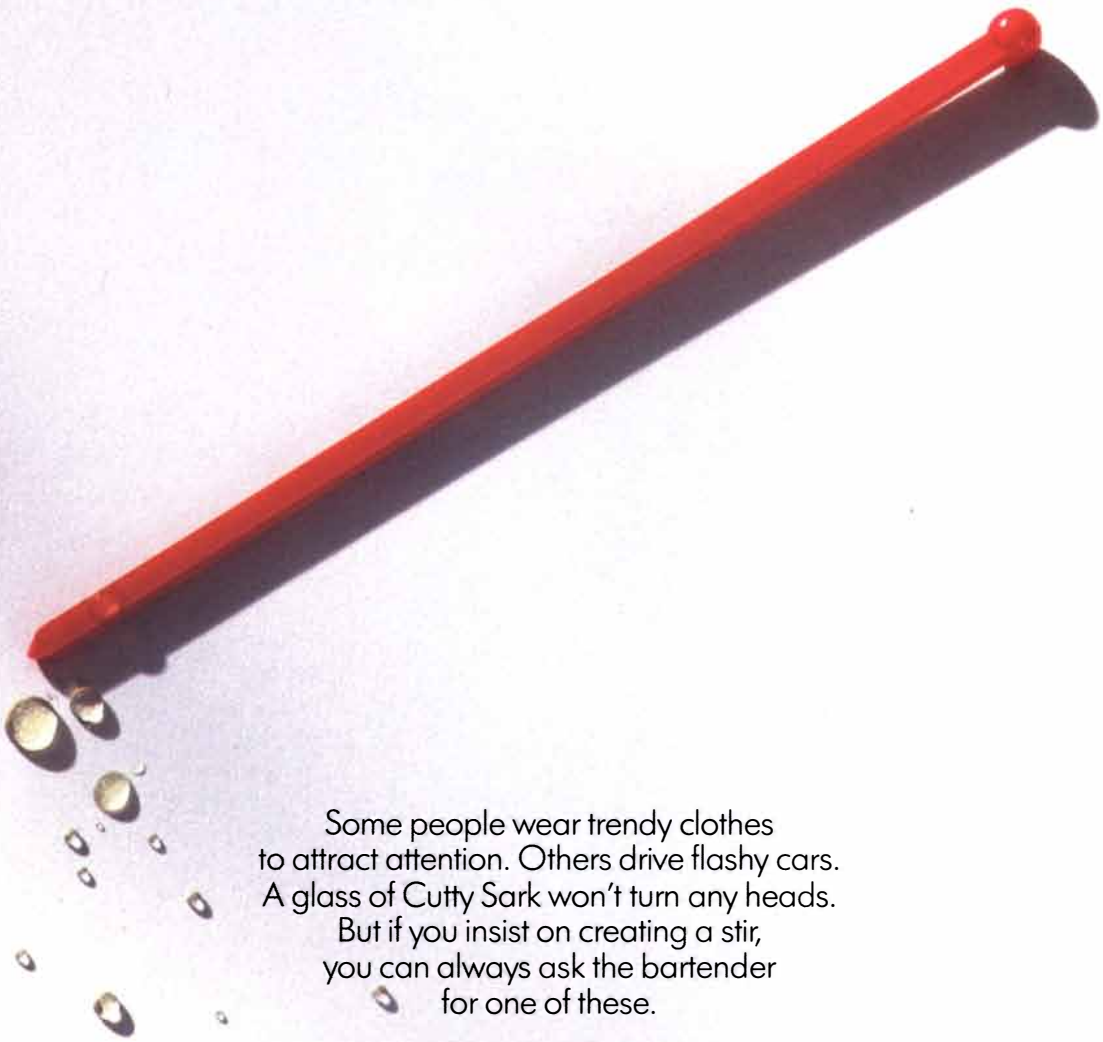
A more promising approach to the problem of identifying conscious mental experiences arises from the versatility and scope of animal communication. Although customarily viewed as incidental by-products of physiological states—roughly analogous to groans of pain—some animal signals are now known to convey semantic information about significant objects and events. These may be statements about immediately present situations, for instance, the alarm calls of vervet monkeys that inform their companions what kind of predator is threatening and thus what escape behavior is called for. In other cases, animal signals identify things that are not part of the immediate surroundings but must be remembered or anticipated. For instance, the waggle dances of honeybees convey the direction, distance and desirability of food sources, water or cavities suitable for a new colony—all of which are removed in time and space from the situation where the communication takes place.

Animal communication often involves reciprocal exchanges, as one animal responds to the signals of another, which then alters its own signals in a simple dialogue. Such communication sometimes leads to important group decisions, like the honeybees' choice of which cavity is most appropriate for establishing a new colony.

When animals make strenuous and mutually adjusted efforts to communicate with others, they may well be reporting simple conscious thoughts as well as emotional feelings. Of course, we cannot rigorously prove this in any specific case, any more than we can be absolutely certain what other people think or feel. But a comparable level of useful and generally accurate inference seems reasonable. An African Grey parrot has been trained to use imitations of English words to ask for things it wants to play with, to answer questions about color or shape, to say whether two objects are the same or different, and, if they differ, whether in shape or color. This parrot, and the several chimpanzees who have learned to communicate by means of keyboards or by gestures derived from American Sign Language, expresses simple statements and requests. Although such communication is extremely limited compared with the tremendous scope of human language, it is quite sufficient to convey uncomplicated thoughts.

The communicative signals that animals exchange provide objective data for scientific analysis. For instance, experimental playbacks of recorded calls and observation of other animals' responses help clarify just what information is conveyed. Thus, the critical interpretation of animal communication constitutes a useful window on animal minds. Communicating animals may be able to tell us directly and explicitly about some of their thoughts and feelings, once we are prepared to listen.

DONALD R. GRIFFIN is best known for his discovery (jointly with Robert Galambos) that bats use echolocation. He has investigated animal behavior for many years while at Cornell, Harvard and Rockefeller universities. Since retiring from Rockefeller, he continues his work at Harvard's Concord Field Station.



Some people wear trendy clothes
to attract attention. Others drive flashy cars.
A glass of Cutty Sark won't turn any heads.
But if you insist on creating a stir,
you can always ask the bartender
for one of these.

CUTTY SARK®

SCOTS WHISKY

UNCOMMONLY SMOOTH