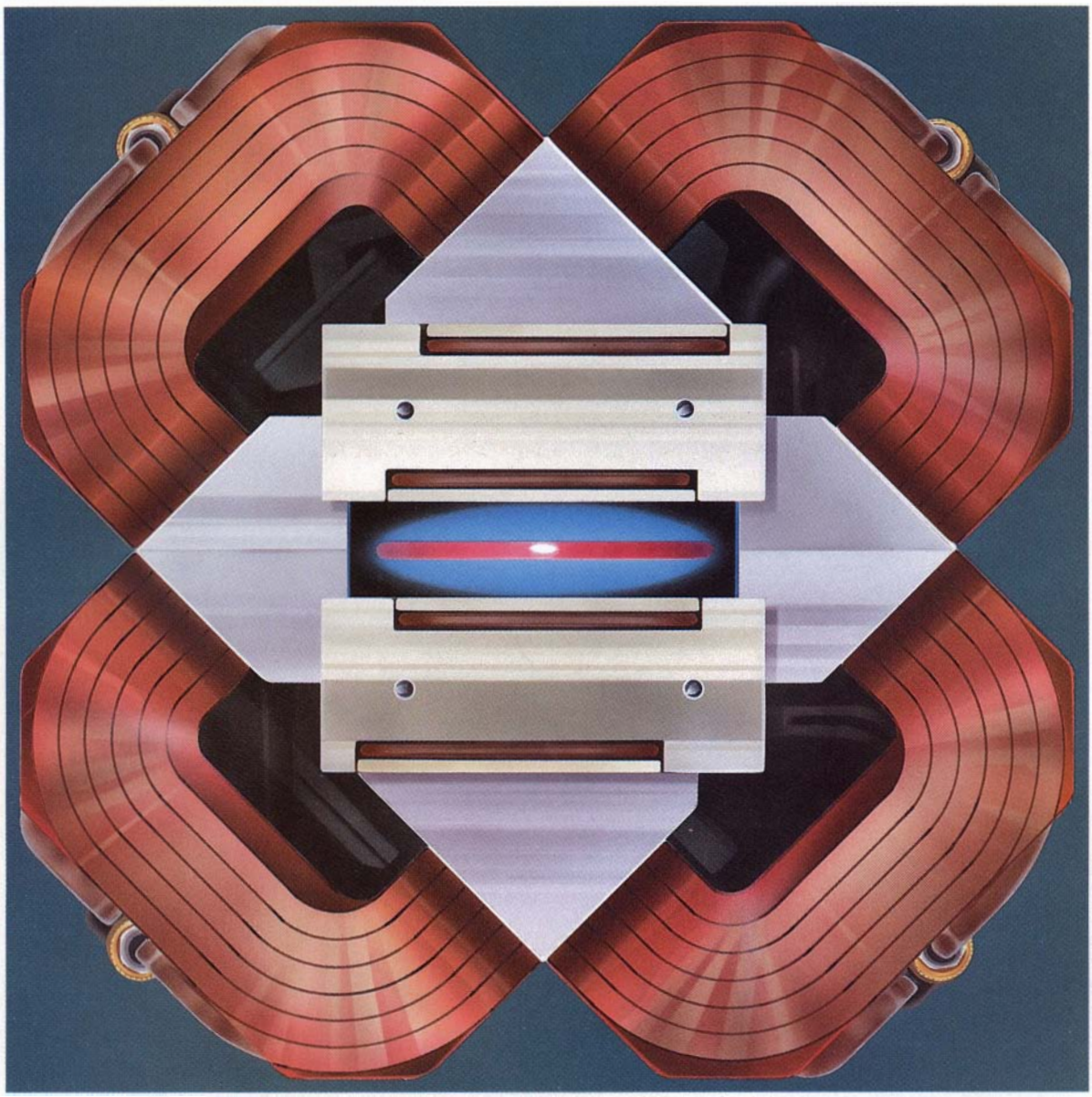
SCIENTIFIC AMERICAN

APRIL 1989
$2.95

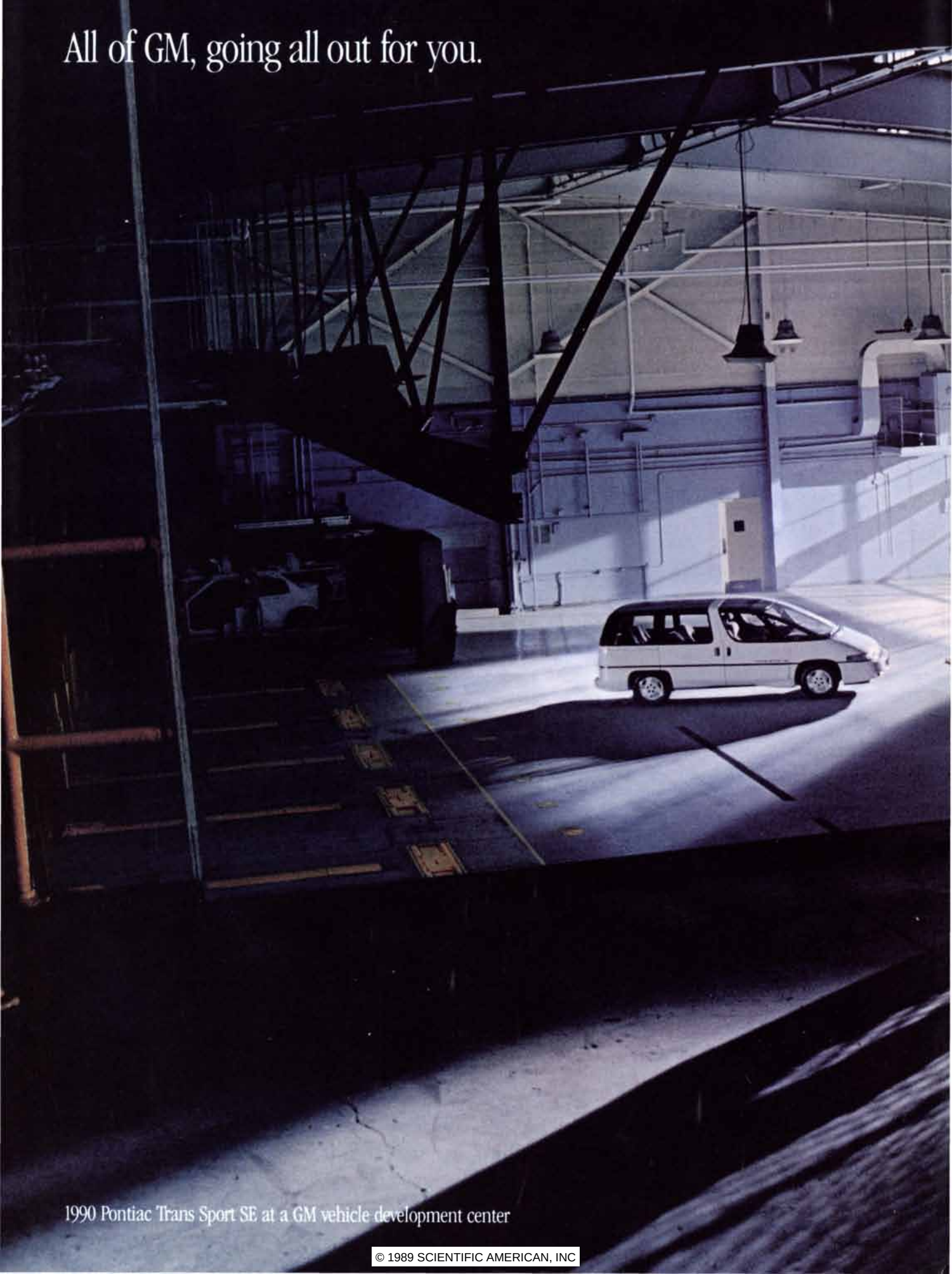
Is the greenhouse effect already warming our planet?

Starlike cells that help to shape the brain.

Why some children are resilient under stress.



Free-electron laser, *seen end on, has potential applications ranging from surgery to fusion power and strategic defense.*



"We've got lots of long hours and hard days wrapped up in this one... taking it all the way from the design studio to reality. You factor it all in... safety, comfort, performance. Make changes, improvements... and make sure it ends up beautiful.

You know, I think we did it."

Fact: GM has been recognized for design excellence by the Industrial Designers Society of America more than any other U.S. company.

GM

MARK OF EXCELLENCE

Chevrolet, Pontiac, Oldsmobile, Buick, Cadillac, GMC Truck

SCIENTIFIC AMERICAN

April 1989 Volume 260 Number 4

DEPARTMENTS

Some cars talk.

Jetta Can we talk?
When it comes to phrases like "your door is ajar" quite frankly we're at a loss for words.

You see, to our way of thinking, cars should carry people. Not conversations.

A Volkswagen is engineered to listen.

To listen when you ask it to accelerate briskly. Stop smartly. Handle vigorously.

To listen come hail or high water.

Whether roads be terrific. Or horrific.

Take our Jetta, for instance.

It has our patented, track-correcting rear axle, front disc brakes and a fuel-

This one listens.

injected engine bred for the Autobahn. So, it willingly does exactly what you ask.

True, our Jetta doesn't talk.

Yet it is undoubtedly a car that speaks to people who love to drive.

To this day it remains the best-selling European nameplate in America.

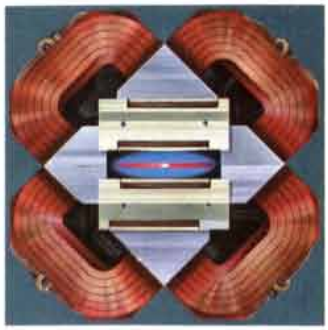
For details on the 1989 Volkswagens call 1-800-444-VWUS.

If you feel like talking things over, we know someone out there who'd be glad to listen.

He's called a Volkswagen dealer.

And like our cars, he's all ears.

German engineering. The Volkswagen way.



THE COVER painting looks down the bore of a free-electron laser, from which both laser light and an electron beam are seen to emerge (see "Free-Electron Lasers," by Henry P. Freund and Robert K. Parker, page 84). The electromagnets contained in the rectangular housings just above and below the beam cavity force the electrons into a side-to-side motion. The slewing beam interacts with the magnetic field of the incoming light wave, slowing down the electrons and amplifying the light in coherent fashion.

THE ILLUSTRATIONS

Cover painting by Ian Worpole

Page	Source
37	Fred Hirschmann
38–40	Thomas C. Moore
41	Markus Aellen, Swiss Federal Institute of Technology
42–43	Thomas C. Moore
49	Mark Henkemeyer and F. Michael Hoffmann, McArdle Laboratory for Cancer Research, University of Wisconsin
50–51	Neil O. Hardy
52	Henry Pitot, McArdle Laboratory (*top left*); Neil O. Hardy (*top right*), Jeffrey Ross (*bottom*)
53–54	Neil O. Hardy
57	Akira Tonomura, Hitachi Ltd., Tokyo
58	Akira Tonomura (*left*), Gabor Kiss (*right*)
59–61	Gabor Kiss
62	Richard A. Webb
67	Andreas Karschin, Heinz Wässle and Jutta Schnitzer, Max Planck Institute for Brain Research, Frankfurt
68–69	Seward Hung
70	Lowell W. Lapham, University of Rochester Medical Center
71	Michael D. Norenberg (*top*), Seward Hung (*bottom*)
72	Dennis M. D. Landis, National Institute of Neurological and Communicative Disorders and Stroke
74–76	Seward Hung
85	Yves Petroff and Jean-Michel Ortega, LURE University of Paris, Orsay
86–87	Ian Worpole
88	Jon Brenneis
91	Deane Renouf
92–94	Patricia J. Wynne
96–97	Paul Lipke
98–99	George Retseck
100	Piergiorgio Sclarandis, Black Star
101–102	George Retseck
103	Piergiorgio Sclarandis, Black Star
107	Mike Teruya/Free Spirit Photography
108–110	Bob Conrad
111	John Wehrheim and Kauai Historical Society
112–114	Michael Goodman
116	Edward Bell
117	Andrew Christie
118	Michael Goodman

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017. Second-class postage paid at New York, N.Y., and at additional mailing offices. Authorized as second-class mail by the Post Office Department, Ottawa, Canada, and for payment of postage in cash. Subscription rates: one year $24, two years $45, three years $60 (outside U.S. and possessions add $11 per year for postage). Subscription inquiries: U.S. only 800-333-1199; other 515-247-7631/32. Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51593.

SCIENTIFIC AMERICAN

Established 1845

EDITOR: Jonathan Piel

BOARD OF EDITORS: Armand Schwab, Jr., *Managing Editor;* Timothy Appenzeller, *Associate Editor;* Timothy M. Beardsley; John M. Benditt; Laurie Burnham; Elizabeth Corcoran; Gregory R. Greenwell; John Horgan; June Kinoshita; Philip Morrison, *Book Editor;* Tony Rothman; Ricki L. Rusting; Russell Ruthen; Paul Wallich; Karen Wright

ART: Samuel L. Howard, *Art Director;* Murray Greenfield, *Associate Art Director;* Edward Bell, *Assistant Art Director;* Johnny Johnson

COPY: Sally Porter Jenks, *Copy Chief;* M. Knight; Michele S. Moise

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

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

ADVERTISING OFFICES: NEW YORK: Scientific American, 415 Madison Avenue, New York, NY 10017; Robert F. Gregory, *Advertising Manager;* Kate Dobson, *Advertising Manager;* Lisa Carden; John Grant; Meryle Lowenthal. CHICAGO: 333 N. Michigan Avenue, Chicago, IL 60601; Patrick Bachler, *Advertising Manager;* Litt Clark, *Midwest Manager.* DETROIT: 3000 Town Center, Suite 1435, Southfield, MI 48075; William F. Moore, *Advertising Manager;* Edward A. Bartley, *Detroit Manager.* CANADA: Fenn Company, Inc. DALLAS: Griffith Group. PRINCETON: William Lieberman, Inc. WEST COAST: Frank LoVerme & Associates

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

INTERNATIONAL: FRANKFURT, GENEVA, LONDON, PARIS: Infopac. HONG KONG/SOUTHEAST ASIA: C. Cheney & Associates. SEOUL: Biscom, Inc. SINGAPORE: Cheney Tan Associates. TOKYO: Nikkei International, Ltd.

ASSOCIATE PUBLISHER/BUSINESS MANAGER: John J. Moeling, Jr.

PRESIDENT OF MAGAZINE DIVISION AND PUBLISHER: Harry Myers

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

PRESIDENT AND CHIEF EXECUTIVE OFFICER: Claus-Gerhard Firchow

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

CHAIRMAN OF THE BOARD: Georg-Dieter von Holtzbrinck

CHAIRMAN EMERITUS: Gerard Piel

The difference between a shattered windshield and a shattered life.

On November 17, 1986, the car in which Tara Meyer was riding was in an accident. Unfortunately, she wasn't wearing her seat belt. So like 150,000 other victims each year, her head hit and shattered the windshield.

But unlike many of the others, Tara suffered no facial cuts or lacerations.

The difference?

A new idea developed by a group of DuPont engineers. They believed that a thin layer of plastic along the inside of the windshield would keep the razor-sharp, shattered glass on the outside—away from the occupants.

Their idea helped make more than just a safer windshield . . . it made a difference.

And that's something we try to do in everything we do.

At DuPont, we make the things that make a difference.

Better things for better living. DU PONT

REG U.S. PAT. & TM. OFF.

LETTERS

To the Editors:

It was a pleasure to read Harvey B. Lillywhite's account of the inability of some snakes to keep up a sufficient blood supply to the head when they are held vertical ["Snakes, Blood Circulation and Gravity," SCIENTIFIC AMERICAN, December, 1988]. This goes a long way toward explaining the episode in Exod. 7:8–13, where Moses and Aaron stand before Pharaoh and Aaron throws down his rod, which turns into a snake and crawls away. The poor snake may have been carried in by the head, hanging limp in a faint; it revived when it was returned to the horizontal. Evidently the Egyptian priests knew the trick too, for they did the same thing with their rods. Aaron must have had a bigger snake, though, for his rod ate theirs.

HANNE DALGAS CHRISTIANSEN

Copenhagen

To the Editors:

I recently read "Educating Poor Minority Children," by James P. Comer [SCIENTIFIC AMERICAN, November, 1988], and I feel the story is not complete without an account of what happens in a black high school in the inner city. I am a teacher of physics and mathematics, and I have taught minority students for more than 12 years.

In inner-city high schools the decent majority of children are constantly subjected to the disruptions of a few students, so that learning cannot always take place. In our school we took a survey of about 500 students, and almost 40 percent said their classes had been disrupted. (Our survey also found that many of our students believe the neighborhood is unsafe, another source of stress that hampers their education.)

Moreover, problem children disrupt the learning environment all through elementary and secondary school. As a result many black children get to high school unable to read at grade level and very weak in elementary arithmetic, such as decimals and fractions. Subjects such as algebra, trigonometry, physics and chemistry build on elementary mathematics as well as on one another.

The hope that schooling offers to children in poverty is slowly dashed when, in part because of the poor learning environment, these children are unable to do physics, chemistry or advanced mathematics. As a result they are locked into jobs that pay far less than jobs in engineering and science requiring a college degree.

It takes just one disruptive student to destroy the learning that is supposed to go on in a classroom. I have never figured out why black parents do not sue to force schools to get rid of the disruptive influences that interfere with the education of their children. Although we must try to help problem children, we must remember that our resources are limited. What we do have should be concentrated on students who attend regularly and are not disruptive.

STEWART E. BREKKE

Paul Robeson High School
Chicago, Ill.

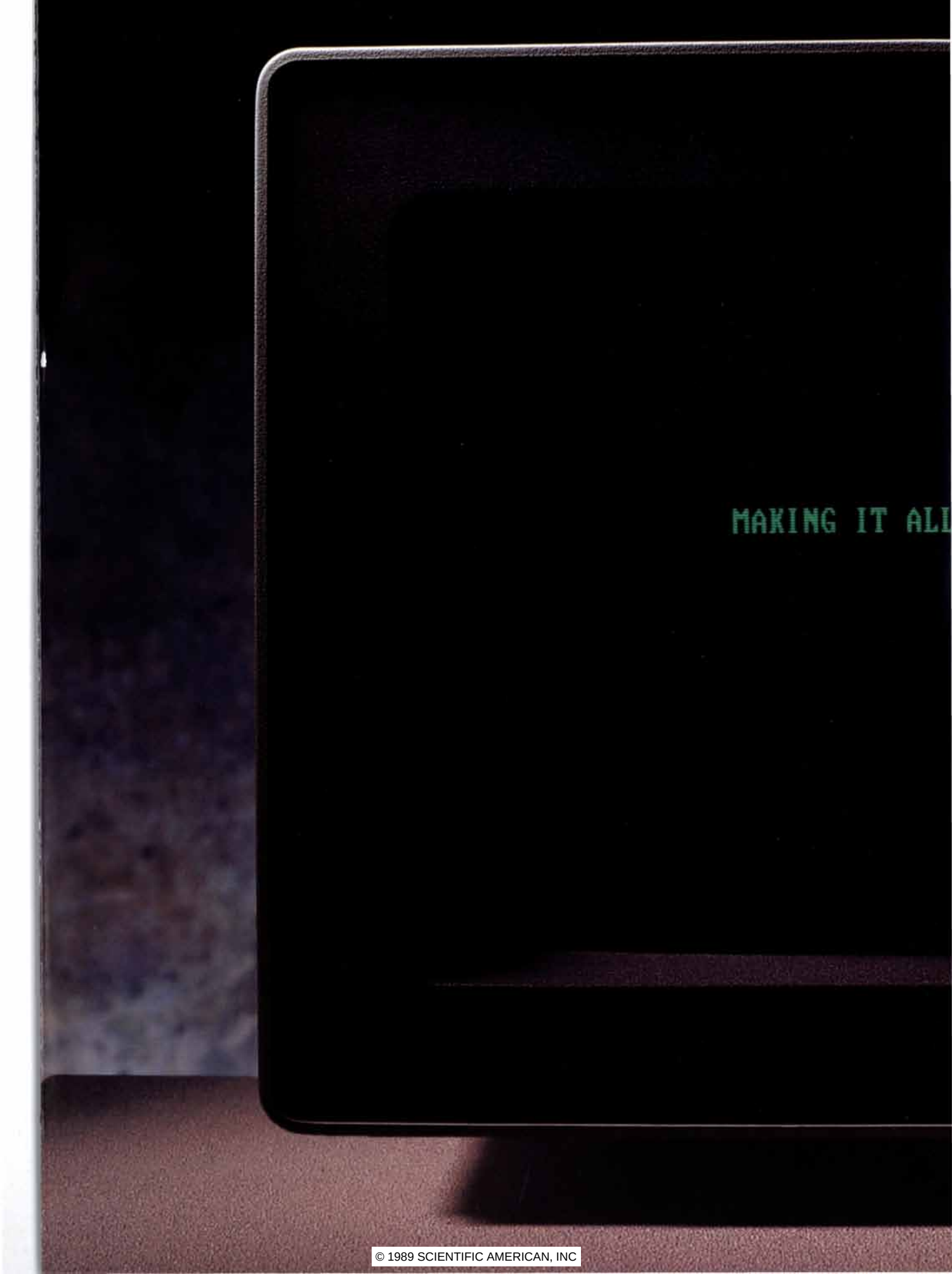
To the Editors:

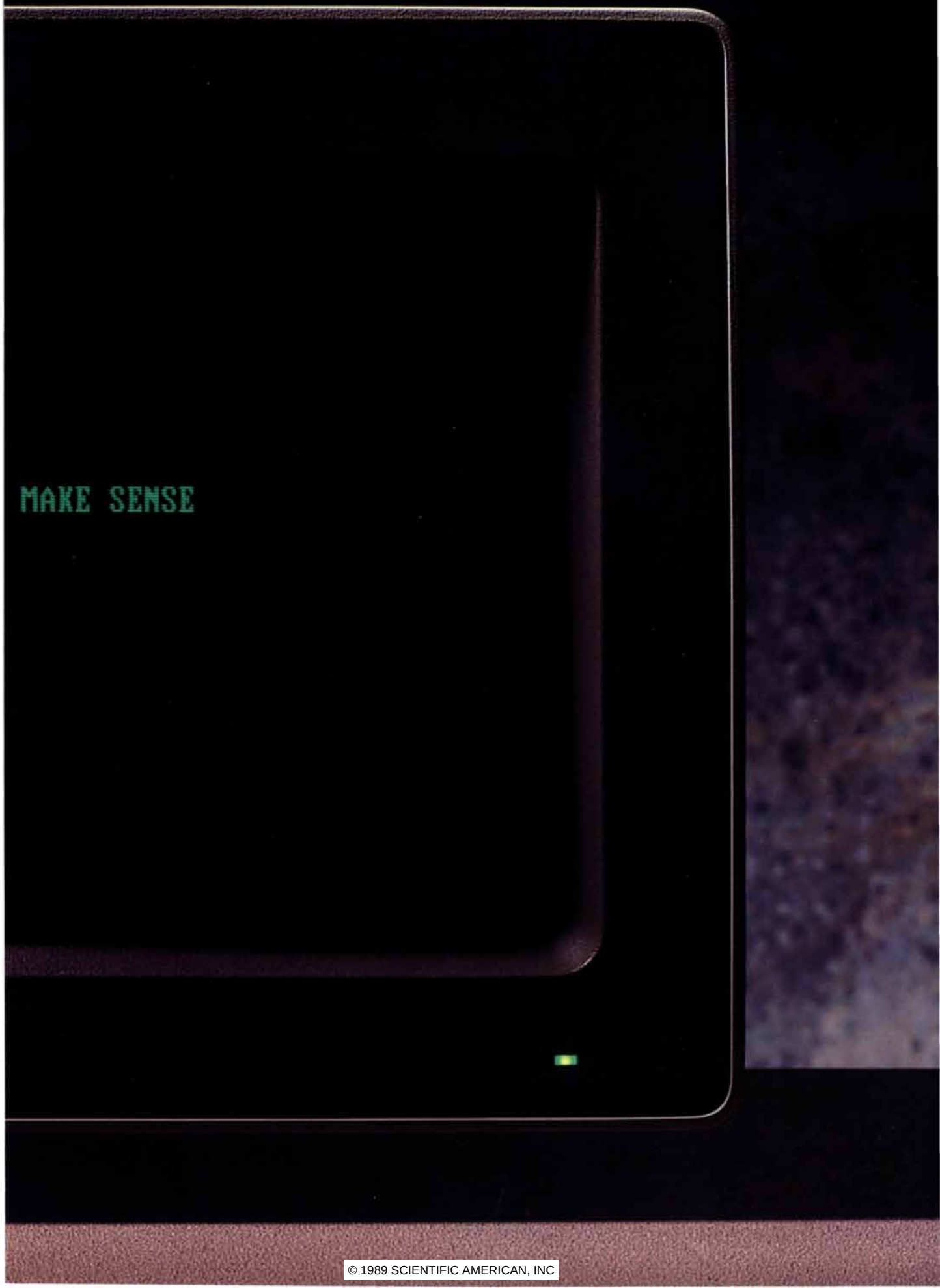
It is true that a significant number of inner-city students are unable to achieve in school or to live lives of dignity and decency in the community because of a small group of highly disruptive students. In the past I have pointed out that a critical mass of students can set a positive or a negative tone in a classroom, school or neighborhood—depending on the skill of the teacher and the presence or absence of other constructive forces. Like Mr. Brekke, I believe we must make learning possible for those students who are prepared to learn. But it is important to remember that most of the disruptive students are the victims of stressful social conditions and inadequate opportunities for personal growth.

Our School Development Program in New Haven, which focuses on improving the social climate of the school and thereby providing personal-growth opportunities, suggests that it is possible to help both students who are prepared to learn *and* those who seem not to be receptive. In New Haven, Chicago, Los Angeles and elsewhere there are examples of middle and high schools that have succeeded in reducing disruptive behavior by improving the overall climate of the school and giving special attention to children who need it.

When we "reclaim" all the young people we can through school programs that support constructive development and growth, the exclusion of recalcitrant students may be justified. But exclusion will probably be less necessary then. Most of the young

The future of personal computing now rests on just one thing.





At Microsoft, first we make it possible.

If you happened to catch the January 1975 issue of *Popular Electronics*, you were one of the lucky few to witness the debut of the personal computer.

Impossible as it seems, a magazine with less than one-tenth the readership of *Time* or *Newsweek* launched a technology race roughly parallel to that of the space program.

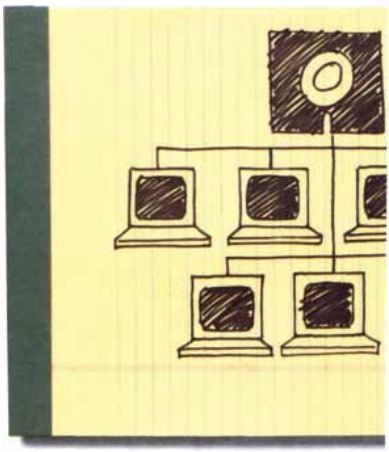
It also launched a company that immediately assumed center stage in the exciting new world of personal computing. The company was Microsoft, and the tenet upon which it was founded was a simple one. To see a computer on every desk and in every home.

To take that rudimentary new contraption that was the early personal computer and turn it into the powerful machine that has literally changed the way we work, required some important steps. The first order of business was to create not simply products, but standards. Microsoft® BASIC became the first universal programming language for the personal computer. And set a standard upon which an industry could grow.

The world wasn't waiting. We were. The PC makes its debut.

Next came what is now the world standard PC operating system, MS-DOS,® developed by us and chosen by IBM for its first personal computers. Today, 20 million machines run on it, and so does a billion-dollar software industry.

And when the Macintosh® was being developed, we were there. That early participation allowed us to write its richest and most important software. These crucial pieces include the powerful



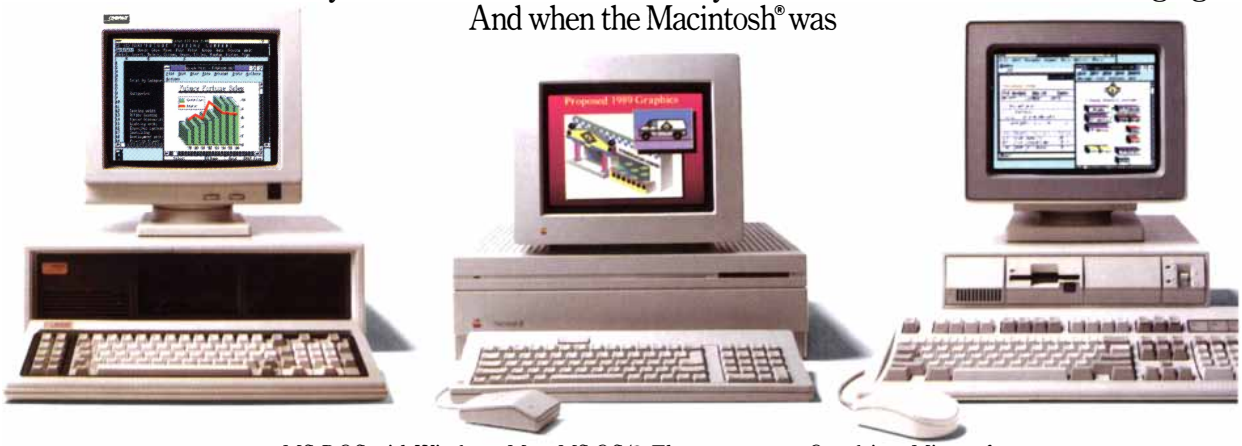
Networking made practical, with softwa

Microsoft Word, the much-applauded Microsoft Excel, and Microsoft Works, the single-solution program for the diverse needs of small business.

As we were contributing to the development of the Mac,® we were also developing a system to put graphics interface technology into the world of IBM® PCs and compatibles.

The introduction of Microsoft Windows in 1985 meant that an easy-to-understand desktop graphical environment now appeared on PCs. Ultimately, this friendly screen will forever replace the cold theater of character-based computing.

But Windows is more than just a useful tool. It is an important technological feat, one that becomes critical to bringing int



MS-DOS with Windows. Mac. MS OS/2. Three ways to go. One driver. Microsoft.

final focus that original Microsoft vision. Through Windows, any number of software applications will seamlessly integrate. Sophisticated spreadsheet programs. Powerful word processors. Interactive databases. All effortlessly accessible.

riving the network.

And in MS® OS/2, the new operating system we developed jointly with IBM, the Windows technology (called Presentation Manager) gets even more exciting. Opening up mega-amounts of power and memory. Opening up your screen to do several tasks at once. And opening up endless possibilities for developers using the Microsoft family of languages.

But all this doesn't end at the desktop. With Microsoft OS/2 LAN (local area network) Manager, it's as easy and natural to work on a network as it is to work alone.

By linking users via software, information can be shared and exchanged by members of a group. Projects are worked on together, instead of bit by bit. And it's amazing how a company communicates once it's joined by electronic mail.

There is no question that the advanced productivity springing from today's personal computer is the direct result of our continued commitment to superior technology.

But even so, that's only half of the equation.

Then w
it prac

The philosophy behind Microsoft includes another, equally important, notion. That all the technology in the world doesn't add up to a hill of beans unless it is practical, useful and, above all, easy.

Easy commands from pull-down menus make window-shopping easy.

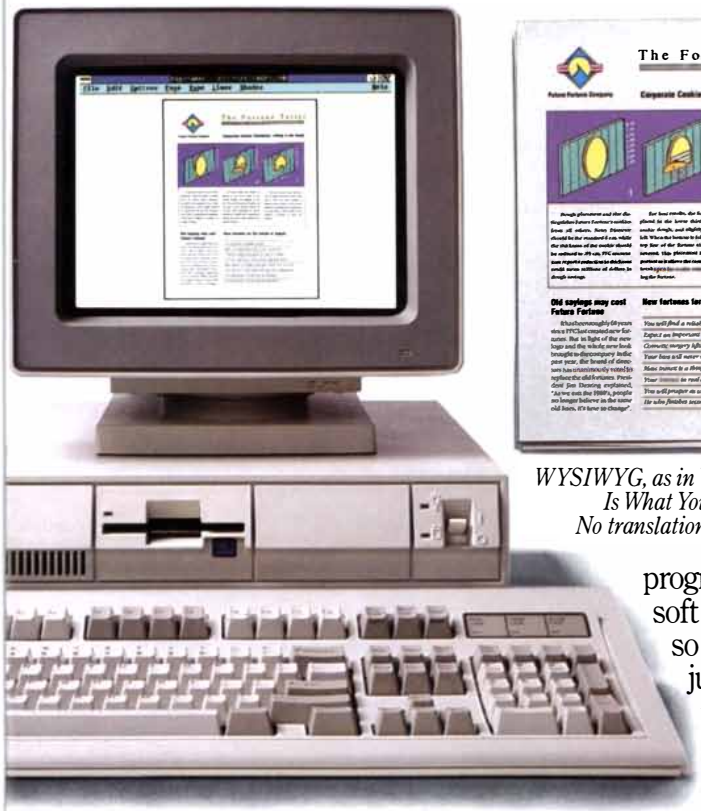
Unless it makes sense.

That's why, whether you're using a Microsoft application on a Mac, an MS-DOS or even an OS/2 machine, it will have a comforting familiarity. Because today's computers share a common software guardian. Microsoft.

Thanks to our groundbreaking work on the graphical interface for the IBM PC and its compatibles, virtually every personal computer can give its user a simpler way to get a lot more done. With a screen that thinks in pictures instead of words, arranged like papers on a desk. Naturally,

e make ctical.

working with pictures makes the work you turn out much more interesting. Which is why the introduction of Microsoft Windows to the IBM PC and compatibles brought with it a whole new category of software with impressive credentials. Like



WYSIWYG, as in What You See Is What You Get. No translation needed.

desktop publishing. And presentation applications that let you create a sophisticated graphics show, from your office instead of the art studio's.

With Windows giving laser printers their marching orders, all manner of documents take on a more finished look. And no matter what application you're using, Windows will take over the job of running your printer.

There is also a hardware complement to graphical applications: the Microsoft Mouse. An unprecedented 1.5 million users have found that a simple point and click eliminates complicated keyboard commands.

Our Windows spreadsheet program, Microsoft Excel, goes so far beyond just simple number-crunching that it has received unparalleled acceptance in corporate America. More powerful than any other, it also easily delivers sophisticated charts, graphs, text and data pulled simultaneously from several sources. And to make it even easier, we built it to graciously accept files and macros from other programs.

The new generation of PCs will run OS/2 with Presentation

A simple point and click replaces mumbo-jumbo keyboard commands.

Manager, taking our graphical screen to even greater heights. By unlocking the capability of these machines, users can easily switch between programs almost instantly. Members of a workgroup can work together on an unlimited number of tasks.

And finally, every kind of program, from spreadsheets to electronic mail to word processing, works in a common way. To the user, learning one is a quick step toward learning them all. To the corporate bottom line, it means far less valuable time and money are spent on training.

But the real practicality of the graphical user interface comes to life when, inevitably, it appears on every computer screen, everywhere. And networking becomes not only possible, but required in this competitive world.

When you think about it, the goal Microsoft set in 1975 of seeing a computer on every desk and in every home seems close at hand. Because, at Microsoft, our fierce pursuit of technology comes with a promise to keep.

Making it all make sense.

Microsoft®
Making it all make sense.™

people who act up and act out in school today want to succeed. We must commit ourselves to new ways of making their success possible.

Important policy changes are needed to reduce the number of families under stress and the number of children growing up underprepared for school. It is not necessary to wait, however, until slow-moving bureaucracies respond to the need. School districts and individual schools can do things to decrease disruptive behavior and improve achievement right now—without suspending or dismissing large numbers of students. After all, where do the suspended students go but into the neighborhood and onto unemployment rolls, sustaining and worsening existing social problems?

JAMES P. COMER

Yale University
New Haven, Conn.

To the Editors:

I very much enjoyed the short piece in "50 and 100 Years Ago" [SCIENTIFIC AMERICAN, January] that was highlighted on the contents page ("1889: A new narrow-gauge railway links mining centers in Colorado").

The railroad link between Ouray and Silverton was never finished, in spite of the enthusiasm of the pioneers. The branch from Silverton got over Red Mountain Pass and as far north as Ironton Park, about 10 or 12 miles short of Ouray. The other branch of the Denver and Rio Grande Railroad stopped in Ouray. Crossing the precipitous canyon of the Uncompahgre River and Red Mountain creek proved too difficult and costly for the railroad builders.

A different link was finally made by Otto Mears, sometimes called "the pathfinder of the San Juan," who built a toll road that became known as "the Million-Dollar Highway." Its principal tollgate was set on a narrow bridge over Bear Creek Falls, which drop several hundred feet from a cliff on the side of the canyon.

Today the narrow-gauge railroad reaches as far as Silverton from the south; the line from the north to Ouray is gone completely. The Million-Dollar Highway still daunts tourists as the only standard automobile link available between the two former mining towns.

MEL GRIFFITHS

University of Denver (emeritus)

This Year, Ou
Wanted to Stop Yo

Jeep® announces our four-wheel anti-lock braking system.

It's called ABS. And it's a Jeep exclusive. It's the only all-wheel full-time braking system available on any four-wheel drive sport utility vehicle. ABS reads the road's surface condition up to fifteen times per second and automatically applies and reapplies brake pressure as needed. It allows drivers the freedom of steering while braking, and turns every stop, even panic stops, into more confident, more controllable stops. That's the beauty of ABS. But it's far from all the beauty of Cherokee Limited. There's the beauty of Selec-Trac®, another Jeep exclusive. It's the only shift-on-the-fly four-wheel drive system that gives you the security of full-time four-wheel drive on any surface, on road or off. And,

r Designers
u in Your Tracks.

there's beauty under the hood in the form of the 177 horsepower Power-Tech Six engine, the most powerful engine in Cherokee's class.

Plus, there's beauty inside with luxurious seating highlighted by lots of genuine leather and surrounded by power everything. And maybe most important, there's the thrill and satisfaction of driving a true American Legend: Jeep. What's more, Jeep Cherokee has a higher resale value than Ford Bronco II and Chevy S-10 Blazer.* And

7/70 now every Jeep comes with Chrysler's exclusive 7-year/70,000-mile Protection Plan.** For further information, call 1-800-JEEP-EAGLE.

Only in a Jeep®

Jeep Eagle

*Source: *Kelley Blue Book*: November/December 1988, analyzing resale values of 1985, '86, '87, and '88 model year vehicles.

**Protects 1989 models' engine and powertrain for 7 years or 70,000 miles and against outerbody rust-through for 7 years or 100,000 miles. See limited warranty at dealer. Deductibles and restrictions apply. Jeep is a registered trademark of Jeep Eagle Corporation. Buckle up for safety.

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

APRIL, 1939: "Broadcasting of regular television programs to the metropolitan area of New York will start this month. The same time has been chosen for marketing the first commercial receivers. Television is bound to have profound effects on our social order. It will affect existing industries and create at least one new one. Someday it may encroach on other entertainment and educational media. Perhaps it will cause a change in the styles of presentation used in other media."

"The Japanese cultured pearl is structurally different from and physically less homogeneous than the natural pearl. Contrary to the widely accepted belief, all specimens examined contain a relatively large mother-of-pearl bead or nucleus and not a 'tiny irritant'—in nearly all cases comprising 80 to 90 percent of the total diameter of the cultured pearl. The process of culturing pearls is, so to speak, the biological analogue of veneering wood or producing a plated metal."

"While the very large, rigid airship is temporarily under a cloud, the much smaller nonrigids (commonly known as 'blimps') are being built by Goodyear for both the Army and the Navy. There is every reason to believe that the blimps will serve a useful purpose in what we must unfortunately call the next war."

"One of the most significant steps in the development of modern highway construction is the consideration now being given by several state legislatures to the building of 'freeways.' The basic essential of the freeway is a right of way to which abutting property does not have access. Because present traffic in built-up areas is slowed to an average of 15 miles an hour owing to stoplights necessary to control it while cross-traffic advances, it is estimated that the rate of travel can be greatly increased and the safety of car drivers can be multiplied many times if intersecting roads are eliminated."

"The Smithsonian Institution is actively engaged in building up a great collection of meteorites for study and encourages people to search for them. The meteorite hunter is advised to look for heavy, dense objects whose surface appears to have been streamlined and to test them either with a horseshoe magnet or a compass."

SCIENTIFIC AMERICAN

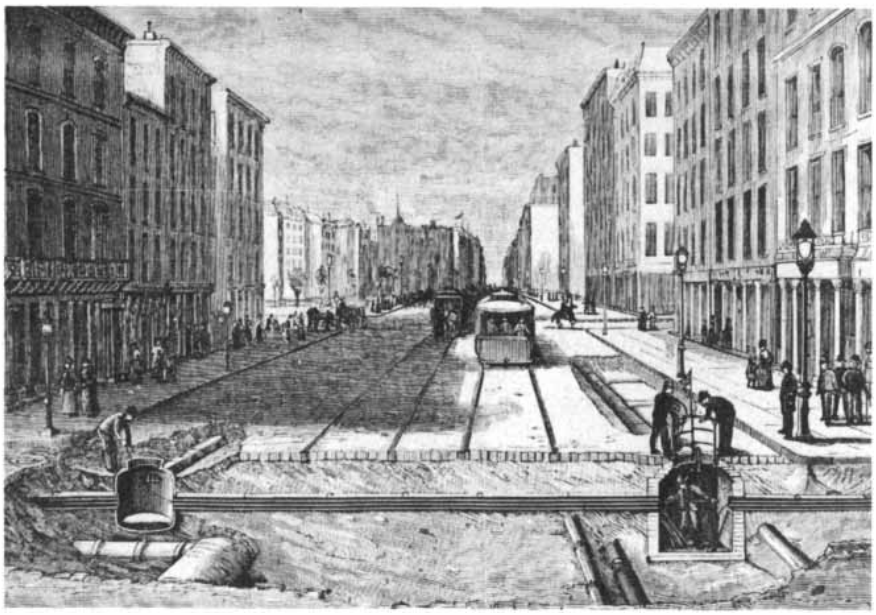
APRIL, 1889: "The opinion of the general public seems to be that oil must, to a great extent at least, take the place of coal as fuel. The advantages are so obvious that it is hard to find a man who does not believe that some method will be found by which a steady, smokeless, economical, and easily controlled flame can be produced from oil, by means of which boilers can be heated. The great public is tired of coal, with its clinkers, smoke, dust, and ashes, and the necessity of constant attention—not to say anything of the regular 'corners,' short weight, and extortionate prices."

"On the question of the right to photograph there is as yet very little judicial opinion. Clearly, every one who sees fit may make pictures of natural scenery. There are objects, however, which cannot be photographed without the consent of the owner. Furthermore, suppose the rejected suitor of the fair Amanda, by means of a detective camera, succeeds in obtaining a picture of his adored one in the very act of being kissed by his hated rival. Would he be allowed to exhibit such a picture? We think that no photograph exposing others to scorn, disgrace, humiliation, or contempt would be tolerated."

"A new method for illuminating from outside some cavities of the body, such as the larynx and nose, has been devised. The instrument used for this purpose is a well-polished (not blackened) glass rod, to one end of which a small electric incandescent glow lamp is attached. The light of the lamp is reflected equally through the whole glass rod to its other end, which can be placed on the skin of the throat. Then the interior of the larynx becomes illuminated sufficiently for laryngoscopy."

"The Eiffel Tower reached its full height on March 31. A newspaper correspondent who went to the top says that the ascent by the staircase took forty minutes, and by elevators it is to be made in five minutes."

"The work of removing the aerial telegraph, telephone, and electric supply lines in New York City, with a view to forcing the electric supply companies to use the subway system in such streets as contain it, has been vigorously prosecuted during the past week. A very impressive feature of the operations is the comparative darkness to which parts of the city are relegated at night. The gas lamps are quite unable to supply sufficient light for the people, who have now been accustomed to electric illumination."



Electric subways of New York City, for telephone, electric and telegraph service

We think our First Class Service is the best in the business.

Too bad more people don't take advantage of it.

TWA's First Class Sleeper Seat.

When you stretch out in TWA's First Class Sleeper Seat, we urge you not to get too comfortable. At least not right away. We wouldn't want you to miss any of the amenities that make our First Class so opulent. And by the way, only TWA offers its luxurious sleeper seat on every domestic and international widebody flight.

TWA's First Class Service.

Before takeoff, you'll be offered complimentary cocktails. Then, on international flights, come the hors d'oeuvres. Everything from American Sturgeon Caviar to Pâté de Canard.

And on all flights you will discover gourmet entrées served to you on china. If you can, save room for dessert or an after-dinner drink.

You won't be disappointed.

TWA's Personal Service Commitment.

TWA is determined to bring you the best personal service in the airline industry. So we've assigned an In-Flight Service Manager to every flight to make sure your trip is as hassle-free and comfortable as possible.

In addition, our Chairman of the Board has put together a Quality Control Team. A group of 30 very picky people who fly TWA, assessing the service they receive, on a checklist of over 100 service items. They report directly back to him with their comments, good and bad, And if anything needs improvement, it gets improved. Fast.

Because at TWA, great service is a top priority.

TODAY'S TWA.
FIND OUT HOW GOOD WE REALLY ARE.℠

TWA

Bruce Sobolov, Director, Equipment Planning and Administration, Election & Survey Unit, CBS News

Laura Gismondi, Account Executive, AT&T Data Systems Group

Bruce Goldberg, Area Technical Manager, AT&T Data Systems Group

Bruce Sobolov of CBS News, Laura Gismondi and Bruce Goldberg, AT&T, savor the afterglow of their own post-election victory. They take us behind the scenes for a glimpse at some of the reasons why CBS was successful on election night.

Skating between terminals put ruts in the rug.

FEBRUARY 15, 1989

AT&T: Afterwards, the critics said CBS was the best, the fastest.

CBS: Right, but we sweated it out for more than a year. With more 20-hour days than I care to remember.

AT&T: Your situation was pretty complicated.

CBS: We were faced with election projections, exit-poll analysis, and other studio programming applications running on IBM hosts.

AT&T: Plus the NewStar system we tied in with our wide-area network, ISN. It's distributed networked computing. Hey, we thrive on this stuff.

CBS: We're impatient around here. Speed is the only way you succeed with election coverage. The first thing we did was provide multi-host access with the 6500 System. Last election, everybody who needed access to two systems used two terminals. Twice the space, twice the cable, additional controllers, added expense, and wasted time.

We had programmers working simultaneously on three host applications, two bisync, one SDLC. They were constantly skating between terminals, wearing ruts in the rug. Now they have access to multiple sessions simultaneously from one terminal.

AT&T: The data moves over twisted pair, the same type wiring the technicians pulled for your System 75 PBX. That made sense.

CBS: An added advantage was having the same dedicated AT&T technicians installing and maintaining our system, providing consistency to my operation.

AT&T: But really, Bruce, why us?

CBS: Your responsiveness. At CBS, we all agreed that what we needed was someone who could deliver it fast, install it, test it, and support it. And you were hungry. You never said, "No, we can't do it." And you never took long to say "yes."

AT&T: You had computer networking problems. Solving them is the house specialty.

CBS: We do distributed computing to the nth degree. Our reporters are all over the country. They call in their results when the precinct closes. Before, we had over a hundred operators standing by, with phones and terminals. That election night we introduced the voice response system running on AT&T PCs.

AT&T: How many calls?

CBS: Thirty, thirty-two calls at once, reporters everywhere having voice response conversations with the IBM host. And all done with the same custom host software we always used. We greatly reduced our cost.

AT&T: The other networks are watching, thinking, "How come CBS has the results already and we don't?"

CBS: It was a good night for us. Now the name of the game is streamlining for 1990. We're talking about a networked computer solution as a gateway into different host systems.

AT&T: With the AT&T Systems already up, running, and in place, we can almost completely automate your survey system.

CBS: That's a real big plus for all of us.

AT&T: Something tells me I've seen that same glint in your eye before. (Laughter)

The CBS Solution:

THE CHALLENGE:

Integrate IBM and DEC host computers and NewStar editorial system. Build an advanced computerized voice response system to speed election-night projections.

THE SOLUTION:

AT&T 6500 Multifunction Communication System with multi-host sync/async 6529 terminals. AT&T CONVERSANT® Voice System for advanced communications running on AT&T WGS computers. AT&T System 75 PBX. AT&T Information Systems Network (ISN), a wide-area network.

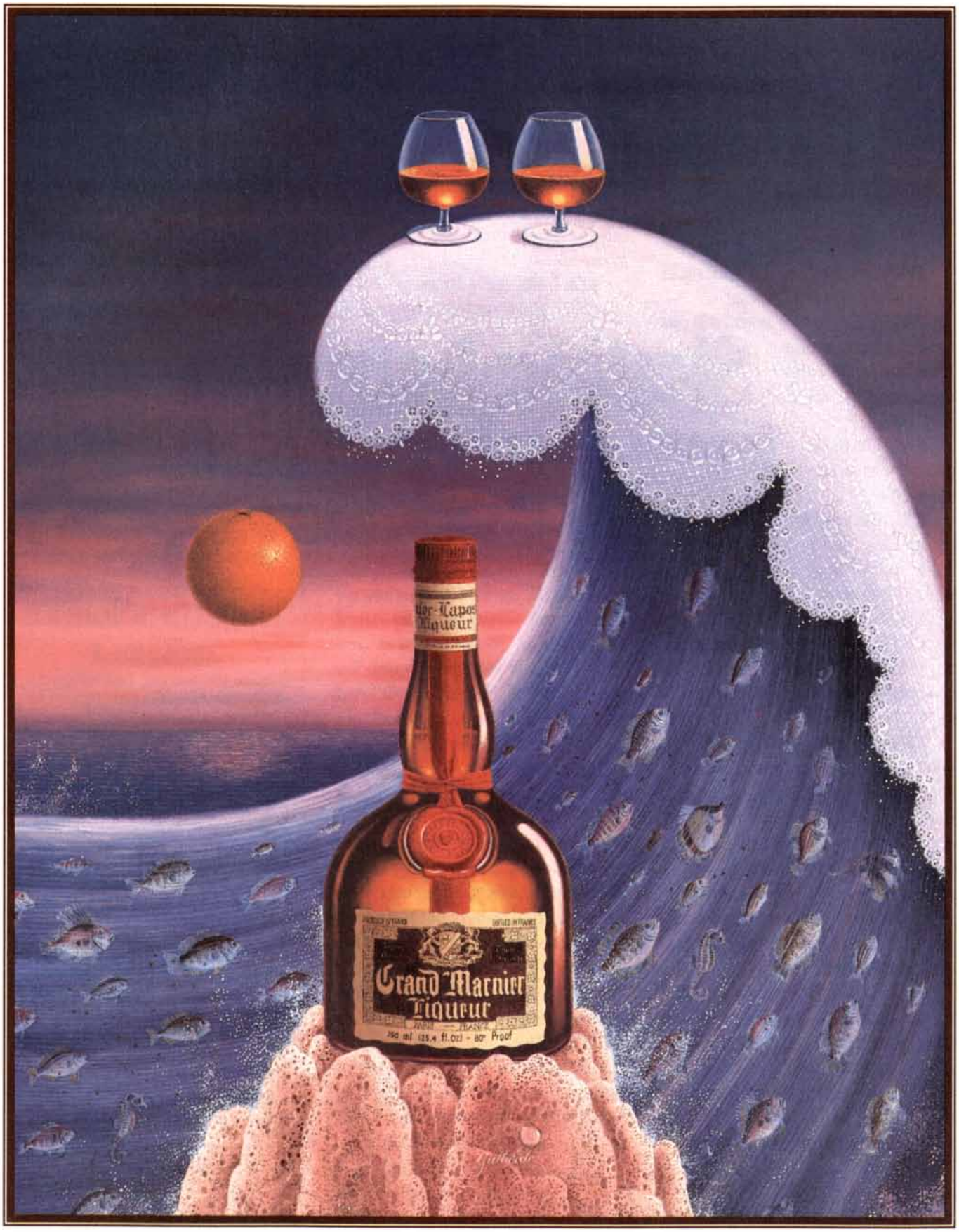
THE RESULT:

CBS News provided fast, accurate election coverage throughout Campaign '88. The *Baltimore Sun* reported that, "CBS was recording results in all sorts of key races faster and with far more authority than either of the other networks."

Call your AT&T Account Executive, Reseller or 1 800 247-1212, Ext. 154.

your computing systems and networking solutions company

AT&T
The right choice.



OCEANS OF GRANDEUR.

For gift delivery of Grand Marnier® Liqueur (except where prohibited by law) call 1-800-CHEER-UP
Product of France. Made with fine cognac brandy 40% alc/vol (80 proof). ©1988 Carillon Importers, Ltd., Teaneck, NJ.

AIDS Counts

Planning for an epidemic whose size is still unknown

A drug addict collapses on the street and is rushed to the nearest emergency room. He is in the last stages of pneumonia. His heart stops. The admitting physician suspects AIDS but does not test the addict's blood for HIV. Rather than resuscitate him so that he can die after weeks of intensive treatment at an uninsured cost of thousands of dollars, the doctor lets him die and records the cause as pneumonia.

This scenario, sketched by a health-care worker on the front lines of AIDS treatment in San Francisco, is being repeated more and more, with variations, in urban centers with large populations of intravenous drug abusers. Rand L. Stoneburner of the New York City Department of Health estimates that perhaps half of the HIV-related deaths among addicts in New York go unrecorded. From 10 to 30 percent of confirmed AIDS cases nationwide are not reported, according to Jeanette K. Stehr-Green of the AIDS Program of the Centers for Disease Control (CDC). The problem may be getting worse as the epidemic increases.

As serious as they are, underdiagnosis and underreporting are probably the least of the problems that plague public-health officials trying to get a handle on the proportions of the AIDS epidemic. Virtually everyone knows that the AIDS virus is spread by sharing needles or through unprotected sexual activity, but no one seems to know much more. Just how risky are these activities? How many people engage in them? How rapidly is the disease—which at first affected chiefly gay men and now also attacks minority drug abusers and their sexual contacts—spreading through the population at large?

A National Research Council (NRC) committee on AIDS, chaired by Lincoln E. Moses, professor of statistics at Stanford University, has called for greatly increased funding to determine the prevalence of HIV and the behaviors that spread it. Unfortunately, because AIDS in the U.S. attacks primarily members of stigmatized groups, gathering more data may be difficult or impossible. Without such information, how will the U.S. (and other countries) plan for treating AIDS victims, target help to those most at risk and determine whether that help is effective?

For example, although it is widely accepted that 50 percent of the gay men in urban centers such as New York and San Francisco are infected with HIV, no one knows what that prevalence implies for gays in the rest of the country, or even how many gay men there are. Last summer New York City health commissioner Stephen C. Joseph provoked public outcry by reducing estimates of the city's gay population by 100,000—and correspondingly reducing the AIDS cases his department must plan for. A report recently published in *Science* estimated on the basis of 1970 data that active gay men constitute not less than 1.4 percent of the male population. In the late 1940's the original Kinsey report estimated the number at more than 4 percent. The small size of the 1970 survey means the percentage could vary between .7 and 2.1. And applying different assumptions to the data, the authors say, could yield a proportion as high as 6.2 percent.

Estimates of the number of intravenous-drug abusers, another high-risk group, range from 1.1 to 1.5 million, of whom perhaps 225,000 were infected with HIV by 1987. Ironically, the Federal Government slashed social-science spending and stopped collecting data on intravenous-drug abusers in 1981, just as the AIDS epidemic was starting.

As a result of these uncertainties, says Charles F. Turner, who directed the NRC study, all anyone can say is that between 500,000 and two million people are currently infected with HIV, with one million a probable figure. Somewhat better estimates can be made by extrapolating backward from the 86,157 AIDS cases known as of February 13, but because of AIDS' long incubation period, this method yields data that are accurate only up to 1985.

The uncertainty in AIDS and HIV numbers makes it impossible to tell whether campaigns to slow the spread of the virus are working, Turner says. The one million people currently infected constitute a horrifying problem, but even more critical is whether next year's count will be 1.01 million or two million. Right now no data are being gathered that could point unequivocally in either direction.

Only one testing program currently under way will help to reduce uncertainty about HIV infection in the general population: the CDC has been testing the blood of newborn infants in selected cities for HIV since 1987; the test yields information about the status of both mother and child. The NRC report says that other CDC-sponsored testing programs—at drug-treatment centers, hospitals, sexually transmitted disease clinics, tuberculosis clinics and family-planning clinics—suffer from potential biases that make generalizing their findings impossible.

Perhaps more promising are plans for a nationwide survey of HIV seroprevalence. It is not easy to establish the degree of trust in a community that makes people willing to give blood samples at a knock on the door, says Daniel G. Horvitz of the Research Triangle Institute, but it can be done. RTI's pilot survey in Washington, D.C., was canceled last summer after it sparked public charges of racism and fears about the misuse of survey data, but a second pilot study, just completed in Pittsburgh, was successful in eliciting cooperation. One key to getting people to provide blood and answer questions about sexual practices, Horvitz says, is guaranteed anonymity: no names are ever attached to blood samples or survey data.

A larger trial survey is slated to begin soon in Dallas, Horvitz says. If the Dallas trial is successful and if Congress approves the funds, a nationwide survey could sample 50,000 people between the ages of 18 and 54 for HIV infection. Such a survey should cost about $25 million and produce a nationwide HIV-infection estimate with a margin of error of about 100,000. Although there might be some systematic errors in the count (much like the undercounting of minorities that plagues the U.S. Census), Turner says, year-to-year trends should still be accurate.

Of course, even the most compelling data will be of little use if politically motivated decisions preclude acting on it. "Overall we're losing the battle" against AIDS, comments Marshall H. Becker, associate dean of the School of Public Health at the University of Michigan. Social scientists understand how to approach problems such as AIDS, Becker points out, but Federal agencies are barred by law (in an amendment introduced by Senator Jesse Helms of North Carolina) from trying to reach high-risk groups in their own idiom and from promoting risk-reducing actions other than total abstention.

Given documented reductions in the spread of AIDS within segments of both the homosexual and the intravenous-drug communities, Becker says, there are clearly some educational and social strategies that work. Becker cautions strongly against believing there is nothing that can be done to change high-risk behaviors simply because "Just say no" campaigns have had such a dismal record. The pattern that must be stopped, he says, is that of doing small intervention studies, publishing a single paper and then dropping the effort because there is no money to continue it. —*Paul Wallich*

Flash Point

U.S. fusion workers chafe under security restrictions

A long-simmering dispute between workers studying inertial-confinement fusion (ICF) and their overseers in the Department of Energy may be coming to a boil. The official goal of the U.S. program—in which lasers or particle beams irradiate pellets of light elements—is to create miniature thermonuclear explosions for probing the effects of nuclear weapons (see "Science and the Citizen," February). In the belief that knowledge of the technology could aid efforts to build hydrogen bombs, the Energy Department has tried to keep much of the research secret.

Yet for years groups in Japan and Europe have been studying and openly discussing aspects of the research that are classified in the U.S. These groups, whose aim is to develop ICF as a source of energy, are also making plans to pool their limited resources in unclassified international projects. Some U.S. investigators, contending that these developments render security concerns moot, have sought permission to discuss their work more openly and to join the international ventures. So far the Energy Department has refused their requests.

More than three years ago a report on ICF by the National Academy of Sciences asserted that classification "impedes progress" and recommended that the Energy Department convene a committee to review its restrictions. Such a committee was finally formed late last year, but Sheldon L. Kahalas, director of the department's ICF division, says the committee "is not anywhere near" completing its review. He also predicts that the department's prohibition on international collaboration will remain in force.

ICF workers are reluctant to criticize the Energy Department publicly for fear they might lose their funding or jobs, but two incidents last year reveal the growing strain over the issue. Last summer Kahalas ordered some U.S. workers not to attend a meeting with Japanese ICF researchers held in Hawaii, according to George H. Miley of the University of Illinois at Urbana-Champaign, who organized the meeting. Later in the year Kahalas rebuked Erik Storm, who heads ICF research at the Lawrence Livermore National Laboratory, for signing a "manifesto" that called for more international cooperation.

William Happer, Jr., of Princeton University, who headed the National Academy's review of ICF, says the U.S. program has already "lost some good people" who felt frustrated by the restrictions and is in danger of losing more. He also disputes the Energy Department's position that the spread of ICF technology would result in nuclear-weapons proliferation. "The engineering involved is completely different," he says.

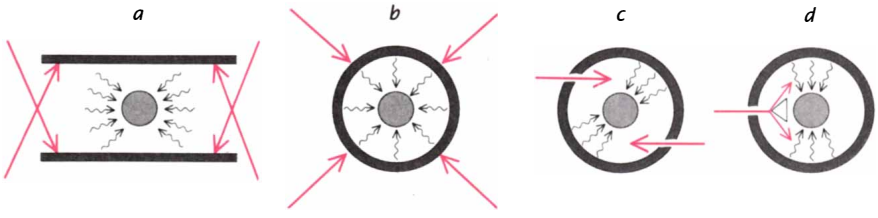
An aspect of ICF research that Energy Department officials consider most sensitive involves a technique called indirect drive, which is loosely based on a design employed in thermonuclear bombs. Rather than being directly irradiated from all sides by beams, the pellet of fuel is placed in a cavity (often called by its German name, *Hohlraum*). When the primary radiation strikes this cavity, it emits X rays that converge on the pellet. In the U.S. only workers at the nuclear-weapons laboratories—the Los Alamos National Laboratory, the Sandia National Laboratories and Lawrence Livermore—are supposed to investigate indirect drive, and they can discuss it publicly only in general terms.

Groups outside the U.S. need not abide by these rules. The February 1984 issue of *Laser Focus*, a magazine published by the PennWell Publishing Co., had a cover story on indirect-drive research done by a group at Osaka University in Japan. Security officials at Livermore considered the cover illustration—which showed an indirect-drive device called a cannonball by the Japanese—so sensitive that they tore it off a copy of the magazine circulating through the laboratory, according to an ICF researcher visiting the laboratory at the time.

Similar articles have appeared more recently. In October the Osaka group delivered a paper on its indirect-drive research at an International Atomic Energy Agency meeting in France. In December *Physical Review A*, a journal of the American Physical Society, carried three papers on indirect-drive experiments by a group at the Max Planck Institute for Quantum Optics in Garching, West Germany.

The Energy Department has tried in vain to persuade these groups to stop publishing such articles. Miley argues that the department should acknowledge that its efforts to keep ICF technology secret have failed and ease its restrictions on U.S. workers. "Right

"Secret" fusion techniques have been published



***INDIRECT-DRIVE CAVITIES** are usually gold spheres a few millimeters wide, although cylinders* **(a)** *have also been tested. Beams of light or particles* **(red arrows)** *enter the cavity through holes or heat it from the outside* **(b)**. *The cavity's lining ablates and emits X rays* **(black arrows)** *that converge on a pellet of fuel. Diagrams* **a–c** *show designs by a group at Osaka University in Japan; diagram* **d** *shows a design by Friedwardt Winterberg of the University of Nevada.*

SCORPIO SCORPIO SCORPIO SCORPIO

LINCOLN-MERCURY DIVISION

IT RUNS IN THE BLACK FOREST. BUT IT'S ALSO NICE TO PEOPLE.

It has road instincts learned in the wild. Yet unlike other German sedans, Scorpio also has an interior designed for the civilized.

Ease into Scorpio's driver's seat and feel the soft available leather gently support your back. Ask your companions how they're enjoying the power-reclining rear seats and over three feet of rear legroom.

Then, for the most comfortable feeling of all, turn the ignition. And put yourself in command of Scorpio's 2.9-liter, autobahn-bred, fuel-injected V-6 and standard anti-lock brakes (ABS).

Scorpio. It's one German touring sedan that treats drivers with utmost respect. Yet still treats passengers with uncommon kindness. **For more information, call 1-800-822-9292.**

GERMAN PERFORMANCE YOU CAN BE COMFORTABLE WITH.

Imported from Germany for select Lincoln-Mercury dealers.

Buckle up—together we can save lives.

now," he says, "we are only harming ourselves in the competition to apply this technology." —*John Horgan*

Talking Tall

Can the Bush Administration deliver more for R&D?

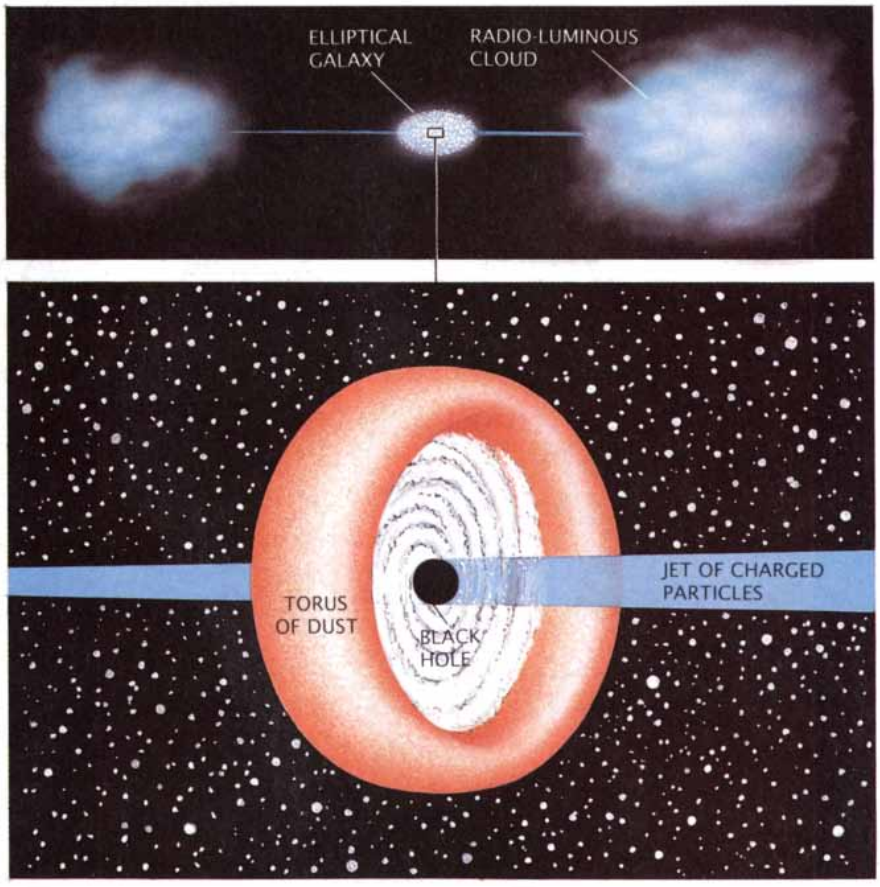
The difficulties with cabinet nominations that have hobbled the new administration could help to explain the longer than expected wait for a nomination of a science adviser to the president. Early signals nonetheless suggest the Administration is attempting to sustain a Reagan-era enthusiasm for big science while at the same time placing a new emphasis on nonmilitary applied research.

The president's proposals for science largely echo those of his predecessor: a 22 percent increase for the National Aeronautics and Space Administration to allow it to develop the space station, $250 million to complete design work on the Superconducting Supercollider and a 13 percent increase for other basic research supported by the Department of Energy and by the National Science Foundation. Bush's budget message reaffirmed the often-repeated goal of doubling the NSF's budget by 1993 and proposed to make permanent the research and experimentation tax credit, an incentive for private-sector domestic research. He also proposed a 6.6 percent increase for basic biomedical research, including $100 million in 1990 for mapping the human genome.

Bush repeated his campaign pledge to designate the director of the Office of Science and Technology Policy (OSTP) an assistant to the president, making the holder of that office a participant in high-level policy-forming groups and giving him or her direct access to the president. Bush also reiterated his pledge to appoint a Council of Science and Technology advisers composed of scientists, engineers and private-sector executives. The budget message undertakes to improve coordination of Federal research with the private sector, an ambition that has been heard before; not heard before was a proposal to strengthen the OSTP by increasing its budget by 28 percent.

Bush succeeds a president who was generally seen to be supportive of science. The impression could be regarded as misleading. Reagan's support was strongly skewed toward military research and big-ticket items. A study by the House of Representatives' budget committee shows that although the Reagan terms were marked by substantial increases in total Federal support for research and development, 84 percent of the $29-billion increase since 1980 was for defense. Support for civilian research and development has actually fallen by 9 percent in real terms since 1980, even though support for all basic research grew by 40 percent in real terms to an estimated $10.5 billion this year.

Applied civilian research and development is the category that has taken cuts since 1980, and that is where Bush appointees seem ready to contemplate changes. Richard G. Darman, director of the Office of Management and Budget, wrote in an answer submitted for his Senate confirmation hearings: "I agree that research and development activities should be increased. Many academic studies have documented that R&D investments lead to longer-term innovation and productivity improvement...." Secretary of Commerce Robert A. Mosbacher repeated the refrain, suggesting that the rapid commercialization of scientific advances "requires closer attention because it appears that this is an area in which we are falling behind our trading partners...." The new administration seems not to be lacking in vision, but how its lofty goals will weigh in the budget scales has yet to be seen. —*Tim Beardsley*



***ROTATING, SUPERMASSIVE BLACK HOLE** encircled by a torus of dust and spewing jets of charged particles from its poles may be at the center of both radio galaxies and quasars. If the torus is edge on to the line of sight, it blocks the region around the black hole; a distant observer perceives a dim elliptical galaxy framed by fully extended radio-luminous jets and clouds: a radio galaxy. If the torus swivels to a position perpendicular to the line of sight, the observer perceives the brilliant maelstrom around the black hole and highly foreshortened jets: a quasar.*

PHYSICAL SCIENCES

Points of View

Quasars and radio galaxies may be two of a kind

Taoist lore has it that the sage Lao-tzu said, "The two are the same, / But after they are produced, they have different names." Might this mystical perception apply

to quasars and radio galaxies? At first glance they appear to be quite unalike. Quasars are highly compact objects that emit vast amounts of both light and radio energy. Radio galaxies are dumbbell-shaped structures that glow primarily in the radio realm and are much larger than any ordinary galaxy. Yet closer scrutiny suggests some relation. Quasars seem to be embedded in starry, elliptical galaxies like those observed at the center of most radio galaxies; indeed, the radio energy from quasars often resembles a truncated radio galaxy.

Such considerations have led Peter Barthel of the University of Groningen in the Netherlands to propose that quasars and radio galaxies are actually the same object seen from different perspectives. His model, which he describes in the *Astrophysical Journal,* pulls together bits of other theories. The central feature is a "powerhouse"—probably a rotating black hole many millions of times more massive than the sun—at the center of a starry galaxy. A thick torus of dust encircles the black hole's equator, Barthel says, "like a tire around a tennis ball." Jets of charged particles stream outward from each pole of the black hole and culminate in two gigantic radio-luminous clouds (the balls of the dumbbell) far beyond the perimeter of the main galaxy.

According to Barthel, the orientation of the dusty torus is crucial. If it is roughly perpendicular to the line of sight from the earth, the torus serves as a frame for the brilliant maelstrom of stars plunging into the black hole. Because this relatively small region outshines the rest of the galaxy, the distant observer sees a point of light: a quasar. The radio-emitting dumbbell, which is aligned with the axis of the torus and therefore with the line of sight from the earth, appears foreshortened; in fact, only the nearer half may be detectable. If the torus is edge on to the earth, however, or tilted at any angle less than about 45 degrees with respect to the line of sight, it blocks the light from the powerhouse. Since the dumbbell is now more or less fully extended, the observer detects a radio galaxy.

This theory explains a major astronomical puzzle: why so many quasars seem to be ejecting matter at so-called superluminal, or faster-than-light, velocities. Acceleration of matter to superluminal speeds is thought to be impossible, but theorists have determined that an effect related to special relativity could create the illusion of superluminal motion—if the matter ejected by the quasar is moving almost directly toward the observer.

Other astronomers are impressed by Barthel's model but suggest that it needs more confirmation. In a recent comment in *Nature,* Virginia Trimble of the University of California at Irvine pointed out that "going around to the other side of a few" radio galaxies and quasars would do the trick. Perhaps there is a mystic—an expert in astral projection, naturally—who would like to volunteer. —*J.H.*

Irish Mist

A tomb in Ireland may be the oldest astronomical structure

Did the world's first astronomers live in Ireland? A new study of an ancient tomb overlooking a valley some 30 miles northwest of Dublin suggests that they might have. The tomb is a tunnel 60 feet long made of dinghy-size rocks buried under a mound of smaller stones and dirt. Carbon dating of charcoal found in the tomb indicates it was built some 5,100 years ago; that makes it at least three centuries older than Stonehenge, which had been considered the oldest known structure aligned with celestial events.

Irish archaeologists have long been intrigued by a local legend that at a certain time of year sunlight penetrated the tomb and illuminated markings in its inner chamber. An excavation of the tomb in the 1960's revealed that in the middle of winter sunlight does occasionally reach the inner chamber, but workers disagreed over whether the alignment was deliberate.

Some 18 months ago Tom P. Ray of the Dublin Institute for Advanced Studies decided to investigate. He found that sunlight reaches the inner chamber during a two-week period in midwinter, and then only briefly each morning. The illumination lasts longest—for about 12 minutes—on the shortest day of the year: at the winter solstice. About four and a half minutes after sunrise a beam enters a window-like aperture built above the entranceway, pierces the narrow passageway and slants halfway across the floor of the domed, high-ceilinged chamber. "It's quite dramatic," Ray says.

Taking into account the periodic wobble of the earth's axis with respect to the sun, Ray determined that 5,100 years ago the effect would have been even more dramatic. At the winter solstice a sunbeam would have entered the chamber precisely at sunrise. It would also have penetrated farther into the chamber. In his report in *Nature* Ray notes that the beam would have indirectly illuminated a carving

Centuries before Stonehenge, Irish farmers erected a structure aligned with a celestial event

***SHAMROCKLIKE CARVING** on a wall of the inner chamber of an Irish tomb would have been indirectly illuminated at daybreak during the winter solstice 5,100 years ago. Photograph courtesy of Conn Brogan, Office of Public Works, Dublin.*

A SPECIAL MICROSCOPE FOR MEDICINE AND BIOLOGY:

AXIOSKOP 20.®

NEW CONVENIENCE AND PRODUCTIVITY FOR EVERY LABORATORY. FROM CARL ZEISS.

Do you demand high standards for your microscopes used for diagnostic specimens and research analyses? Microscopes that deliver the best possible images, easy operation in a relaxed position, and outstanding results, particularly in fluorescence microscopy.

The latest microscope in the family of Zeiss "Pyramids", the Axioskop 20 microscope, meets all these demands in a new, very special way.

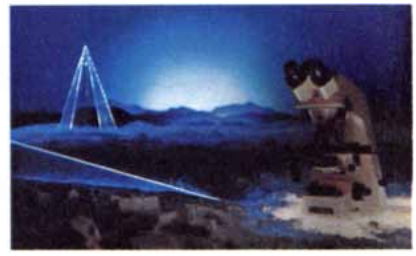
Zeiss ICS optics (Infinity Color-Corrected System) guarantee excellent image quality. All the optical components of the microscope are integrated in a single, optimized system, including the unique, patented Zeiss objective and tube lens combination.

Especially important for routine work: the new Achroplan or Plan-Neofluar objectives for all fluorescence and transmitted-light microscopy. Flatness from edge to edge over the large field of view. Exceptional image contrast and superb color rendition, even for photomicrography.

The high luminous intensity of the new built-in 6 V 20 W illuminator provides a bright image for transmitted-light microscopy.

The SI (System-Integrated) design is user-friendly and guarantees fatigue-free operation for many hours. Changing microscopy techniques is fast and easy. For example, a single move of a slider is all you need to change from fluorescence to transmitted-light microscopy, without any compromise in image quality. When focusing, your hands rest conveniently on the table top – where they naturally belong – and not on the edge of a hand rest.

The Pyramids of Zeiss: The New Geometry for Microscopes.



For the full story and a demonstration, contact your Zeiss dealer. For the name of the nearest dealer, call the Microscopy Division at (914) 681-7755.

See for yourself how easy, how productive it is to work with the new Axioskop 20 microscope from Carl Zeiss.

Carl Zeiss, Inc.

One Zeiss Drive
Thornwood, NY 10594
914 · 747 · 1800

ZEISS
West Germany

What if they all ran on different fuel?

The streets of the world are filled with different types of vehicles. Fortunately, most of them run on gasoline.

IBM is committed to bringing the same consistency to the "fuel" that runs our computers. That "fuel" is, of course, software.

The framework for accomplishing this is called Systems Application Architecture. SAA will enable customers and software vendors to build consistent software applications for different types of IBM computers—from personal computers to mainframes. SAA will make it possible for everyone in an organization to access information regardless of its location.

What's more, all software written to SAA specifications will provide similar screen layouts, menus and terminology. This will make it easier for someone trained to use one type of IBM computer to learn to use others. These are only some of the benefits of SAA—benefits that make it easier for you to "talk" to computers and for computers to "talk" to each other.

IBM is committed to developing state-of-the-art technology. So customers of all sizes can depend on us to offer more state-of-the-art solutions. That's IBM technology at work.

IBM®

on the inner chamber's rear wall, "as mentioned in legend."

The odds that all this happened by accident are—well, astronomical, according to Ray. He suggests that Stone Age farmers, who must have been intimately attuned to the seasons, probably built the tomb. In other respects the tomb's past remains somewhat misty: How was it built? What purpose did it serve? And how did the legend of its secret last for so long? —*J.H.*

It's a Pulsar!

Astronomers may have detected supernova 1987A's flashy core

Imagine a star one and a half times as large as the sun crushed into a sphere only 10 miles across. Electrons, the jittery particles that fluff out ordinary matter, are pressed so tightly against their corresponding nuclei that they merge with the protons and form neutrons. Indeed, the star resembles one gigantic atomic nucleus, held together by gravity rather than the strong nuclear force; a thimbleful scooped from its interior and transported to the earth would weigh about 100 million tons. Now imagine that the object is spinning—so rapidly that it flings charged particles from its surface. These particles, accelerated along magnetic lines of force trillions of times more powerful than the earth's, emit a beam of electromagnetic radiation that spins in time with the object, turning it into a kind of cosmic lighthouse.

An international team of astronomers says it has glimpsed this fantastic object—a rapidly spinning neutron star, or pulsar—nestled deep within the expanding envelope of supernova 1987A. The observations of the supernova, which exploded into public awareness more than two years ago, were made on the night of January 18 at the Cerro Tololo Inter-American Observatory in Chile. Only after the data were analyzed by supercomputers at the Los Alamos National Laboratory did the team realize that the light from the supernova was pulsating.

The failure of two follow-up observations to detect the pulsar raised some concern that the original readings resulted from a defect in the instrumentation. But the workers believe a much likelier explanation is that debris has temporarily blocked the pulsar. "There's nothing marginal about the data," says Jerome A. Kristian of the Mount Wilson and Las Campanas Observatories, one of the observers. "It's a whopping big signal."

Like other observations of supernova 1987A, the sighting of the pulsar provides both a stunning confirmation of the general theory of stellar evolution and an assault on specific aspects of it. Theorists have predicted that supernovas leave behind a neutron star, but none of their models envisioned a remnant like this one. First, its pulse rate indicates that it is spinning nearly 2,000 times per second. That is more than twice as fast as any other known pulsar; a point on its equator would move at about one third the speed of light.

The pulsar's speed also seems to contradict a recent proposal that the fastest-spinning pulsars begin as binary systems and come up to speed only over long periods; matter falling from the companion onto the pulsar supposedly whips it to higher speeds. The discovery last year of a very rapid pulsar that is periodically eclipsed by a companion seemed to provide support for this proposal (see "Science and the Citizen," July, 1988).

Subtle Doppler shifts in the data present another puzzle. They suggest that every eight hours or so the source of the flashes moves away from the earth and then back to its original position. Kristian notes that the pulsar's axis might be wobbling slowly with respect to the earth. Another possibility, he says, is that the pulsar is orbiting a companion about as massive as Jupiter and less than a million miles distant. The companion (if it exists) could not predate the supernova, since its orbit falls within the radius of the progenitor star; it could represent a chunk of matter ejected from the pulsar or dragged out of the surrounding cloud of debris.

"This is exciting because it's so strange," comments Stanford E. Woosley of the University of California at Santa Cruz, a supernova theorist. "If it holds up, it will be the astronomical event of the year." —*J.H.*

Unsuperconductivity

An unholey variant poses some crucial questions

Theoreticians have been having a difficult time explaining how superconductivity occurs at such high temperatures in the new ceramic superconductors. For one thing, the classic theory of superconductivity is based on metals. Furthermore, experiments with the new materials are technically demanding, so that theorists

The ultimate in sound...
The ultimate in savings...

GET 6 COMPACT DISCS FOR THE PRICE OF 1

...with nothing more to buy ever!

Bon Jovi: New Jersey 100516

100586

200596

100711

100715. R.E.M.: Green Orange Crush, Pop Song 89, etc. (Warner Bros.)

100602. Elton John: Reg Strikes Back • Elton's 22nd gold album! (MCA)

264134. D.J. Jazzy Jeff & The Fresh Prince: He's The D.J., I'm The Rapper (Jive)

200478. Metallica: And Justice For All • #1 Speed metal band! (Elektra)

223559. The Beach Boys: Endless Summer • 21 timeless hits! (Capitol)

100532. Diane Schuur: Talkin' 'Bout You • Cry Me A River, etc. (GRP)

100603

115436

100707

105392. Pops In Space John Williams & The Boston Pops. Music from Star Wars, The Empire Strikes Back, more. (Philips *DIGITAL*)

153582. Tracy Chapman Fast Car, Talkin' 'Bout A Revolution, etc. (Elektra)

164165. Bobby McFerrin: Simple Pleasures • Don't Worry Be Happy, etc. (EMI)

244006. Simon & Garfunkel: The Concert In Central Park • All-time classics! (Warner Bros.)

125179. Tchaikovsky, 1812 Overture; Romeo And Juliet; Nutcracker Suite Chicago Symph. Orch./ Solti. (London *DIGITAL*)

100459. Cocktail (Original Soundtrack) • (Elektra)

100604. Heifetz: The Decca Masters, Vol. 1 Golliwog's Cakewalk, Clair de lune, many more. (MCA)

100035. Robert Palmer: Heavy Nova • Simply Irresistible, More Than Ever, etc. (EMI)

123385. The Best Of Eric Clapton: Time Pieces (Polydor)

100579. K. T. Oslin: This Woman • Hold Me, Money, title song, more. (RCA)

100470. Vangelis: Direct New Age Meditations, The Motion Of Stars, The Will Of The Wind, etc. (Arista)

153983. Charlie Parker: Compact Jazz • Now's The Time, NightAnd Day. (Verve)

154135. The Best Of Steely Dan: Decade 14 hits. (MCA)

104871. Supertramp: Classics (14 Greatest Hits) • The Logical Song, Give A Little Bit, more. (A&M)

144578. The Judds': Greatest Hits • (RCA)

115356. Pinnock: Vivaldi, The 4 Seasons • Simon Standage, violin; etc. (Archiv *DIGITAL*)

114780. Cinderella: Long Cold Winter • Gypsy Road, Don't Know What You Got, more. (Mercury)

134347. Huey Lewis: Small World • (Chrysalis)

173406. Jazz CD Sampler Over 67 minutes of jazz. (Polygram)

100467. Beethoven, Symphony No. 9 (Choral) London Classical Players/ Norrington. (Angel *DIGITAL*)

123721. Jimmy Page: Outrider • Led Zeppelin guitarist's solo flight! (Geffen)

134321. Led Zeppelin: Houses Of The Holy (Atlantic)

153606. INXS: Kick • Need You Tonight, Devil Inside, etc. (Atlantic)

100517. Phil Collins: Buster/Soundtrack • Groovy Kind of Love, Two Hearts, etc. (Atlantic)

134420. John Cougar Mellencamp: The Lonesome Jubilee • Paper In Fire, more. (Mercury).

100008. Randy Travis: Old 8x10 • Honky Tonk Moon, more. (Warner Bros.)

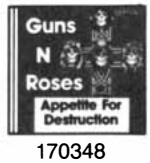
100927

115457

100713

182522

115311



170348

270106. An Evening With Louis Armstrong • (GNP Crescendo)

262889. Virgil Fox: The Digital Fox • Organ music by Bach, others. (Bainbridge *DIGITAL*)

209468. Perlman: Brahms, Violin Sonatas (Angel *DIGITAL*)

120768. 20 Great Love Songs Of The 50s & 60s, Vol. 1 • (Laurie)

153740. Genesis: Invisible Touch • (Atlantic)

163579. Andrés Segovia Plays Rodrigo, Ponce & Torroba • Fantasia para un Gentilhombre, Concierto del Sur, Castles Of Spain. (MCA)

100679. Steve Earle: Copperhead Road • (UNI)

134267. Marriner: Mozart, Overtures • Academy of St. Martin. (Angel *DIGITAL*)

125360. By Request...The Best Of John Williams & The Boston Pops • Olympic Fanfare, Liberty Fanfare, more. (Philips *DIGITAL*)

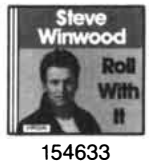
134627. Classic Old & Gold, Vol. 1 • 20 hits! (Laurie)

104857. Benny Goodman: Sing, Sing, Sing • (RCA)

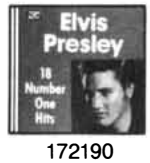
115306. Pinnock: Handel, Water Music • The English Concert. "A winner."—Ovation (Archiv *DIGITAL*)

173233

100714



154633



172190

SAVE 50%

INSTANT HALF-PRICE BONUS PLAN

Unlike other clubs, you get 50%-off Bonus Savings with every CD you buy at regular Club prices, effective with *your first* full-price purchase!

START NOW WITH 4 COMPACT DISCS!

Yes, pick any 4 compact discs shown here! You need buy just one selection at regular Club prices (usually $14.98–$15.98)...and take up to one full year to do it. Then you can choose another CD free as a bonus. That's 6 compact discs for the price of 1 and there's nothing more to buy...ever! (Shipping & handling added to each shipment.)

HOW THE CLUB OPERATES

You select from hundreds of exciting compact discs described in the Club's magazine and mailed to you approximately every 3 weeks (19 times a year). Each issue highlights a Featured Selection in your preferred music category, plus alternate selections. If you'd like the Featured Selection, do nothing. It will be sent to you automatically. If you'd prefer an alternate selection, or none at all, just return the card enclosed with each issue of your magazine by the date specified on the card. You will have at least 10 days to decide, or you may return your Featured Selection at our expense for full credit. Cancel your membership at any time after completing your membership agreement, simply by writing to us.

FREE 10-DAY TRIAL

Listen to your 4 introductory selections for a full 10 days. If not satisfied, return them with no further obligation. You send no money now, so complete the coupon and mail it today.

BMG Compact Disc Club
6550 E. 30th St., Indianapolis, IN 46219-1194. CD692

COMPACT disc DIGITAL AUDIO

YOUR SAVINGS START HERE

Mail to: BMG Compact Disc Club
P.O. Box 91412/Indianapolis, IN 46291

YES, please accept my membership in the BMG Compact Disc Club and send me the four Compact Discs I've indicated here, billing me for just shipping and handling under the terms of this ad. I need buy just 1 CD at regular Club prices during the next year—after which I can choose a FREE bonus CD! That's 6 for the price of 1...with nothing more to buy ever! (Shipping & handling is added to each shipment.)

RUSH ME THESE 4 CDs (Indicate by number):

D	D	D	D

I am most interested in the musical category checked here—but I may always feel free to choose from any (check one only):
1 ☐ **EASY LISTENING** (Instrumental/Vocal Moods)
2 ☐ **COUNTRY** 3 ☐ **HARD ROCK**
4 ☐ **POP/SOFT ROCK** 5 ☐ **CLASSICAL**

☐ MR.
☐ MRS.
☐ MISS First Name Initial Last Name (PLEASE PRINT)

Address Apt.

City State Zip

BMG COMPACT DISC CLUB

Telephone ( ) Area Code

Signature YDT65 (BU)

Limited to new members, continental USA only. Current CD Club members not eligible for this offer. One membership per family. We reserve the right to request additional information or reject any application. Local taxes, if any, will be added.

have relatively few undisputed facts with which to work. As a result imagination has free rein and theories have proliferated, splitting into new versions every time they encounter an inconvenient fact. Nevertheless, facts are accumulating, and a new kind of high-temperature superconductor recently described by Japanese investigators promises to generate some new challenges.

With few exceptions the ceramic superconductors discovered up to now contain distinctive layers of copper and oxygen atoms. In these materials other adjacent layers of atoms take some electrons away from the copper-oxygen layers, leaving behind positively charged "holes" that somehow become the entities creating a resistanceless flow of current. The materials could therefore be called hole superconductors.

The new material, first made by Y. Tokura and his colleagues at the University of Tokyo and described in *Nature,* is different: instead of taking electrons away from the copper-oxygen layers, the adjacent layers are doped with a metal, cerium, that donates electrons. Hence electrons, rather than holes, carry the current. Yet in spite of this fundamental difference the new material still shows superconductivity at the respectably high temperature of 24 degrees Kelvin.

John M. Tranquada and others at the Brookhaven National Laboratory have thickened the plot still further. They report in *Nature* that they made the cerium-doped electron superconductor according to Tokura's recipe and observed how it absorbs X rays. The pattern of absorption reveals the electronic structure of the material's atoms. Tranquada found that whereas the holes of a hole superconductor are linked to oxygen atoms in the copper-oxygen layers, the electrons in the electron superconductor are associated with copper atoms.

The two teams have thus presented theorists with some interesting questions. Is the mechanism of high-temperature superconductivity the same in hole superconductors and electron superconductors? If it is, does it make any difference that the holes are associated with oxygen atoms whereas the electrons are linked to copper atoms?

At first glance the second question could be troublesome for theories that treat the charge carriers in superconductors as if they all occupy the same energy "band" regardless of whether they come from the copper or the oxygen atoms. Several well-known theoretical models fall into this "sin-

ANY SET FOR ONLY $4.95

Choose one of these three sets as your introduction to the

Library of Computer and Information Sciences

You simply agree to buy 3 more books—at handsome discounts—within the next 12 months.

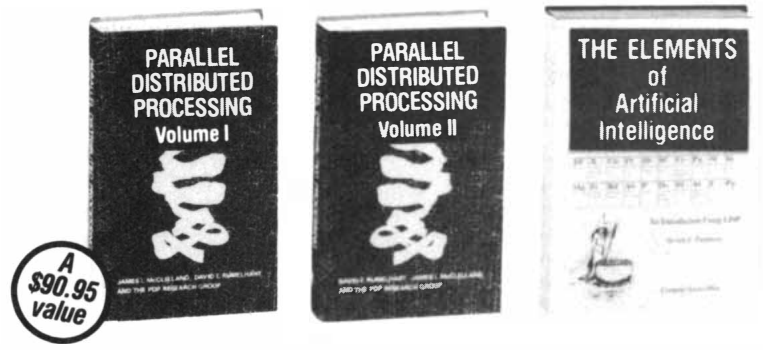
Values to $93.00

1. INTRODUCTION TO C LIBRARY

- Learn to write, compile, and execute programs in ANSI C
- Use previously undocumented tips, tricks, and short-cuts to create powerful, portable programs in any environment
- Apply such advanced concepts as character testing, string handling, encryption, random number generation, and more

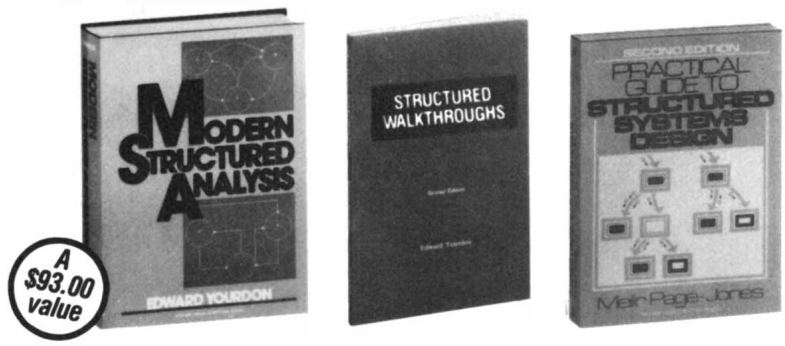
2. LIBRARY OF ARTIFICIAL INTELLIGENCE

- Understand the newest principles in artificial intelligence and neural networking
- Develop natural language, expert system, computer vision, and probabilistic reasoning software
- Apply the latest knowledge in human cognition to create programs that understand colloquial speech, distinguish elements in a scene, and more



3. SYSTEMS DESIGN LIBRARY

- Use the updated Yourdon Methodology for prototyping, data modeling, artificial intelligence, and more
- Develop realistic designs with proven structured analysis tools
- Test, debug, and improve your software through peer review



MEMBERSHIP BENEFITS • In addition to getting 1 of 3 sets for only $4.95 when you join, you keep saving substantially on the books you buy. • Also, you will immediately become eligible to participate in our Bonus Book Plan, with savings of 65% off the publishers' prices • At 3–4 week intervals (16 times per year), you will receive the Library of Computer and Information Sciences News, describing the coming Main Selection and Alternate Selections, together with a dated reply card. • If you want the Main Selection, do nothing, and it will be sent to you automatically. • If you prefer another selection, or no book at all, simply indicate your choice on the card and return it by the date specified. • You will have at least 10 days to decide. If, because of late mail delivery of the News, you should receive a book you do not want, we guarantee return postage.

THE LIBRARY OF COMPUTER AND INFORMATION SCIENCES is the oldest, largest book club especially designed for computer professionals. In the incredibly fast-moving world of data processing, where up-to-the-moment knowledge is essential, we make it easy to keep totally informed on all areas of the information sciences. What's more, our selections offer you discounts of up to 30% or more off publisher's prices.

If reply card is missing, please write to The Library of Computer and Information Sciences, Dept. 7-FL9, Riverside, NJ 08075, for membership information and an application.

Scientific American 4/89

gle band" category, including those proposed by Nobel laureates Philip W. Anderson of Princeton University and J. Robert Schrieffer of the University of California at Santa Barbara.

Victor J. Emery, also of Brookhaven, is inclined to think a more complicated "two-band" model is called for, at least for hole superconductivity. On the other hand, single-band theorists could fight back by justifying their assumption that the origins of charge carriers are not important for superconductivity.

Tranquada emphasizes that his results represent only a start toward describing the electronic structure of the electron superconductor. Many more measurements will be necessary before a clear idea of the mechanism of high-temperature superconductivity emerges. —*T.M.B.*

Arctic Angst

No Arctic "ozone hole," but conditions could lead to one

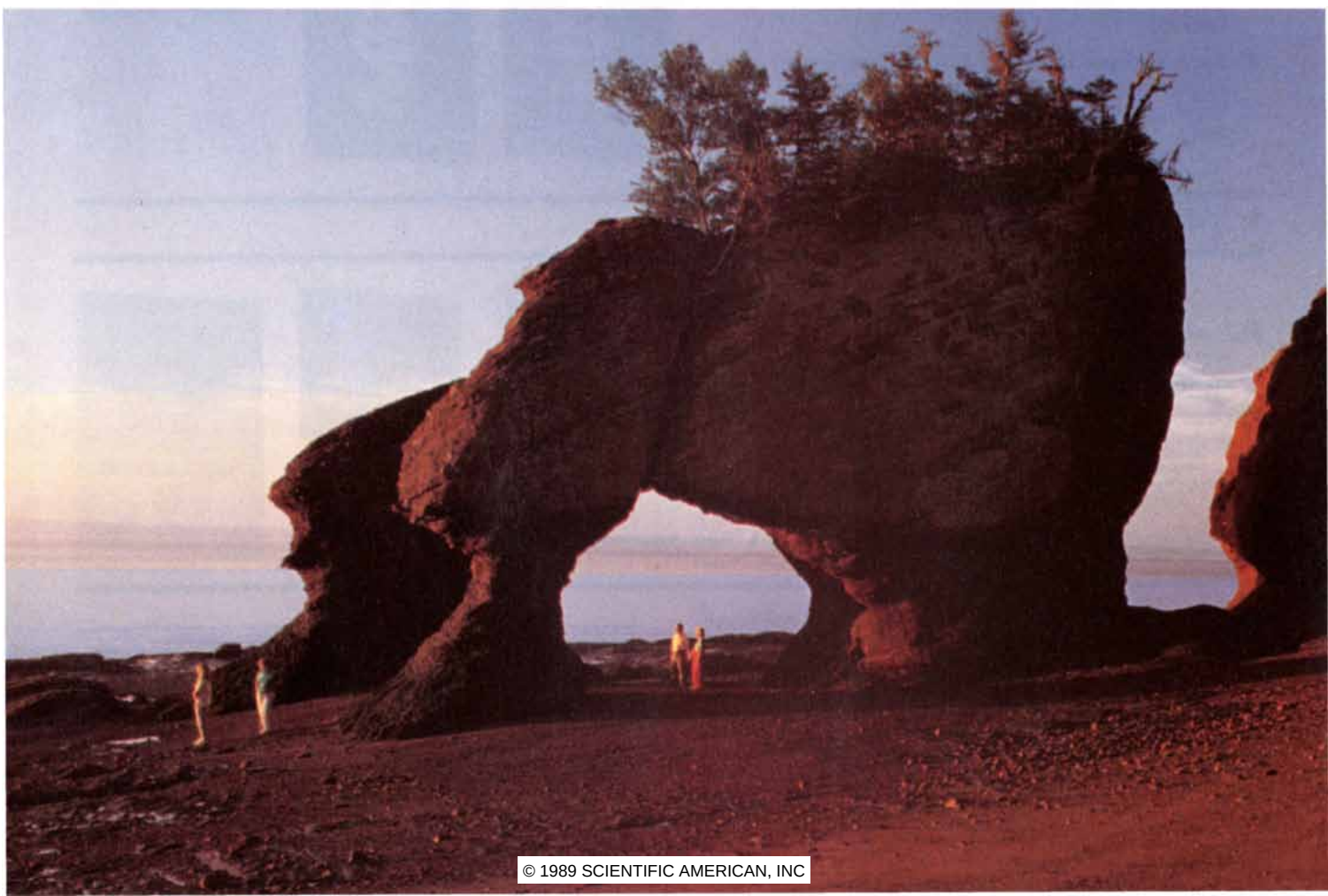
The discovery of an "ozone hole" over the Antarctic has provided dramatic proof that the layer of ozone in the earth's upper atmosphere, which protects living organisms from solar ultraviolet radiation, is dangerously vulnerable. Does a similar hole exist over the Arctic region? Preliminary results from an expedition that examined the Arctic stratosphere in January and February indicate the answer is negative. But "highly perturbed" chlorine chemistry was observed. The conditions, according to expedition scientists, could lead to significant damage to the Arctic ozone layer.

In northern latitudes the winter ozone layer has declined by about 6 percent since 1970, a rate three times greater than standard theories predict. The so-called Airborne Arctic Stratospheric Expedition was aimed at finding out whether elevated levels of certain chlorine compounds, which are responsible for producing the ozone hole over Antarctica every year, also occur in the Arctic. Atmospheric chlorine compounds result mostly from the breakdown of chlorofluorocarbons, the widely used synthetic refrigerants and solvents. The expedition was fielded by the National Aeronautics and Space Administration and the National Oceanic and Atmospheric Administration (NOAA), helped by workers from Norway, Great Britain, the Soviet Union and West Germany.

During the expedition two specially equipped research aircraft made a total of 28 flights northward from Stavanger, Norway, during a period of six weeks that began in early January. The two aircraft, a DC-8 and a high-altitude modified spy plane known as an ER-2, were equipped with a variety of analytical instruments.

In the Antarctic, chemical reactions that take place on the minute particles making up Polar Stratospheric Clouds (PSC's) increase the concentrations of chlorine monoxide and chlorine dioxide; these active substances lead to destruction of ozone, which occurs when sunlight arrives after winter. PSC's are less common in the Arctic, since the Arctic stratosphere is warmer. Nevertheless, Adrian F. Tuck of the NOAA, the expedition's project scientist, said investigators were startled to find chlorine monoxide and chlorine dioxide in abundances more than 50 times higher than normal over a wide range of altitudes. Nitrogen compounds that inhibit ozone destruction were depleted.

These conditions are similar to conditions seen in the Antarctic. Although the preliminary results showed only equivocal indications of actual ozone loss, Robert T. Watson of NASA, the chief program scientist, said ozone loss would be expected if the mass of



cold air over the Arctic known as the Arctic polar vortex persisted into spring, when the sun rises after the polar winter.

The new results will increase pressure to strengthen the Montreal protocol, which came into force this year and should halve emissions of chlorofluorocarbons by 1998. A scientific reappraisal of the protocol is due to be completed in August. Watson points out that the potential for drastic depletion of Arctic ozone is clear and that under the existing protocol the concentration of stratospheric chlorine will still double. "I think there's a rather strong message there for policymakers," he says. —*T.M.B.*

TECHNOLOGY

Dirty Business

The Department of Energy still thinks it can clean up its mess

Pollution at the Department of Energy's weapons-production facilities has been recognized as one of the nation's most pressing environmental problems. In spite of severe budgetary pressures, President Bush proposes spending $401 million in 1990 on studies to determine what cleanup actions are necessary, an increase of $242 million over this year's figure. The Energy Department has tentatively estimated that total cleanup costs might reach $64 billion.

If the funds are made available, is the department up to the job? Officials concede that the task will certainly take decades. Indeed, in spite of the political pressure created by recent publicity, the department has no legal authority to initiate a broad cleanup immediately, according to the department's Walter A. Frankhauser. Indeed, Frankhauser believes the news media have exaggerated the urgency and scale of the problems. In any case, the department is still only at the first stage of characterizing the "remediation" task ahead: "We are in no sense close to recommending what will be required," Frankhauser remarks. He suggests that the timetable for action will be driven largely by legal considerations outside the department's control.

Cleanup operations are only part of a broader program of safety and waste-management improvements affecting dozens of sites across the country. One of the main tasks is converting high-level radioactive liquid waste leaking from storage tanks at the Hanford site in Washington and the Savannah River plant in South Carolina into solid form for disposal in a deep repository. For that purpose the department has begun to build three vitrification plants, which capture radioactive material in glassy blocks. It is also building installations that will cement low-level waste into blocks that can be stored in special vaults; now much low-level waste is lying buried in cardboard boxes in shallow trenches. Lack of information compounds the problem. In most cases there are no good records of what was dumped where.

Perhaps the greatest challenge—judged by sheer size—will be contaminated soil. In some places radioactive isotopes including plutonium, as well as toxic organic solvents, have for decades been simply poured into onsite burial pits and seepage basins on the assumption—since shown to be faulty—that dilution would render them harmless. Now the materials are spreading. At the Fernald plant in Ohio pollution has seeped beyond the site boundary. The volume of polluted earth is much too large to be dug up and put somewhere: there is no place big enough. Therefore many of the new concepts in waste management

Decontamination techniques have been developed. Can the Government exploit them effectively?

***VITRIFICATION PROCESS** being developed by workers at the Battelle Pacific Northwest Laboratories converts soil into an obsidianlike substance.*

stress reducing the volume of the contaminated material. Two leading candidate technologies for treating polluted soil, in situ vitrification and the plasma centrifugal reactor, have been adopted by the Geosafe Corporation of Kirkland, Wash. (a division of the Battelle Memorial Institute), and by Retech, Inc., of Ukiah, Calif.

In situ vitrification is accomplished by inserting four large electrodes into the ground. They pass enormous electric currents through the soil until it melts from the surface down to the desired depth. This happens at about 1,600 degrees Celsius, so that all organic materials conveniently burn off. The radioactive elements are captured in an impermeable glassy mass that can be left in place and covered over.

In a plasma centrifugal reactor, soil contained in a revolving drum moves slowly past a plasma torch (an electric arc) until it melts. When the drum stops revolving, the slag runs out. By adding the right materials to it one can render the slag hard and impermeable, according to Alan D. Donaldson of the Idaho National Engineering Laboratory. Small pilot models have been built; in two years the laboratory expects to have a centrifugal reactor in service that can melt a ton or more of earth in an hour.

The Energy Department has many other technological fixes in development for decontaminating wastes of all kinds as well as soil. But even if the fixes work, critics wonder about the department's fitness to manage a cleanup. After all, they point out, the department and its predecessors created the mess in the first place.

Department officials parry by noting that standards of acceptable practice were different during the Cold War. Yet a staff member of Senator John Glenn's Committee on Governmental Affairs argues that the department has still not set proper environmental standards for its sites and that it still makes environmental safety and health decisions on a local level. Glenn plans to introduce legislation that would mandate some external scrutiny of the department's efforts. —*T.M.B.*

Deep Thought

Chess-playing computers are beating grand masters

It was inevitable. Just as John Henry, the hammer-wielding railroad laborer of legend, finally succumbed to a steam-powered drill, so have the world's best chess players begun to bow before computers.

Not long ago chess-playing computers were considered unimaginative hacks, capable of beating only middling human players. Then last September Hitech, a computer built at Carnegie-Mellon University, beat former U.S. champion Arnold Denker in an exhibition match, marking the first time a machine had ever beaten a grand master.

Two months later a computer created by another team at Carnegie-Mellon eclipsed Hitech's feat. Called Deep Thought, it tied for first place in a tournament in which some of the best players in the world competed. The computer beat Bent Larsen, a grand master who ranks among the world's top 50 players; it also racked up enough points in the tournament to qualify as an international master.

The latter accomplishment has won Deep Thought a $10,000 "intermediate prize" drawn from a fund created nine years ago by the noted computer scientist Edward Fredkin. The computer's creators, who include four graduate students and one recent Ph.D., do not intend to rest on their laurels. They plan to increase from two to eight the number of custom-made microprocessors with which Deep Thought ponders its moves. Some experts think this bigger brain might enable the computer—after a few more years of practice, perhaps—to win the ultimate Fredkin prize of $100,000 by beating the world's top-ranked player. —*J.H.*

Beating the Spread

Energy beams that stay in focus indefinitely can be generated

If you have ever cupped your hands around your mouth and shouted "Hello!" to a friend across a busy street—only to have the volume of your voice dwindle as the sound waves radiate in all directions—then you know how scientists feel who have attempted to prevent the energy transmitted by waves from spreading out from its source. Such frustration may be fading, at least in the world of high technology. Several investigators have developed techniques that generate pulses or beams of acoustic and electromagnetic waves that can remain in focus over their entire path. The work could yield new and more efficient applications in remote sensing, communications, power transmission and directed-energy weapons.

Beams of light and sound exhibit diffractive spreading because of the way waves bend around edges. Light, for instance, diffracts off the components of a laser. This diffraction caus-

es the width of the laser beam to spread as it travels through space. The divergence can be reversed by a simple lens, but once the beam is brought into focus it again spreads out very rapidly.

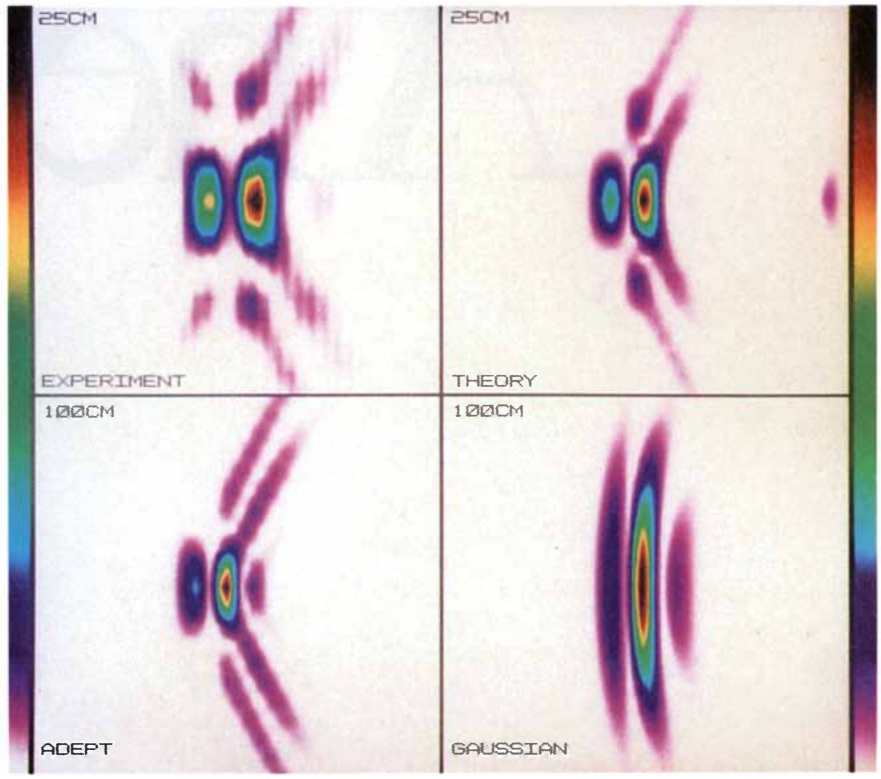
More than a decade ago James N. Brittingham of the Lawrence Livermore National Laboratory began investigating whether it is possible to propagate a beam or pulse that does not spread. In 1983 he showed in principle how such a system might work. Five years later Richard W. Ziolkowski and D. Kent Lewis of Lawrence Livermore and Bill D. Cook of the University of Houston began an attempt to simulate a nondiffracting pulse by producing a series of acoustic waves in a water tank. The workers created an electrical signal that caused a crystal to vibrate and generate a wave in the tank. Then they measured the intensity of the wave at a given distance from the source. They changed the signal, moved the crystal and generated another wave. Next they measured the new wave and repeated the process. After a number of waves had been launched the investigators added up the measurements, which gave them a picture of what the pulse would look like if all the waves had been launched simultaneously.

Ziolkowski and his colleagues report in *Physical Review Letters* that the simulated pulse did not spread out over a portion of the tank. As the group had predicted, the intensity of the pulse began to decay after traveling a critical distance. The workers say this is not a drawback because the distance can be made larger by manipulating the characteristics of the individual waves.

While the Livermore-Houston group worked on acoustic waves, James Durnin of the University of Rochester concentrated on electromagnetic radiation. Specifically, he explored the possibility of generating a continuous beam of light that did not exhibit diffractive spreading. He and his coworkers, Joseph J. Miceli, Jr., and Joseph H. Eberly, proposed in 1986 that such a beam could be produced by passing light through a circular slit and lens. Eberly says Durnin's predictions have been confirmed; a report of the work will be published soon.

"If these technologies can be developed further," Ziolkowski remarks, "the resulting beams would revolutionize directed-energy systems as well as remote sensing and communications media." Since a nondiffracting beam can be detected only by a receiver that is directly in its path, the beams could be used to carry private communications. If pulses of energy can be made to propagate over great distances, he says, a satellite that gathers solar radiation and sends it in beams to receivers on the earth might even be feasible. —*Russell Ruthen*

A series of underwater waves combine to form a pulse that remains in perfect focus as it travels



***NONDIFFRACTING PULSE** from an underwater source is simulated. In each panel the pulse's intensity is shown in cross section; the darker the color, the greater the intensity. At a distance of 25 centimeters from the source the nondiffracting pulse remains focused, as reconstructed experimentally* **(top left)** *and predicted theoretically* **(top right)**. *In a computer simulation the pulses in the two bottom panels are initially the same size. At a distance of 100 centimeters the nondiffracting pulse* **(bottom left)** *has not spread out, whereas a conventional pulse, called a Gaussian* **(bottom right)**, *has increased 2.8 times in size.*

BIOLOGICAL SCIENCES

Virus de Résistance

Drug-resistant viruses surface in AIDS patients

Ever since the discovery of penicillin, bacteria have evaded antibiotics by means of mutations that altered their susceptibility. Now it appears that viruses can resist drugs too. The first severe, progressive infections caused by drug-resistant viruses have been documented in AIDS patients. Although none of the resistant strains identified so far is likely to invade the general population, some virologists say the findings raise the possibility that AIDS patients could host other resistant pathogens posing a greater threat.

If viruses have been slow to display resistance to drugs, it is in part because they have not had much to resist. At present only about half a dozen antiviral drugs have been approved by the Food and Drug Administration. One of them is acyclovir, a drug licensed in 1982 for the treatment of herpes-simplex virus (HSV) infections. Acyclovir is the treatment of choice for initial outbreaks of genital herpes, which afflicts more than 20 million people in the U.S.

In an article in the *New England Journal of Medicine*, John Mills of the University of California at San Francisco, Clyde S. Crumpacker of Beth Israel Hospital in Boston, Gregory J. Mertz of

A new art

It is subtle, refined. Flawless function is heightened by clean aerodynamic form.

movement.

It is undeniably powerful and it will move you. The Accord Coupe. HONDA

the University of New Mexico School of Medicine and their colleagues report the clinical progression of HSV infections that did not respond to acyclovir in 12 patients who also had AIDS. In spite of increased doses of acyclovir and a derivative compound called ganciclovir, the herpes lesions continued to spread until the patients died.

In the same issue another article by Henry H. Balfour, Jr., of the University of Minnesota Health Sciences Center, Sunwen Chou of the Oregon Health Sciences University, Karen K. Biron of the Burroughs Wellcome Company and their colleagues describe resistance to ganciclovir in infections of cytomegalovirus (CMV). In most adults infection with CMV causes no more than a cold, but it can lead to life-threatening illnesses as well as blindness in people with AIDS. The workers found three patients—two with AIDS and a third with leukemia—whose CMV infections did not respond to ganciclovir.

No one is surprised that viral resistance made its first appearance in victims of AIDS and leukemia. Individuals whose immune systems have been weakened cannot defend themselves against even the least potent of viruses, and so many of them receive chronic antiviral therapy—which increases the opportunity for resistant strains to surface. Mills estimates that acyclovir-resistant HSV may infect one out of every 1,000 AIDS patients. Other immunocompromised people, such as transplant patients, may also be at risk.

The risks for people with normal immune systems are not as clear. No resistant strains of either CMV or herpes virus have been found in people who are not immunocompromised, and transmission of resistant HSV or CMV strains has not been reported. Resistant strains of HSV are usually less pathogenic than their drug-sensitive relatives. Resistant strains of influenza virus, however, have been found to cause illness in people with normally functioning immune systems. In the March issue of the *Journal of Infectious Diseases* Robert B. Belshe and his colleagues at the Marshall University School of Medicine in Huntington, W.Va., and at Hoffman-La Roche Inc. describe the spread of such strains among family members—the first reported transmission of a drug-resistant virus.

Robert T. Schooley of the Massachusetts General Hospital says resistant strains of HSV and CMV are unlikely to escape to the general population, because they, like AIDS, cannot be spread by casual contact. But Schooley says problems could arise in the future if AIDS patients harbor resistant pathogens that are more easily transmitted. "As the prognosis for patients with AIDS improves, it is likely that an even wider array of pathogens will emerge that are resistant to conventional therapy," he said in an editorial accompanying the reports in the *New England Journal of Medicine*.

Martin S. Hirsch of Massachusetts General, who coauthored the essay, says resistant viruses could probably be defeated by an assortment of antiviral drugs administered sequentially or in combination. "We have to put more effort into developing drugs that act by a variety of mechanisms," he says. He and other workers also stress the importance of prescribing antiviral drugs selectively, so that the viruses have limited opportunity to develop resistance. "If you start using them prophylactically," points out Robert F. Betts of the University of Rochester School of Medicine and Dentistry, "you're shooting yourself in the foot." —*Karen Wright*

A Sparrow's Fall

DNA of the dusky sparrow tells a cautionary tale

On June 16, 1987, the last dusky seaside sparrow died in a zoo at Walt Disney World, near Orlando, Fla. Now DNA salvaged from the bird has brought good news and bad about the dusky's demise: the loss may not have been as great as was feared, but a major effort to preserve the sparrow's genes may have rested on a faulty pedigree.

The dusky seaside sparrow lived in the salt marshes of Brevard County, Fla., on the Atlantic coast near Cape Canaveral. In the 19th century its dark color led to its being classed as a breed distinct from the other populations of seaside sparrows found along the Atlantic and Gulf coasts. During the 1950's and 1960's the number of dusky sparrows fell as the wetlands were drained, and by 1980 only six birds—all males—were left.

In a rescue effort mandated by the Endangered Species Act, five of the birds were captured and mated with seaside sparrows from a Gulf Coast population. By repeatedly crossing the male duskies with their own, hybrid daughters, workers at the University of Florida and at zoos in Gainesville and at Walt Disney World built up a population carrying mostly dusky genes. The birds may be released to reestablish a wild "dusky" population.

John C. Avise and William S. Nelson of the University of Georgia have applied molecular tests to the premise of the effort: that the dusky sparrow is evolutionarily quite distinct from other seaside sparrows. They analyzed DNA in mitochondria (organelles outside the cell nucleus that contain their own genetic material) from the last dusky sparrow and from other seaside sparrows found all along the Atlantic and Gulf coasts.

Mitochondrial DNA provides a sensitive test of relatedness because it evolves rapidly: populations separated from one another quickly develop differences in the DNA's sequence of nucleotides. The workers analyzed each bird's DNA by digesting it with restriction enzymes, which cut DNA at specific nucleotide sequences. They then sorted the fragments by size and compared the resulting "restriction patterns." A difference in the patterns indicates a sequence difference that creates or eliminates a cutting site.

At least in the stretches of DNA recognized by the restriction enzymes, the dusky sparrow was genetically indistinguishable from most of its Atlantic fellows. Birds from the Gulf Coast, on the other hand, all differed from the Atlantic birds by between six and nine restriction sites. The genetic differences, Avise and Nelson suggest, reflect a long separation between the Atlantic and Gulf populations. Other workers have suggested that the two populations stopped interbreeding several hundred thousand years ago, when falling sea level exposed the Florida peninsula.

If this picture of the seaside sparrow's taxonomy had been available earlier, the authors write in *Science*, "management programs for the species would likely have been very different." A major effort to preserve the dusky's heritage might not have been undertaken. In any case the captive duskies probably would not have been mated with birds from the Gulf Coast, from which they are genetically distinct, but with their practically identical neighbors on the Atlantic side.

Avise says he "wouldn't want to be understood as condemning the management program, which was based on the information available at the time." He points to a larger lesson: other conservation efforts may be based on taxonomies as dated as the one that guided the dusky-sparrow program. Such efforts may end up preserving spurious species or subspecies while overlooking real ones. "If we

A new microcomputer-controlled wire bond pull tester will enable manufacturers of hybrid microcircuits to perform fully automatic non-destructive pull testing of wire connections. Hughes Aircraft Company's automatic pull tester, designated Model 2600, features advanced vision and machine capabilities. Faultless testing of each wire is assured by gathering data on wire location from three sources: the global wire bond program, the wire bonder itself, and a wire path detection vision system that actually "sees" the wire being tested. In contrast to traditional manual pull methods, the Model 2600 will require fewer operators, completely log each hybrid under test by serial number, and assure one hundred percent visual and mechanical testing of every wire.

An advanced targeting aid will allow pilots to launch multiple Maverick missiles and provide maximum combat effectiveness during low-altitude, high-speed attacks. The pod-compatible Automatic Target Recognizer, under development by Hughes, receives imagery from an advanced infrared sensor and then uses statistical pattern recognition algorithms, combined with high-speed digital processing, to automatically detect, classify, and prioritize targets in the field of view. If desired, the system can make target selection and automatically fire the missiles. Several missiles could be launched in seconds, enabling the pilot to complete the mission quickly, possibly in a single pass.

A torpedo's guidance and control subsystem directs the torpedo to the target to successfully complete search, homing, and attack missions. The Mk-48 Advanced Capability (ADCAP) torpedo, designed and built by Hughes, initially receives its attack information from the submarine's fire control system. After launch, a long, thin communications wire between the torpedo and submarine serves as a real-time relay for changes in the torpedo's attack functions. Should the torpedo lose sonar contact with its target, its own tactical data processor returns to the search mode to relocate the target. If the target is able to initially evade the weapon, the ADCAP torpedo will re-enter and maintain an attack mode until interception.

More reliable transmission capabilities for domestic telecommunications customers will be one result of the combination of the Westar and Galaxy satellite systems. The Westar system, under purchase by Hughes, includes three C-band communication satellites in orbit, and one replacement satellite scheduled for launch in 1989. These satellites will be combined with three Hughes-owned and operated Galaxy satellites producing important operational savings. The six-spacecraft combined system will provide customers with significantly more reliable transmission capability because of the increased redundancy and extra channel capacity.

Hughes' Combat Systems Engineering Facility in San Diego, California has immediate openings in advanced development and training to support the Navy Command and Control Processor (C2P) and Advanced Combat Direction System (ACDS) Programs. Experience desired for Combat Systems Engineers includes 7-9 years of system level development of military systems, preferably Surface Navy Combat Systems. For Computer Programmers/Instructors the level of experience desired is 4-5 years of designing, coding and debugging computer software. Teaching or training experience is desired. Applicants must have a B.S. Degree in Computer Science or the equivalent. Please send your resume to Hughes Aircraft Company, Ground Systems Group, Dept. S2, P.O. Box 4275, Fullerton, CA 92634. Equal opportunity employer. U.S. citizenship required.

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

HUGHES

Subsidiary of GM Hughes Electronics

want to manage for biological diversity," Avise concludes, "we should be sure taxonomy rests on a firm genetic foundation." —*Tim Appenzeller*

Computerized Canaries

Wildlife monitoring provides an ecological early-warning system

In an era of multibillion-dollar particle accelerators, opportunities for amateurs to contribute to scientific research are few. Ornithology is an exception. The British Trust for Ornithology is harnessing the energies of Britain's indefatigable bird-watchers with the number-crunching power of the computer to create a biological monitoring system. The system should alert investigators to ecological disturbances resulting from pollution or destruction of important habitats before the damage becomes obvious—and harder to remedy.

Because birds occupy the top of many food chains, they are vulnerable to environmental insults. For example, a sudden drop in the number of peregrine falcons in Britain and related raptors in the U.S. first warned the world in 1961 of the dangers of persistent organochlorine pesticides. Investigation revealed that the birds were accumulating the compounds, which were present in their prey. The poisoning caused them to lay eggs with thin shells that broke easily.

At the heart of the monitoring system will be the largest collection of nest records in the world: 750,000 cards sent to the trust by enthusiastic amateurs since 1939. More than 30,000 cards continue to arrive at the trust's headquarters every year. The records have already proved their value in a number of studies. A recent study showed that cuckoos—European birds that lay their eggs in the nests of other species (as cowbirds do)—have changed their preferred species of dupe foster parents. Other trust schemes encourage birders to send in particulars of banded birds they find, and to trap and record birds regularly at selected sites.

Together with census data, information from these schemes enables workers to estimate three crucial numbers that determine the future size of a bird population: production of young, survival rate and present population size. The trust plans to develop computer-based population models for some two dozen "indicator species" selected for their sensitivity to a variety of environmental changes. The models will be able to identify any threatening trends in the populations and perhaps make it possible to set action thresholds. A similar project for the U.S. is being debated at Cornell University's Laboratory of Ornithology. The computer models might even point to currently unrecognized causes of environmental change. —*T.M.B.*

Multiple Choice

Viral fingerprints are found in the blood of MS patients

Finding the cause of multiple sclerosis, an incurable disease that afflicts 250,000 people in the U.S. alone, has been an intractable task. Investigators have long suspected that an infectious agent may be involved. Yet no one has a clear idea what destroys the myelin sheath around neurons in the nervous systems of MS victims, producing scarring that over the years leads to muscular tremors, paralysis and sometimes fatal complications.

In a case with no suspect any new lead is valuable. That explains the excitement stirred by two recent reports that some MS patients carry in their cells DNA sequences similar to genetic

sequences of a known virus. The discoveries suggest that the virus, or one related to it, could play a role in the disease. If it does, workers might have a target at which to aim new therapeutic strategies.

One report comes from E. Premkumar Reddy, Hilary Koprowski and their colleagues at the Wistar Institute of Anatomy and Biology in Philadelphia and Magnhild Sandberg-Wollheim of the University of Lund in Sweden. They reported in 1985 that some MS patients have antibodies that react with HTLV-I (human *T*-lymphotropic virus type I)—a meager clue. Early this year they presented in *Science* direct evidence of a virus: by means of a powerful technique that amplifies known DNA sequences they detected HTLV-I-like sequences in blood samples taken from a series of six Swedish MS patients. Only one out of 20 healthy people had the sequences. Koprowski says that "the association [between virus sequences and MS] is very strong." He believes the viral sequences in MS patients might be from HTLV-I itself or from a related virus.

The other report is the work of Steven J. Greenberg of the National Cancer Institute and Bernard J. Poiesz and Garth D. Ehrlich of the State University of New York Health Science Center at Syracuse. They say they will publish their results soon in *Proceedings of the National Academy of Sciences*. These workers also amplified DNA to search for several fragments of HTLV-I-like DNA. Greenberg's group found one of the fragments in blood from six out of 21 MS patients but not in blood from 35 healthy people. (Two other fragments they searched for could not be found either in the patients or in healthy people; a third fragment was found in every sample tested.) Thomas A. Waldmann, one of the investigators at the NIC, says the results lead him and his colleagues to believe "a novel human retrovirus related to but distinct from HTLV-I might be present in some cases of MS."

Circumstantial evidence also seems to strengthen the case: HTLV-I is known to cause at least one other degenerative disease of the nervous system and one form of leukemia. Other viruses of the same general type, the retroviruses, cause slow-acting diseases in animals as well as AIDS in human beings.

Neither group contends that a virus is the single cause of MS. Dale E. McFarlin of the National Institute of Neurological and Communicative Diseases and Stroke agrees: he observes that MS is commonest in northern regions where HTLV-I is rare. He concludes it is unlikely that HTLV-I alone causes MS. One theory is that the sequences might represent a viral accomplice that is activated when another agent triggers an immune reaction; the immune reaction might then turn against the patient.

There are other sources of doubt. McFarlin points out that the DNA-amplification technique is so sensitive that the most careful workers could easily contaminate samples and thereby produce invalid results. Furthermore, a disease called tropical spastic paraparesis caused by HTLV-I is hard to distinguish from the type of MS studied by Greenberg's team.

The knowledge that 20 different viruses have been implicated in MS over the years (three of them by Koprowski) further tempers the excitement, as does the fact that retroviral fragments are common in human DNA. Byron H. Waksman of the National Multiple Sclerosis Society points out that normal human genetic material contains genetic fragments from at least five families of retroviruses.

Nevertheless, these new findings mean the pursuit has once again been joined; this time the quarry, and perhaps hope for MS patients, may prove less elusive. —*T.M.B.*

SCIENTIFIC AMERICAN April 1989 Volume 260 Number 4

Global Climatic Change

Evidence suggests that production of carbon dioxide and methane from human activities has already begun to change the climate and that radical steps must be taken to halt any further change

by Richard A. Houghton and George M. Woodwell

The world is warming. Climatic zones are shifting. Glaciers are melting. Sea level is rising. These are not hypothetical events from a science-fiction movie; these changes and others are already taking place, and we expect them to accelerate over the next years as the amounts of carbon dioxide, methane and other trace gases accumulating in the atmosphere through human activities increase.

The warming, rapid now, may become even more rapid as a result of the warming itself, and it will continue into the indefinite future unless we take deliberate steps to slow or stop it. Those steps are large and apparently difficult: a 50 percent reduction in the global consumption of fossil fuels, a halting of deforestation, a massive program of reforestation.

There is little choice. A rapid and continuous warming will not only be destructive to agriculture but also lead to the widespread death of forest trees, uncertainty in water supplies and the flooding of coastal areas. When the ice now covering the Arctic Ocean melts, further unpredictable changes in the global climate will ensue. There may be controversy over whether the data are adequate and whether the warming is caused by changes in the atmosphere. Yet there is an unusually powerful consensus among climatologists that the dominant influence on global climate over the next centuries will be a warming driven by the accumulation of heat-trapping gases. The consequences are threatening enough so that many scientists, citizens and even political leaders are urging immediate action to halt the warming.

RICHARD A. HOUGHTON and GEORGE M. WOODWELL have collaborated for more than 20 years on topics of environmental concern. Houghton is an ecologist and a senior scientist at the Woods Hole Research Center in Woods Hole, Mass. For the past 10 years he has been concerned with the global carbon cycle and has specialized in the response of ecosystems, particularly forests, to climatic change. Woodwell is also an ecologist and is the director of the Woods Hole Research Center. He and Houghton hope to be able, along with their Woods Hole colleagues, to advance improved models for the management of renewable resources.

The fact that heat-trapping gases have been accumulating in the atmosphere is well established. Since the middle of the 19th century the amount of atmospheric carbon dioxide has increased by about 25 percent. The increase has come about because human activities, especially the burning of coal and oil and the destruction of forests, have released greater quantities of carbon dioxide into the atmosphere than have been removed by diffusion into the oceans or by photosynthesis on land [*see illustration on page 38*].

The increase in carbon dioxide appears trifling when one considers that the total amount in the atmosphere is a little more than .03 percent by volume. But in spite of its low concentration, carbon dioxide and several other gases present in even smaller amounts have an important role in determining the temperature of the earth. In contrast to both nitrogen and oxygen, which together make up more than 99 percent of the atmosphere, these trace gases absorb infrared radiation, or radiant heat. Since in this regard they act much like the glass over a greenhouse, they are commonly referred to as greenhouse gases.

Because the total amount of greenhouse gases is small, their concentrations are easily changed. An increase in the concentration of any one of them increases the atmosphere's capacity to retain heat and raises the temperature at which the atmosphere comes into equilibrium with the energy it receives from the sun. In recent years investigators have recognized that the atmospheric burden of greenhouse gases other than carbon dioxide, such as methane (CH_4), nitrous oxide (N_2O) and the chlorofluorocarbons (CFC's), is also growing at an increasing rate. By the mid-1980's, in fact, these gases had reached levels at which their combined effect approached that of carbon dioxide.

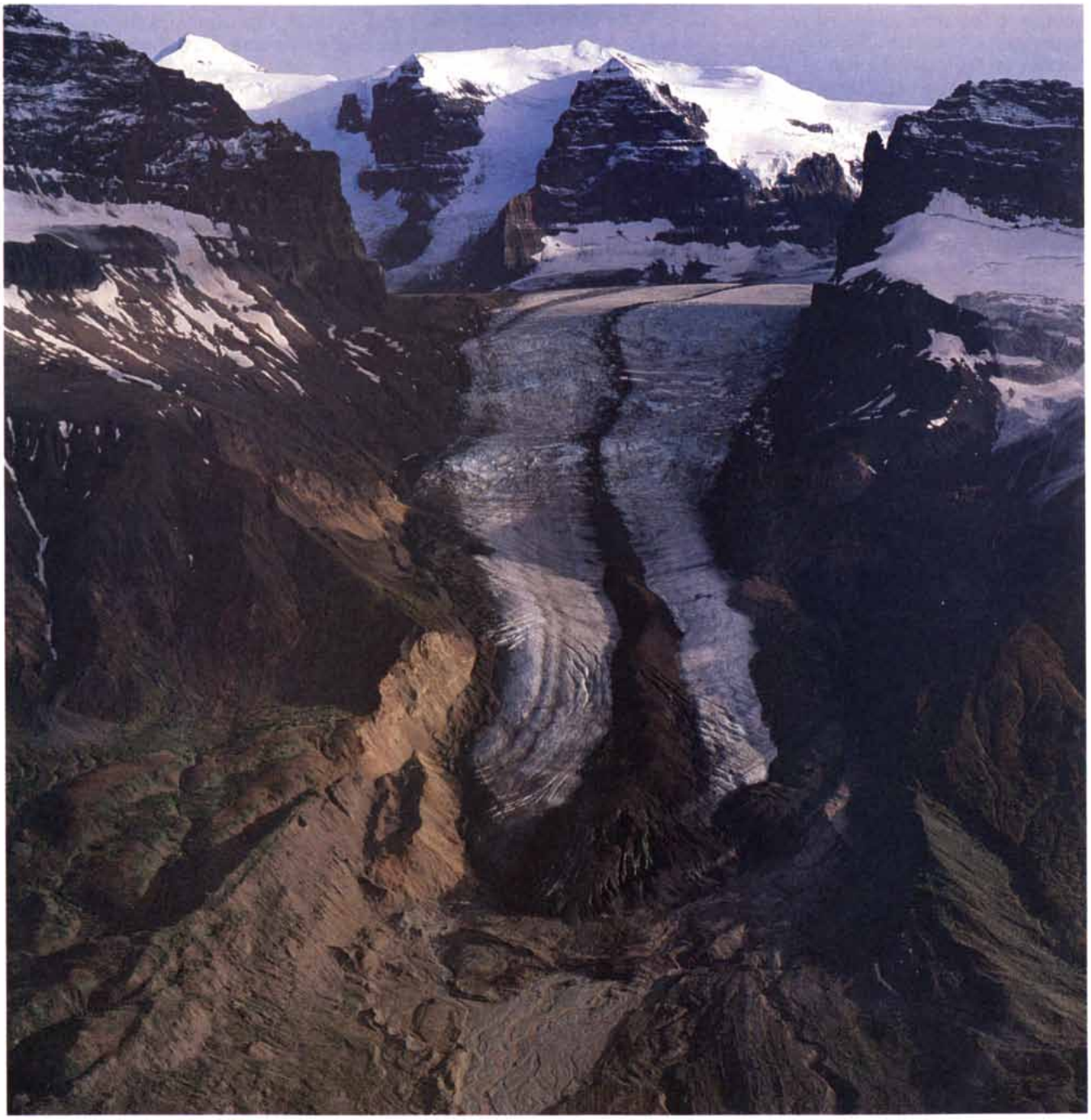
In this article we emphasize the role of carbon dioxide and methane because they are the principal contributors to the current warming, because their concentrations are strongly influenced by biological processes and because slowing or stopping the global warming will require control of carbon dioxide emissions in particular.

Global warming due to the accumulation of heat-trapping gases, particularly carbon dioxide, was predicted at the turn of the century by Svante Arrhenius in Sweden and Thomas C. Chamberlin in the U.S. Systematic research on the atmospheric accumula-

tion of carbon dioxide began only in 1958. Since then Charles D. Keeling of the Scripps Institution of Oceanography has provided a continuous record of the carbon dioxide level at various stations, the best-known of which is at Mauna Loa in the Hawaiian Islands [*see illustration on page 39*].

Information on the earth's temperature has been more difficult to accumulate. Strong evidence for global warming became available by late 1988. The most direct evidence lies in temperature records from around the world. James E. Hansen of the National Aeronautics and Space Administration's Goddard Institute of Space Studies and his colleagues have analyzed temperature records going back to 1860. Their analyses suggest that the average global temperature has increased by from .5 to .7 degree Celsius since that year. The greatest increase has taken place in the past decade; this recent warming is both statistically significant and consistent with their experience based on theory and on models of the global climatic system.

Thomas M. L. Wigley and his colleagues, working independently at the University of East Anglia in England, have also shown the increase in average global temperature. The rise has



HOLE-IN-THE-WALL GLACIER in Wrangell–St. Elias National Park, Alaska, is shown in an aerial view. The exposed ground at the foot of the glacier exhibits striations and moraines, the piles of debris left by a moving glacier; the ground has not had time to grow vegetation. All these are signs of recent glacial retreat. Similar behavior of a number of glaciers around the world (*see illustration on page 41*), the increasing depth to permafrost and other data suggest that warming has continued since the last glacial period. Its cause, however, extending back long before the current buildup of greenhouse gases, remains a puzzle.

not been observed in all regions: a recent analysis of climate records by Kirby Hanson and his colleagues at the National Oceanic and Atmospheric Administration shows no trend in temperature for the contiguous U.S. Such regional variation is not unexpected; the contiguous U.S. covers only 1.5 percent of the globe's surface.

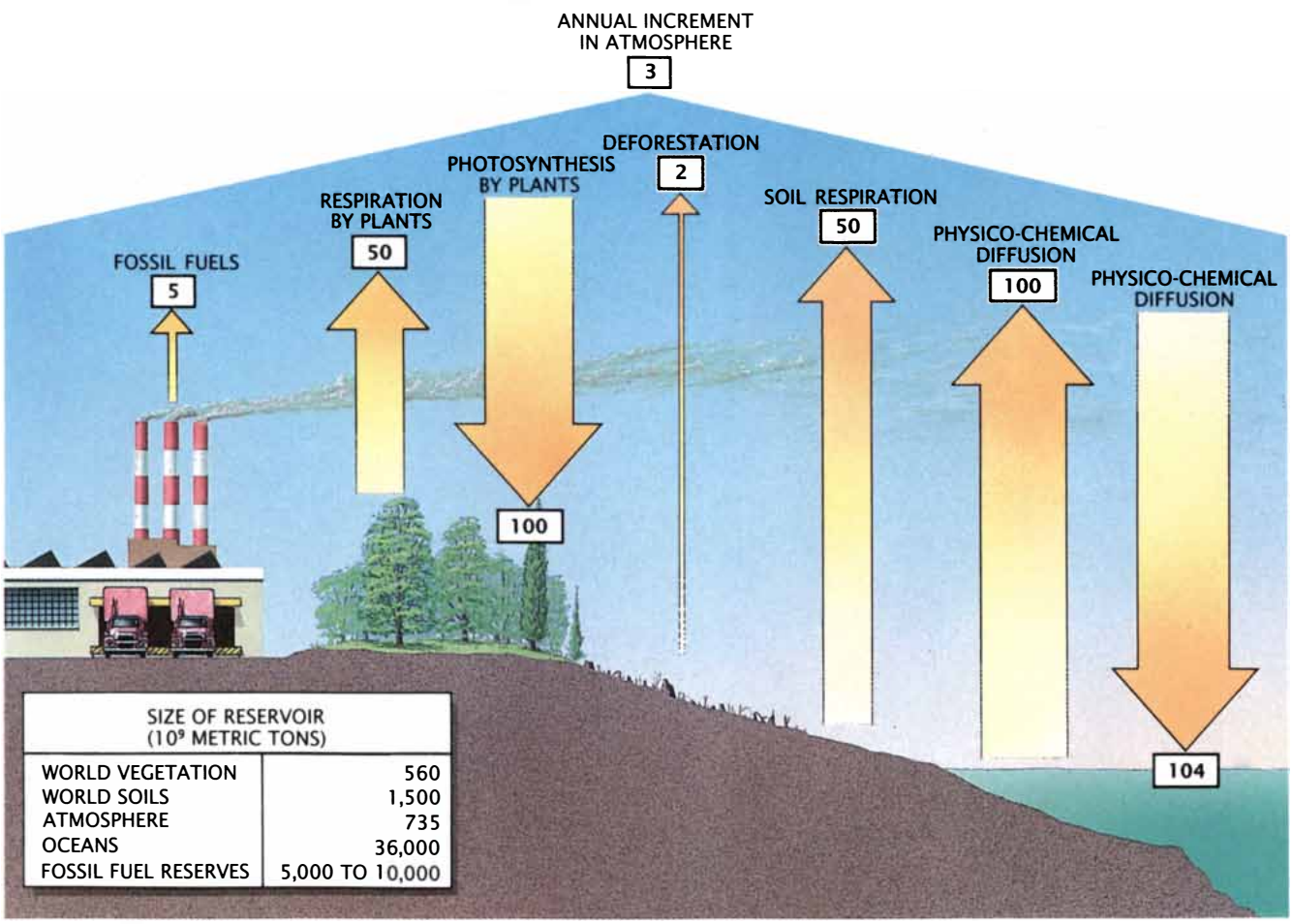
The observed rise in global temperature has not been steady and is clearly not simply a response to the accumulation of greenhouse gases. There was, for example, a decline in the mean global temperature between 1940 and 1965 in spite of the continued increase of heat-trapping gases in the atmosphere. Nevertheless, Phil D. Jones, one of Wigley's collaborators, has just recently reported that the global temperature has risen about .5 degree C since the beginning of the century and that the six warmest years on record were 1988, 1987, 1983, 1981, 1980 and 1986 in that order.

If a .5-degree temperature change seems insubstantial, one should remember that in 1816, the "year without a summer," the mean global temperature drop was also less than one degree. It was nonetheless sufficient to cause frosts in June in New England and widespread crop failures [see "The Year without a Summer," by Henry Stommel and Elizabeth Stommel; SCIENTIFIC AMERICAN, June, 1979]. The heat and drought that have afflicted North America and other regions of the earth in recent years are consistent with the predictions of a global warming trend.

There are other indications of an accelerated warming. According to Arthur H. Lachenbruch and B. Vaughn Marshall of the U.S. Geological Survey, the depth to permafrost in the Alaskan and Canadian Arctic has increased in recent decades. The average temperature of Canadian lakes has increased; the annual maximum extent of sea ice surrounding the Antarctic continent and in the Arctic seas appears to be declining; inland glaciers throughout Europe and elsewhere have receded.

These observations are consistent with predictions made by climatologists on the basis of theory aided by general circulation models. Several such global models exist and, although analyses based on them do not agree in detail, the general predictions are consistent with theory and experience. Climatologists expect that the greatest warming will occur at higher latitudes in winter. In these latitudes the warming, according to the models, will probably be at least twice the global average. In addition it is expected that the upper atmosphere will cool as the lower atmosphere warms and that there will be less precipitation and less moisture in the soil at lower latitudes. All these trends have been reported in recent years.

Data such as those are always open to further analyses, interpretation and augmentation. They invariably appear to suffer from inadequacies of



ANNUAL CARBON FLUXES are shown in units of one billion (10^9) metric tons. Photosynthesis on land removes about 100 billion tons of carbon from the atmosphere annually in the form of carbon dioxide. Plant and soil respiration each return about 50 billion tons. Fossil-fuel burning and deforestation release into the atmosphere respectively about five and two billion tons. Physicochemical processes at the sea surface release about 100 billion tons into the atmosphere and absorb about 104. The net atmospheric gain is about three billion tons annually. The table lists the world's major carbon reservoirs.

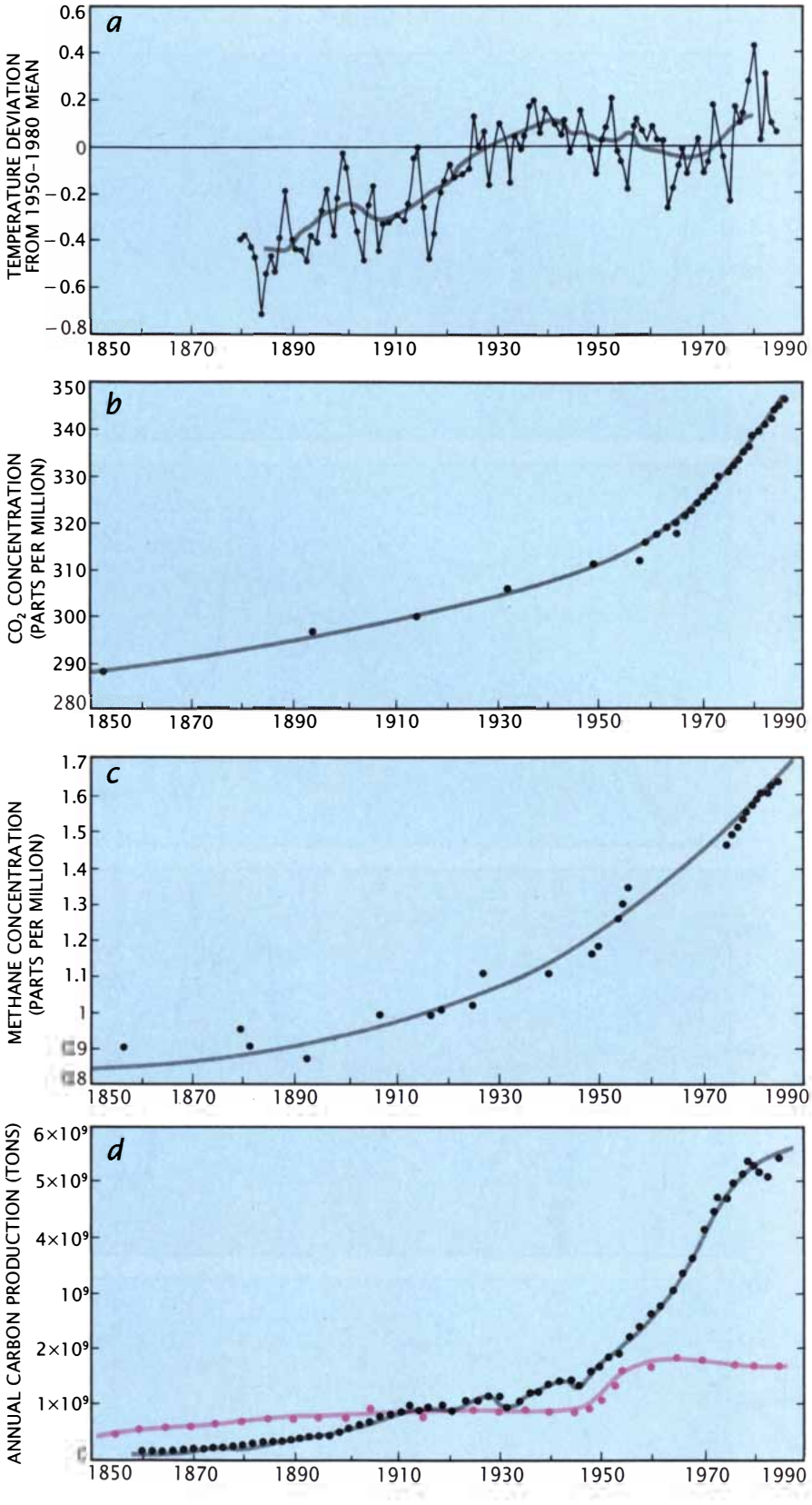
measurement and uncertainties about whether the period over which the measurements were taken was long enough to be significant. Investigators are currently improving the data and the analyses, but the fact remains that the observations described above, taken together with the rising concentration of greenhouse gases, constitute strong evidence that the process anticipated nearly a century ago by Arrhenius is under way.

One can learn much about potential future changes in climate by examining past climatic change. A mere 15,000 years ago glaciers covered much of North America and northern Europe. Were changes in the composition of the atmosphere involved in the great climate swings that brought glacial and interglacial periods? The answer is not completely clear, but one of the most important advances in recent years has been the ability to determine atmospheric composition in previous eras from tiny samples of air trapped in glacial ice. In particular, determination of the atmospheric composition during periods of glacial expansion and retreat has been made possible by data obtained from an ice core drilled by a joint French-Soviet team at the Antarctic Vostok station.

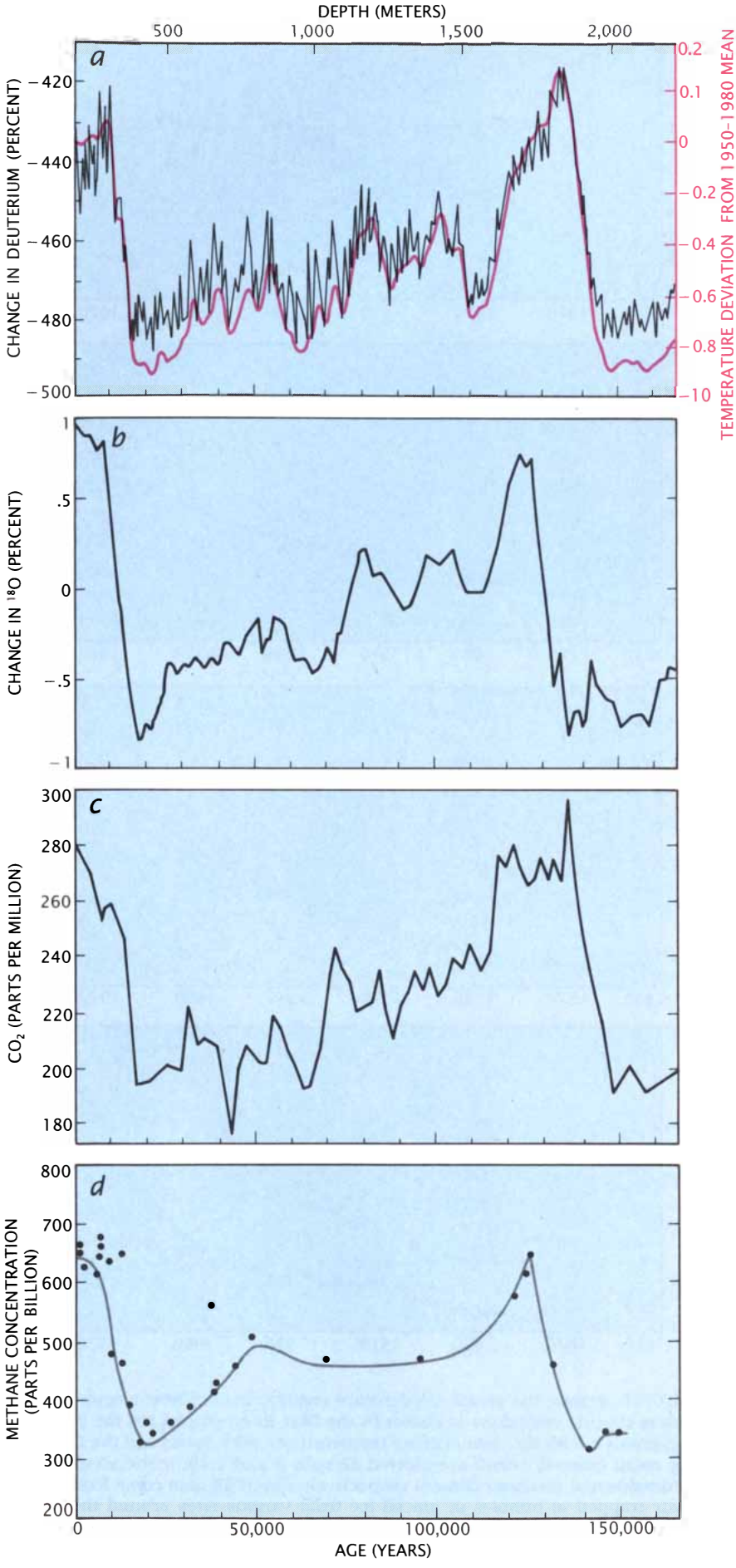
The Vostok core, as it is called, was 2,000 meters in length, long enough to sample ice dating through the past 160,000 years [*see illustration on next page*]. The data show fluctuations in temperature of up to 10 degrees; such fluctuations are derived from changes in the isotopic ratios in the core. It is well established, for example, that the ratio of the two common isotopes of oxygen, ^{18}O and ^{16}O, in cores of marine sediments reflects past temperature changes.

The Vostok data also show how the abundances of atmospheric gases have fluctuated with temperature over the past 160,000 years: the higher the temperature, the greater the concentration of carbon dioxide and vice versa. To be sure, the correlation of carbon dioxide with temperature does not establish whether changes in atmospheric composition caused the warming and cooling trends or were caused by them. Although the carbon dioxide content follows temperature very closely during periods of deglaciation, it apparently lags behind temperature during periods of cooling.

Although there is tight statistical coupling between carbon dioxide and temperature throughout the record, the temperature changes are from five



CORRELATION among the global temperature change, level of heat-trapping gases and carbon dioxide emissions is shown in the first three graphs for the past 140 years. In graph *a* both the annual mean temperature (*spiky curve*) and the five-year running mean (*smooth curve*) are plotted. Graphs *b* and *c* show the atmospheric carbon dioxide and methane content respectively. Pre-1958 data come from analyses of air trapped in bubbles of glacial ice from various sites around the world. The annual production of carbon from fossil-fuel burning (*black*) and from change in land use (*color*) is shown in *d*; the last data were obtained from historical sources.



to 14 times greater than would be expected on the basis of the radiative properties of carbon dioxide alone. This relation suggests that quite aside from changes in greenhouse gases, certain positive feedbacks are amplifying the response. Such feedbacks might involve ice on land and sea, clouds or water vapor, which also absorb radiant heat.

Other data from the same Vostok core sample show that methane also closely follows temperature and carbon dioxide. The methane concentration nearly doubled, for example, between the peak of the penultimate glacial period and the following interglacial period. Within the present interglacial period it has more than doubled in just the past 300 years and is rising rapidly. Although the concentration of atmospheric methane is more than two orders of magnitude lower than that of carbon dioxide, it cannot be ignored: the radiative properties of methane make it 20 times more effective molecule for molecule than carbon dioxide in absorbing radiant heat. On the basis of Hansen's radiative-convective model, which includes chemical feedbacks, methane appears to have been about 25 percent as important as carbon dioxide in the warming that took place during the most recent glacial retreat 8,000 to 10,000 years ago.

How can a global rise in temperature be expected to cause greater releases of carbon dioxide and methane into the atmosphere? In the process of photosynthe-

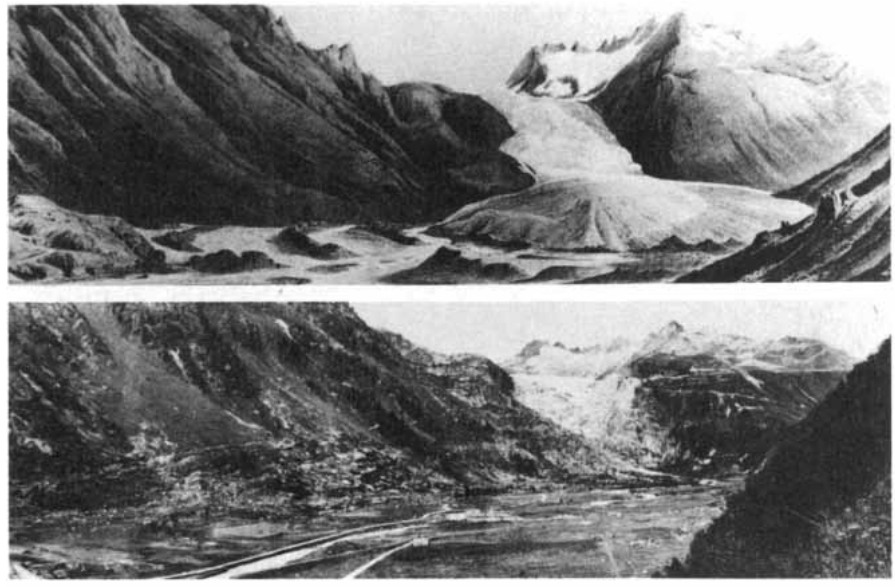
VOSTOK ICE-CORE DATA reveal a correlation between certain gas concentrations and temperature over the past 160,000 years. The ice core, 2,200 meters long, contains bubbles of air with carbon dioxide and methane that were trapped at different depths (*top scale*) and hence at different times (*bottom scale*). Several independent methods have established that the deuterium concentration in ice is a good measure of past temperature; both temperature and deuterium level are plotted in *a*. More traditional is the use of the oxygen isotope ^{18}O to track temperature; curve *b* is almost identical with *a*. The remarkable agreement with the shape of the Vostok-station carbon dioxide curve *c* argues that carbon dioxide can also serve as a global thermometer. Data on Antarctic methane compiled in 1985 and 1986 from several stations (*d*) strengthens the conclusion that levels of greenhouse gases are positively correlated with temperature and may actually influence it.

sis terrestrial plants remove about 100 billion tons of carbon from the atmosphere per year, or about 14 percent of the total atmospheric carbon content. An approximately equal amount of carbon is returned to the atmosphere through the processes of plant respiration and decay of organic matter. Because the fluxes are a substantial fraction of the carbon dioxide already in the atmosphere at any time, a change of a few percent in either the photosynthetic or the respiratory flux would soon significantly alter the atmospheric carbon dioxide content. Will global warming produce such an imbalance?

The answer is unclear and probably will remain so until after the climate has changed considerably more than it has already. Nevertheless, the general picture is probably as follows. The rate of photosynthesis is affected by many factors, particularly the availability of light, water and nutrients. It is not, however, very sensitive to temperature change. The rates of plant respiration and decay, on the other hand, do strongly depend on the temperature. A one-degree temperature change in either direction often alters rates of plant respiration by from 10 to 30 percent.

These observations suggest that a global warming will speed the decay of organic matter without appreciably changing the rate of photosynthesis. That will increase the release of carbon dioxide into the atmosphere. A warming will also result in more methane, because methane is produced by respiration in regions where oxygen is not freely available, such as swamps, bogs and moist soils. In recent years there has been a rise in the concentration of atmospheric methane of more than 1 percent per year. The increase is both rapid and significant because, as noted above, methane is 20 times as effective as carbon dioxide in trapping heat. The wet soils where methane is produced as a result of anaerobic decay probably represent the world's major source of methane. The global warming that has already occurred has undoubtedly stimulated anaerobic decay and the production of methane as well as carbon dioxide.

It is possible to estimate the size of the resulting increase in carbon production at least crudely. A significant fraction (from 20 to 30 percent) of global respiration on land takes place in the forest and tundra of the middle and high latitudes, where the warming is expected to be greatest. If we assume that the mean global warming to date has been .5 degree C, and that in the middle and high latitudes the rise has been one degree, then plant respiration in these latitudes and the decay of organic matter in soils has increased significantly. If the increase in respiration is between 5 and 20 percent over 20 to 30 percent of the total area respiring, then total global respiration will increase between 1 and 6 percent above normal. Once again assuming that the annual flux of carbon into the atmosphere is 100 billion tons and that the rate of photosynthesis remains unchanged, the warming that has already taken place has meant an injection of between one and six billion tons of carbon per year. Over the past century from 20 to 30 billion tons of carbon may have been released in this manner.



RHÔNE GLACIER in Switzerland is shown in a lithograph after an 1848 watercolor by Henri Hogard (*top*). Four sets of moraines are clearly visible. The outermost set has been dated to 1602, the second to 1818, the third to 1826 and the fourth to 1848; the pattern indicates that the glacier had been retreating up the valley for at least 250 years. A photograph from 1970 (*bottom*) shows that the glacier has retreated still farther up the valley. The retreat is additional evidence for recent global warming.

That estimate is probably high, because the average warming may have been less than assumed and because photosynthetic response will tend to reduce the release of carbon dioxide. Yet the estimate is probably not high by as much as a factor of two, and it serves to emphasize the importance of biotic feedback mechanisms.

How does the value just computed compare with amounts of carbon released by other known processes? The release from the burning of fossil fuels is approximately 5.6 billion tons per year; deforestation adds an amount estimated at between .4 and 2.5 billion tons per year. The total carbon injected into the atmosphere from these two sources added to a temperature-enhanced respiration is not known, but it appears to be more than six billion tons annually and may approach 10 billion.

The release of carbon due to changes in the respiratory rate could fluctuate appreciably; a gradual warming, such as that experienced over most of this century, would change the respiratory rate slowly enough so that year-to-year changes would be inconspicuous. On the other hand, a sudden warming or cooling over a period of several years might result in an observable change in the carbon dioxide content of the atmosphere. In the past 15 years the annual rate of accumulation of atmospheric carbon dioxide has been about 1.5 parts per million, equivalent to a global accumulation of about three billion tons of carbon.

According to data recorded on Mauna Loa and at the South Pole by Keeling, however, over the past 18 months the accumulation rate has risen to about 2.4 parts per million, equivalent to about five billion tons of carbon. Keeling expects that the surge will prove transitory, as a lesser surge in 1973 and 1974 did. Nevertheless, the implication we assign to the observations at the moment is that the surge is a result of the high temperatures that have marked the 1980's, delayed by the time necessary to warm the soil. Whether this interpretation is correct remains to be seen.

Any climatic change can also be ex-

"GLOBAL WARMING FLOODS FLORIDA" could be a tabloid headline if the polar ice caps began to melt. Florida is shown here as it might look if sea level rose 4.6 meters above (*bluish green*) or 7.6 meters above (*light green*) the present level (*blue*). In either case Miami and Lake Okeechobee are submerged. A rise of four to five meters might be expected if the West Antarctic Ice Sheet broke up under global warming.

pected to affect the ability of the terrestrial biota, in particular forests and soils, to retain carbon. At warming rates that are lower than the rates at which forests develop, forests may actually expand, and with them the capacity to store carbon. But if the warming rate exceeds the rates at which forests migrate into more climatically favorable regions, widespread mortality of trees and other plants is likely to follow. The net result of such destruction of forests is difficult to predict, but it will probably mean a further release of carbon dioxide through the decay of plants, animals and organic matter in soils.

The amount of carbon dioxide that could be injected into the atmosphere would depend heavily on the rate of climatic change in the forested zones of the middle and high latitudes. Although it is impossible to make any accurate calculation, an upper limit is given by the amount of carbon in these forested latitudes: approximately 750 billion metric tons, or about the same amount of carbon as there is in the atmosphere currently.

Is it possible that a global warming could stimulate the growth of forests? In this case the spread of forests to high latitudes and tundra regions would result in a greater uptake of carbon dioxide from the atmosphere and a greater accumulation of carbon dioxide in the soil. Such a transition is unlikely. Forests require centuries to develop, especially where soils are thin and nutrients are in short supply. They also require climatic stability and sources of seeds. The climatic transitions currently under way, unless they are checked, are rapid by any measure and can be expected to continue into the indefinite future. They do not offer the conditions under which forests are able to develop on new land and remain for long periods.

Might the warming at least stimulate existing forests to store additional carbon in plants and soil? Perhaps. The boreal forest and other coniferous forests may indeed be sufficiently resilient to respond to warming with increased photosynthesis and growth. Whether the carbon taken up by photosynthesis will be stored or simply released through increased respiration remains an open question.

There is also the possibility that the tundra, the treeless plain found in arctic and subarctic regions, will respond to a warming in surprising ways, including an increase in the production of carbon and its storage in

peat. The nature of the response will largely hinge on the availability of water. A wetter tundra might store additional carbon in soils; a drier tundra might release it through the decay of organic matter in long-frozen soil or soil that is normally frozen for most of the year. W. Dwight Billings of Duke University believes global warming will speed the decay of peat in tundra soils and precipitate that ultimate breakdown of the tundra known as thermal karst erosion, which allows flowing water to erode the tundra in great acre-size chunks. Not only is the tundra devastated but also substantial amounts of carbon dioxide and methane that were stored in the peat as carbon are released into the atmosphere.

The evidence indicates that under rapid planetary warming respiration rates will increase more than photosynthesis rates. The changes will lead to the release of additional carbon dioxide and methane into the atmosphere. The magnitude of the release will hinge strongly on the rate of warming: the faster the warming, the larger the release. Such behavior is consistent with (but not proved by) the data from the Vostok core.

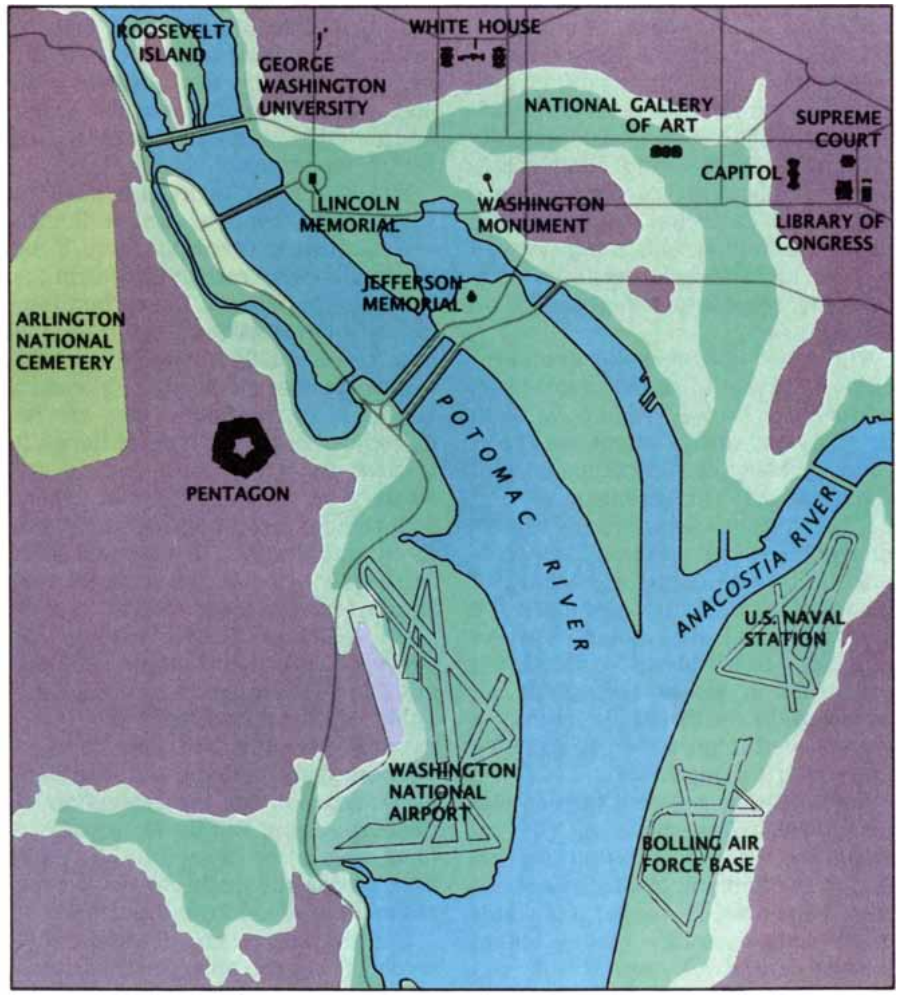
What will be the consequences of a continued global warming? In 1985 a group of meteorologists meeting under the auspices of the World Meteorological Organization (WMO) and the United Nations Environment Programme (UNEP) demonstrated that without the respiratory feedback mechanisms addressed above, the combined effect of the greenhouse gases would warm the earth by an average of from 1.5 to 4.5 degrees C before the middle of the next century. The conclusion was recently confirmed in a review written by more than 50 scientists who met in Villach, Austria, in 1987 and was published by the WMO and the UNEP.

Seldom has there been such a strong consensus among scientists on a major environmental issue. The warming, unless consciously checked by human effort, will be rapid and will be felt differentially over the earth. Winter temperatures in the middle and high latitudes can be expected to rise by more than twice the world average. If the mean global temperature were to rise by from two to three degrees C by the year 2030, the winter temperature increase in Minneapolis might approach from four to six degrees C, or about one degree per decade. Summer temperatures would also rise, but less severely. A one-degree change in temperature is equivalent to a change in latitude of from 100 to 150 kilometers. The prairie-forest border, which is now south and west of Minneapolis, might be expected to migrate north at a rate of between 100 and 150 kilometers per decade, or between 400 and 600 kilometers by the year 2030.

Such changes are likely to be difficult for most of the world's peoples. First, the changes will be continuous. Unless the warming stops, efforts to adapt to climatic changes are likely to be responses to conditions that no longer exist. Second, the changes in climate will be irreversible for any time of interest to us or our children. There is no way to cool the earth or to lower sea level; we cannot return quickly to an atmosphere with lower concentrations of greenhouse gases. The best we can do is to reduce current emissions. If that step is taken immediately, a further warming of more than one degree can be expected as the full effects of the heat-trapping gases already present are felt.

Finally, the effects are open-ended. Although most modeling to date simulates a doubling of the atmospheric carbon dioxide content, there is simply no reason to assume that the concentrations will stop at twice the current levels. Estimated reserves of recoverable fossil fuels in themselves are enough to increase the atmospheric concentration of carbon dioxide by a factor of from five to 10.

Can anything be done to slow the climatic change that is now under way? The immediate need is to stabilize the greenhouse-gas content of the atmosphere. Regardless of its source, over the past decade carbon has been accumulating in the atmosphere at a rate of about three billion tons annually. (The remainder is being absorbed by the oceans or stored in forests and soils.) If current fluxes were reduced by three billion tons annually, the atmospheric carbon dioxide level would be stabilized for a few years. The stabilization would not



WASHINGTON, D.C., is depicted here under the same conditions as in the preceding illustration. Washington National Airport and the Lincoln Memorial are inundated. The 7.6-meter contour reaches almost to the Capitol steps and to the White House.

be permanent, however. The rate of accumulation in the oceans is determined by how fast they can absorb carbon dioxide from the atmosphere; this in turn depends on the difference in carbon dioxide concentration between the atmosphere and the ocean. As the flux of excess carbon is reduced, the difference is also reduced and the ocean becomes less capable of absorbing excess carbon; carbon dioxide emissions would have to be reduced still further to prevent additional atmospheric accumulation.

The largest source of carbon dioxide emissions is the combustion of fossil fuels, which releases about 5.6 billion tons of carbon into the atmosphere annually. Industrial nations contribute about 75 percent of these emissions; steps toward stabilizing the composition must begin in the industrialized world. A recent study carried out under the auspices of the World Resources Institute and led by José Goldemberg, president of the University of São Paulo in Brazil, suggests that the consumption of energy from fossil fuels in the developed nations could be halved by a program of conservation and improved efficiency alone.

Although developing nations produce less carbon dioxide, their contributions are growing; if economic development follows conventional patterns, their potential contributions are very large. The second step toward the stabilization of greenhouse gases will require innovations in economic development that lessen dependence on fossil fuels.

The other known major source of carbon dioxide is deforestation, predominantly in the Tropics. By 1980 about 11,000 square kilometers of forest were being cleared annually, with the result that in 1980 between .4 and 2.5 billion tons of carbon (as carbon dioxide) were released into the atmosphere. The rate of deforestation has increased over the past decade. If the present release of carbon is near the upper end of the above range, halting deforestation would reduce carbon emissions by the three billion tons per year needed immediately to stabilize atmospheric composition.

Reforestation will also help to stabilize the composition of the atmosphere. The reforestation of from one to two million square kilometers (about the area of Alaska) will result in the annual storage of one billion tons of carbon. Although this area is large and productive land in the Tropics is at a premium, there may be as much as 8.5 million square kilometers of once forested land available for reforestation. Of this land, about 3.5 million square kilometers could be returned to forest if permanent agriculture were to replace shifting cultivation. Another five million square kilometers of deforested land are currently unused, and there reforestation could in principle be implemented immediately. Forests established to store carbon would, of course, have to be maintained: neither harvested nor destroyed by toxic effects or change in climate.

Each of the measures to stabilize the atmospheric carbon dioxide level would have salutary effects locally, regionally and nationally, quite apart from its effects on climatic change. An improvement in energy-use efficiency, a step that might have been taken long ago with benefits to all, would bring economic and material advantages to both individuals and nations. An improvement in efficiency would lessen reliance on fossil fuels; this in turn would reduce sulfur and nitrogen oxide emissions, acid deposition and the release of other toxins. Halting deforestation would help to maintain the genetic diversity of the planet, reduce erosion, stabilize local and regional climates, cleanse water and air and preserve opportunities for future generations.

No one remedy by itself is likely to stabilize the levels of carbon dioxide and methane in the atmosphere. If the accumulation of carbon dioxide in the atmosphere persists, the carbon burden will have shifted from three billion tons annually to five billion tons and will be that much more difficult to address. The measures that are required can begin at home, although it is clear the world must join in the effort if it is to be effective. There are precedents for international action on similar issues. The Limited Test Ban Treaty of 1962 was an agreement among certain nations to avoid atmospheric tests of nuclear weapons. It has been effective. Nations that did not sign it (France and the People's Republic of China) have yielded to international pressure and now conduct weapons tests underground. The Vienna Convention for the Protection of the Ozone Layer and the Montreal Protocol, the latter negotiated in 1987, have moved the world far toward the elimination of chlorofluorocarbons.

There is no reason to assume that similar progress cannot be made with carbon-based fuels and deforestation. With that end in view a series of steps has already been undertaken: 50 specialists in international diplomacy and law met recently under the auspices of the Woods Hole Research Center to outline approaches that might work. The greatest problem is gaining the active and effective support of the developing nations, which are poised for a massive increase in fossil-fuel consumption. Development need not, however, follow historical paths. To cite one example, the low-latitude countries stand to gain immeasurably as techniques for exploiting solar energy are perfected. Solar-powered electrolysis of water can produce hydrogen, which in turn can run automobiles and other machinery. There are few places in North America where domestic hot water cannot now be produced by solar energy at little or no cost throughout most of the year. Nor is it to the advantage of nations to allow their forests to be destroyed.

Conferences are under way in the developing nations to explore alternatives to the present course. The first was held in New Delhi in February; the second is planned for São Paulo in September under the leadership of Goldemberg. The conferences will explore the possible responses of developing nations to a world in which conventional energy sources are limited. There are extraordinary opportunities for industrial innovations, particularly in energy efficiency and solar power. But developing countries cannot be expected to shoulder the entire burden; the developed nations, which are responsible for most of the problem, must do their share.

These issues will persist throughout the next century and dominate major technical, scientific and political considerations into the indefinite future.

FURTHER READING

GLOBAL DEFORESTATION: CONTRIBUTION TO ATMOSPHERIC CARBON DIOXIDE. G. M. Woodwell, J. E. Hobbie, R. A. Houghton, J. M. Melillo, B. Moore, B. J. Peterson and G. R. Shaver in *Science*, Vol. 222, No. 4628, pages 1081–1086; December 9, 1983.

THE FLUX OF CARBON FROM TERRESTRIAL ECOSYSTEMS TO THE ATMOSPHERE IN 1980 DUE TO CHANGES IN LAND USE: GEOGRAPHIC DISTRIBUTION OF THE GLOBAL FLUX. R. A. Houghton, R. D. Boone, J. R. Fruci, J. E. Hobbie, J. M. Melillo, C. A. Palm, B. J. Peterson, G. R. Shaver, G. M. Woodwell, B. Moore, D. L. Skole and N. Myers in *Tellus*, Vol. 39B, Nos. 1–2, pages 122–139; February–April, 1987.

GLOBAL TRENDS OF MEASURED SURFACE AIR TEMPERATURE. James Hansen and Sergej Lebedeff in *Journal of Geophysical Research*, Vol. 92, No. D11, pages 13345–13372; November 20, 1987.

INTRODUCING THE NEW DODGE SPIRIT.

THE ALL-NEW DODGE SPIRIT ES. *Dodge Spirit ES is a true driver's sedan. You can tell by the way it looks. And by the way it moves. Power comes from a new 2.5L fuel-injected turbo. Or choose a 3.0L V-6 and new electronic 4-speed Ultradrive automatic. You also get responsive power steering, a sport suspension and our exclusive 7 year or 70,000 mile Protection Plan.* Plus, Spirit prices start at just $9,995, while the premium ES (shown) is $12,495.† So see your Dodge dealer. Then let your Spirit soar.* 7/70

Dodge **THE NEW SPIRIT OF DODGE**
THE PERFORMANCE DIVISION OF CHRYSLER MOTORS

*See this powertrain limited warranty & its restrictions at dealer †Sticker prices exclude tax, destination charge & options.

BUCKLE UP FOR SAFETY.

THE LIFESTYLE RESOURCE

At The Lifestyle Resource we give you all the facts and details necessary to make an informed purchase. Your satisfaction is our primary concern. If your purchase doesn't meet your expectations, return it in original condition within 30 days for prompt refund.

WORTH CUTTING THE CORD FOR

A leading consumer magazine article likens a person's first conversation walking around talking on a cordless telephone to the exhilaration of that first bike ride minus training wheels. In the magazine's current 1988 Buying Guide, the Southwestern Bell FF-1700 Cordless Phone is still rated tops for range in controlled tests among 21 brands and models. By handling incoming and outgoing calls to a range of 1000′ (the article rated a maximum of 1500 ft.), with outstanding speech quality and convenience features, the FF-1700 ended up on top of the consumer magazine ratings reports. Base unit serves as freestanding speakerphone with dialpad, so you get two phones in one. Plus intercom, paging and 10-channel selection. Digital security code protects line from outside access. **$179.95** #2130. To take your freedom a step further, Southwestern Bell's FA-450 Telephone Answering Machine gives you the latest technology and newest features at a most attractive price. Single cassette operation, call screening, household memo function, voice-activated record, one-touch playback. Two-way record for messages or conversations. Beeperless remote lets you retrieve messages from any pushbutton phone, also allows remote announcement changes. These Freedom Phones connect you in without tying you down. **$99.95** #2140.

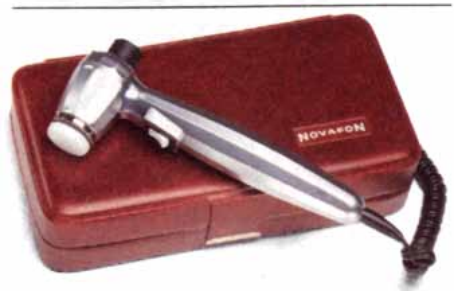
FROM CHINA TO YOUR HEALTH

Ancient mandarins dating back 800 years believed that these Chinese Exercise Balls induced well-being of the body and serenity of spirit. These treasured gifts were given to President Reagan and his wife while visiting the Peoples Republic of China. The Chinese say that rotating the balls in the palm of each hand stimulates the fingers and acupuncture points, and improves circulation of vital energy throughout the body. Sports enthusiasts, musicians, computer users and health-conscious people everywhere consider them great muscle conditioners. Arthritis sufferers feel a decided benefit from this gentle but challenging exercise. Very effective for relaxation and meditation, Chinese Exercise Balls emit a distantly mysterious chime as you turn them. Beautifully handcrafted, 45mm. hollow polished chrome balls are perfectly weighted and fit comfortably into the average man's or woman's hand. In exquisite silk brocade box. **$29.95** #1701.

A WATCH FOR ALL REASONS

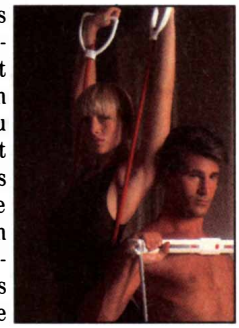
This handsome digital-analog chronometer will swim with you (to a depth of 150 feet), fly with you, calculate your speed, wake you in the morning, and tell you the day and date. If you're overseas, keep the digital timekeeper on home time, and adjust the analog hands to the new time. If you like, a push of a button will convert the digital display to 24-hour timekeeping, to conform to European standards. All functions—digital and analog timekeeping, calendar, alarm, chronometer and tachymeter—are powered by a highly accurate electronic quartz movement. The analog hands and hour markers are luminous. The stopwatch times to 1/100th of a second, and the rotating bezel can be used for such practical matters as telling you when the parking meter will need another quarter. Adjustable fit. **$39.95.** #1051 Teflon coated durable black matte anodized stainless steel; #1061 Stainless steel and gold-plated.

SOOTHING SOUND MASSAGE



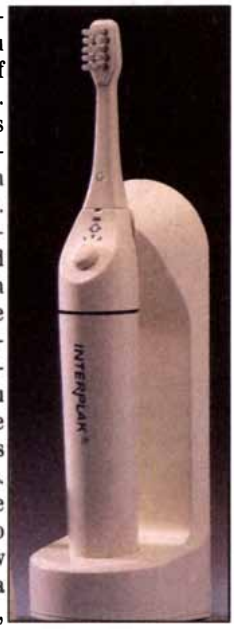
For more than 20 years, people, clinics and health centers around the world have used the Novafon sound massager. Novafon's sound waves penetrate up to 2¼″, and they help a variety of conditions because of their mixed frequencies. When used as directed, Novafon will bring relief from pain, loosen joints and soothe tired muscles, help circulation, speed recovery from exercise and over-exertion. The Novafon is lightweight (8 oz.) small (8″length). Adjustable intensity control, choice of massage heads (disc-type and ball-type). It comes in a fitted case, perfect for carry along. 1 year manufacturer's warranty. A precision made instrument with no interacting parts to wear out, the unit will give many years of service. **$169.95** #1750.

THE LIFELINE GYM



The Lifeline Gym is the most space efficient, time efficient home fitness device on the market today. You can simulate just about any of the exercises done on the expensive machines found in health spas and gymnasiums — the reason is latex. With this unique stretchable latex cable, resistance increases with movement—your muscles are challenged through their full range of motion. These variable resistance exercises provide an excellent workout for a beginner as well as a challenge for a pro. The Lifeline Gym is so successful it is used by members of the Chicago Bears, the U.S. Ice Hockey team, and the UCLA track team to name a few. Adjustable for individual strength levels and for different body shaping and toning. The gym includes a lifting bar, resistance cable, stirrups, exercise belt, door attachment and a comprehensive fitness guide that outlines over 25 exercises. Everything packs into a compact carry case, so you can enjoy the benefits of a complete workout anywhere. **$49.95** #1840 Ladies or #1850 Men's.

YOUR SMILE IS BEAUTIFUL



But you're only fooling yourself if you think it's proof of healthy teeth and gums. The unseen enemy is a natural mouth substance deposited on the teeth — plaque. Unchecked, plaque develops into rock-hard tartar behind which decay breaks down the tooth enamel and invades the soft, sensitive substructure. You have to remove the plaque before it hardens to prevent the process. Until recently there was no easy solution to this problem, but now there's Interplak, a precision - engineered, university-tested, three-speed plaque removal instrument. Its 10 independently driven bristle tufts rotate around their own axes: each tuft spins in the direction opposite to the spin of its neighbor at 4,200 rpm. The bristles are soft and rounded, adhering to dentistry recommendations, so that virtually every particle of plaque — *98% of it* — is removed gently and without irritation. Interplak even reaches between teeth and around orthodontic braces. Children think it's fun, so they're eager to brush. Interplak is cordless and easily stores in a free-standing or wall-mounted recharger/holder. Individual color-coded brush heads are available. The Interplak Home Plaque Removal Instrument with charger and two brush heads. **$99.95** #1110. Two additional brush heads **$29.95** #1111.

The counter-rotational movement cleans plaque.

SIT UP STRAIGHT!

Sitting in a conventional chair, bent over a desk or table for long periods of time can stress your spine and back muscles. Posture or "back" chairs have long been accepted as a way to relieve your back from the pain of cramping or kinking up due to this stress. Our posture chair is designed to allow your body to properly align itself by shifting weight from your lower back to your knees, making long work periods far more comfortable. *The Lifestyle Resource* is one of the first to introduce the Criss-Cross Posture Chair. This new and innovative design with a criss-crossed base provides added stability, and an easy exit from your seat. The chair is crafted of laminated hardwood with an attractive oak finish. The seat and kneepad are thickly padded and fully upholstered in a classic neutral grey fabric. The swiveling, carpet-friendly, double wheel casters make maneuvering a breeze. Suited for use at any workstation, desk or table. Requires minimal assembly; a screwdriver is all you need. **$79.95** #2260.

BRING MOUNTAIN TOP FRESHNESS INDOORS

The new Bionaire 700, no bigger than a table-model radio, will clean and recharge your stale indoor air to virtually mountain top freshness. Get relief from breathing allergy causing dust, pollen, tobacco smoke, animal hair and dander, cooking odors, soot, and mold spores. The Bionaire 700 will clean and rejuvenate the air in a 12x12x8 ft. room 4 times an hour, while the filtering system removes up to 99% of all particulate pollutants as small as .01 microns in the process. The filtering process begins with an activated carbon pre-filter that helps remove odors and large particles; next, with the patented electret main filter, the Bionaire removes particles as small as 1/10,000th the diameter of a human hair. Finally, Bionaire's unique negative ion generator—which not only precipitates any remaining particles, but also generates millions of negative ions—reproduces the stimulating effect of fresh mountain air. Manufacturer's 2-year limited warranty, UL listed, weighs 5.2 lbs., **$149.95** #2070. Set of two replacement filters **$19.95** #2071.

ZONE OF CALM

Noise produces stress in the human organism. Today the volume of civilization nearly everywhere seriously interferes with our abilities to relax, read, sleep, concentrate or otherwise function at optimum efficiency. But you needn't be victimized by noise pollution. The *new* Marsona Sound Conditioner electrically synthesizes a variety of pleasing natural sounds that help mask out and reduce the annoyance of unwanted noise. A 5 inch speaker brings you the sounds of ocean surf, summer rain, mountain waterfalls — as close to nature as you will find. You control not only the volume but also wave pattern, wave or rain rhythm, the seeming nearness or distance of the source. Solid state electronics designed for continuous use. UL listed. Manufacturer's 1 year limited warranty. **$139.95** #2201.

THE LIFESTYLE RESOURCE™

FOR FASTEST SERVICE — CREDIT CARD ORDERS CALL TOLL-FREE 24 HRS A DAY **800-872-5200**

DEPT. SFAD09; 921 EASTWIND DR. SUITE 114: WESTERVILLE, OH 43081

ITEM NO.	QTY.	DESCRIPTION	ITEM PRICE	TOTAL PRICE

SEND TO (PLEASE PRINT)

ADDRESS

CITY STATE ZIP

☐ CHECK OR MONEY ORDER ☐ MASTERCARD ☐ VISA ☐ AMEX SIGNATURE

ACCT. # EXP. DATE

Shipping Charge covers UPS, handling and insurance for guaranteed delivery.
UPS Second Day available for an additional $7.50 per order.

Up to $20	$ 3.95	$50.01 to $60	$ 7.95
$20.01 to $30	$ 4.95	$60.01 to $70	$ 8.95
$30.01 to $40	$ 5.95	$70.01 to $100	$10.95
$40.01 to $50	$ 6.95	Over $100	$12.95

Canadian residents please call (614) 794-2662 for ordering information.

SUB TOTAL	
SHIPPING (See table at left)	
TOTAL	

ORDER WITH CONFIDENCE

- *Most orders ship within 48 hours of receipt.*
- *Credit Card orders billed only upon shipment.*
- *No risk 30-day return privilege.*

UPS 2ND DAY AIR

THE MEDIA DEVELOPMENT GROUP

©1989 C22 ADVERTISEMENT

The Turnover of Messenger RNA

The synthesis rate of many proteins is influenced more by how rapidly their messenger-RNA templates decay, or turn over, than by how rapidly the RNA's are synthesized. What regulates the decay?

by Jeffrey Ross

Cells maintain themselves and fulfill their role in an organism by making proteins. Precise regulation of protein synthesis is therefore crucial to the normal functioning of individual cells and the organism as a whole. Let even one protein be overproduced or underproduced and the effect can be devastating. For example, an excess of certain growth factors in the body can cause cells to proliferate abnormally and become cancerous; a lack of the hormone insulin, made by the so-called beta cells of the pancreas, results in diabetes.

In higher organisms proteins are made in the cytoplasm on templates known as messenger-RNA molecules, or mRNA's, which carry information stored in and copied from genes in the DNA of the cell's nucleus. It is now clear that the rate of synthesis of any protein is in general proportional to the quantity of its corresponding mRNA. Hence the factors that control the abundance of mRNA's play a major part in regulating protein levels.

At one time biologists generally assumed that the rate at which an mRNA was synthesized was the major determinant of the amount of the mRNA, and therefore of its protein, in the cell. Yet work done in many laboratories during the past 10 years or so has demonstrated that the rate of mRNA turnover, or degradation, can also significantly affect the rate of protein synthesis. That is, the production of proteins reflects not only how fast their mRNA templates are made but also how fast they are destroyed. Such studies have also shown that controlled changes in the turnover rates of various mRNA's are important to the ability of a cell to replicate and differentiate normally and to respond quickly to changes in its environment.

These and other findings raise an obvious question: What controls the turnover of mRNA? None of us have a complete answer, but what we do know is intriguing. As will be seen, mRNA turnover is determined in part by the structure of each mRNA molecule. Moreover, a host of substances extrinsic to an mRNA itself, including hormones and occasionally even viruses, issue direct or indirect regulatory signals that affect mRNA turnover.

An understanding of the evidence requires a basic understanding of how protein is made and the role of mRNA in that process. When a cell needs to produce a protein, it activates the corresponding gene, which specifies a particular sequence of the amino acids that are the building blocks of proteins. The amino acid sequence is itself encoded in the DNA of the gene as a sequence of nucleotides, which are distinguished by their bases: adenine, guanine, thymine or cytosine. Under the right conditions the genetic information stored in the double-strand DNA is transcribed, or copied, into mRNA, which is an exact copy of a single DNA strand (with the base uracil substituting for thymine).

This RNA transcript is modified in the nucleus (for instance, long nucleotide stretches are excised), after which it is transported to the cytoplasm. There it joins large, globular organelles called ribosomes to form a protein-making factory known as a polyribosome. Each ribosome travels down the mRNA, translating it into protein by stringing together amino acids in the specified order. Eventually the mRNA is degraded, although how quickly that happens can vary from minutes to months.

The fact that the regulation of this last step, mRNA turnover, once drew little attention is not overly surprising. The investigators who initially studied the molecular biology of protein synthesis generally examined proteins encoded by specialized genes: ones expressed (transcribed into mRNA and copied into a protein) only in certain cell types. The steady-state levels of the mRNA transcripts for such proteins are in fact often better correlated with mRNA-synthesis rates than with mRNA-turnover rates.

Not until the early 1980's did a new perspective begin to emerge. In 1984, for example, Ueli Schibler and Mauro Carneiro of the Swiss Institute for Experimental Cancer Research and Roy J. Britten and Eric H. Davidson of the California Institute of Technology examined the relative contributions of mRNA-synthesis and mRNA-turnover rates to the steady-state levels of certain mRNA's representing "housekeeping" genes. Such genes code for proteins crucial to the basic functioning of most or all cells; they include the genes for such well-known enzymes as the DNA polymerase responsible for synthesizing new DNA prior to cell division.

Those investigators found that the steady-state level of each mRNA under study correlated with its half-life—the time it takes to degrade half of the mRNA present—and not with its synthesis rate. If one mRNA was synthesized rapidly and another was synthesized slowly but both were degraded rapidly (had short half-lives), both mRNA's were scarce. Similarly, if a rapidly and a slowly synthesized molecule were both degraded slowly (had long half-lives), they accumulated to high levels.

JEFFREY ROSS is professor of oncology and pathology at the McArdle Laboratory for Cancer Research of the University of Wisconsin at Madison. He earned his undergraduate degree at Princeton University in 1965 and his M.D. from the Washington University School of Medicine in St. Louis in 1969. He then spent a year as a medical intern and was a postdoctoral fellow at the National Institutes of Health for four years before moving to the University of Wisconsin in 1974. In his spare time, Ross says, he aspires toward his dream, "illusory though it may be," of becoming the world's greatest jazz saxophone player.

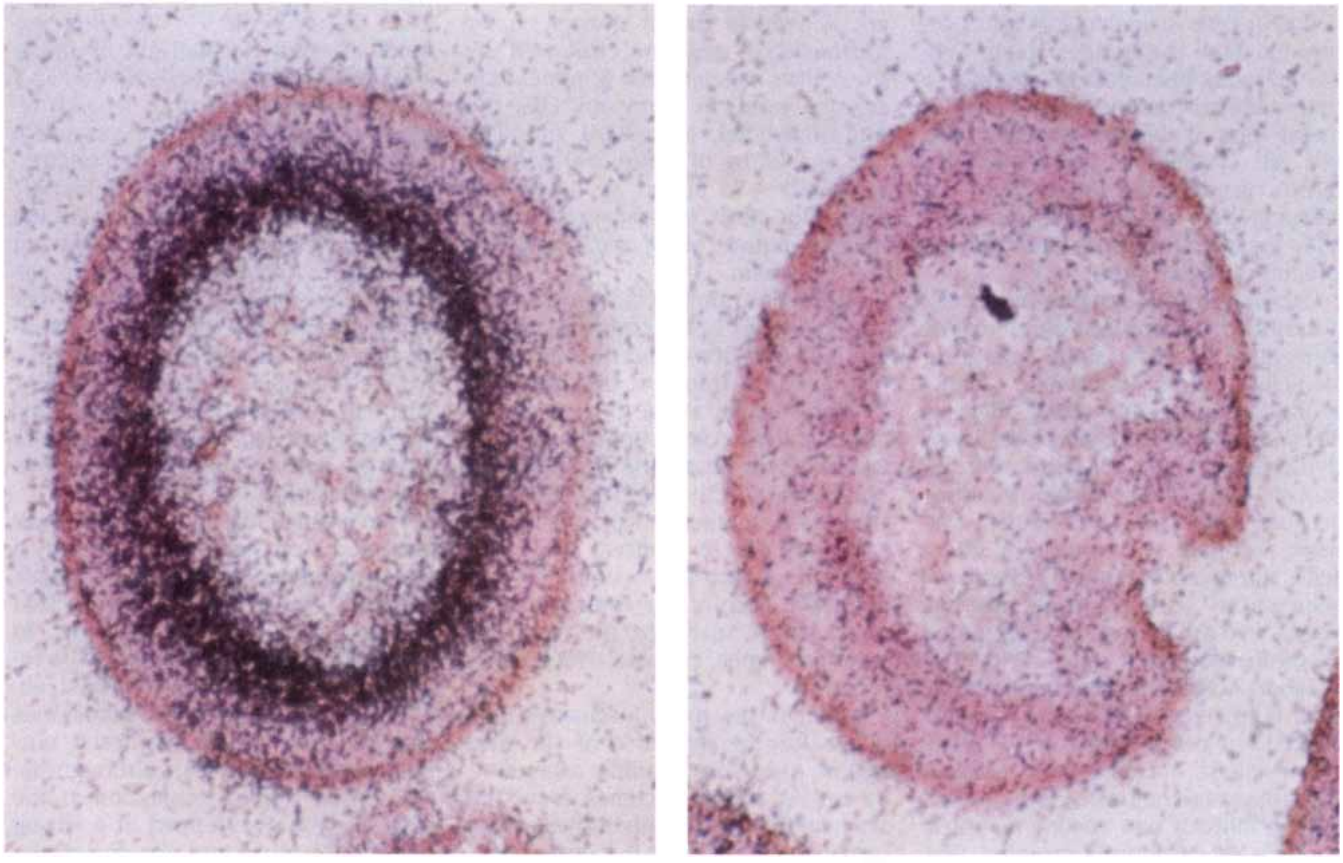
A good way to envision how mRNA turnover can influence mRNA levels in a cell is to imagine trying to fill a leaky bucket with water. As water flows into the bucket, some fraction also escapes through the holes. Eventually a steady-state level of water is reached, as the amount of water going in comes to equal the amount going out. That level depends both on the rate at which water pours into the bucket (gene transcription) and on the size of the holes (RNA turnover). One could run the water full blast, but if the holes were large (turnover was great), little water (mRNA) would accumulate in the bucket (cell). If the holes were small, the water would reach a higher steady-state level.

In terms of cell biology, it is easy to understand the advantage to the cell of regulating mRNA degradation. Consider the differentiation of unspecialized bone-marrow cells into red blood cells. The main function of red cells is to carry oxygen, and so approximately 95 percent of the protein inside a fully differentiated red cell is hemoglobin (the oxygen transporter). In order to accumulate that much hemoglobin the unspecialized parent cell of the red cell must begin abundantly producing the protein components of hemoglobin and at the same time must decrease the production of most other proteins. The cells could reduce the synthesis of unneeded proteins by halting the synthesis of other mRNA's alone, but then long-lived mRNA's would still continue to produce unwanted proteins for some time.

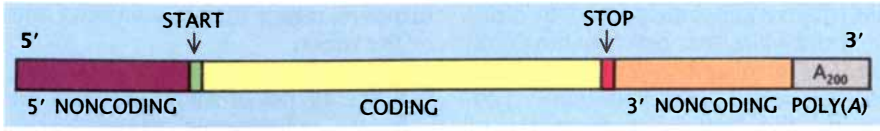
Instead the cells employ a double strategy, according to studies done separately by Haim Aviv of the Weizmann Institute of Science in Israel and George Brawerman of Tufts University and their colleagues. In addition to halting mRNA synthesis the unspecialized parent cells also minimize the activity of the remaining nonhemoglobin mRNA's by destabilizing, and therefore rapidly eliminating, most of them. These and other experiments indicate that a cell can regulate protein production most efficiently by varying both mRNA synthesis and turnover, rather than just varying one or the other.

The agents of mRNA destruction are cytoplasmic ribonucleases (RNA-degrading enzymes) acting in concert with other factors that bind to mRNA molecules. Unfortunately not much is known about these enzymes. For instance, most cells contain 10 or more different ribonucleases, but which of them degrade mRNA's and which degrade other types of RNA molecules is not clear. Moreover, no one knows whether certain mRNA-degrading ribonucleases act only on specific mRNA's and whether several different ribonucleases act together to degrade individual mRNA's.

Work to address such issues is under way, however. It should be facilitated by a new experimental approach, the cell-free system, in which mRNA turnover is studied not in whole cells but in test-tube mixtures that include the protein-making polyribosomes of the cell. My group at the McArdle Laboratory for Cancer Research of the Uni-



MESSENGER-RNA MOLECULES (mRNA's) encoding the protein Drosophila *abl* are abundant in the blastula stage of an embryo of the fruit fly *Drosophila melanogaster* (*left*), as is indicated by a radioactive label (*black dots*). Yet the mRNA's are sparse in an embryo just 30 minutes older (*right*), in the gastrula stage, indicating that fruit-fly embryos rapidly degrade Drosophila *abl* mRNA's early in development. Because the synthesis rate of a protein is generally proportional to the amount of its mRNA present, eliminating selected mRNA's can help a cell to alter quickly the array of proteins it synthesizes. Factors regulating mRNA turnover can include hormones, cellular growth factors and viruses. Mark Henkemeyer and F. Michael Hoffmann of the McArdle Laboratory for Cancer Research at the University of Wisconsin at Madison made the micrographs.



MOLECULE of mRNA typically includes four segments, each of which can influence mRNA turnover. They are the 5′ noncoding segment, the coding segment, the 3′ noncoding segment and, in many mRNA's, the poly(*A*) tail, which consists of a string of approximately 200 adenine (*A*) nucleotides. (Nucleotides, the building blocks of both DNA and RNA, are distinguished by their nitrogen-rich bases, of which adenine is one.) Only the coding segment is translated into protein. Such translation proceeds in the 5′ to 3′ direction, from the "start" signal (*green*) to the "stop" (*red*).

versity of Wisconsin at Madison isolates the polyribosomes by breaking open cells and spinning the resulting cell lysate in a centrifuge. After that initial step ribonucleases are usually found associated with the polyribosomes. We can then study the turnover patterns of the mRNA's in those polyribosomes.

To examine the effect of individual ribonucleases, we can solubilize, or free, the enzymes (by exposing the polyribosomes to a solution high in salt) and separate and purify them by standard techniques. Once we have purified an enzyme, we can then incubate it with various mRNA's and determine how it degrades them. Cell-free systems also make it possible to determine how various agents, such as metal ions or certain proteins from the cell cytoplasm, might influence mRNA turnover; we simply allow the ribonucleases to remain associated with the polyribosomes and add the substance under study.

Similar cell-free systems have been described by Lawrence I. Slobin of the University of Mississippi School of Medicine and by Kathryn L. Calame of the Columbia University College of Physicians and Surgeons. We know that the systems are a reasonable approach to studying mRNA turnover because mRNA half-lives in such systems correlate with those in intact cells. Moreover, the mRNA's that have been studied both in cell-free systems and in vivo are known to be degraded by similar pathways.

Robert A. Kratzke, a former postdoctoral fellow in my laboratory, has exploited our cell-free system to purify and evaluate a cytoplasmic enzyme that appears to be involved in degrading the mRNA's that encode histones: a group of structural proteins that bind to and organize chromosomal DNA in the cell nucleus. Both in cells and in our cell-free system, histone mRNA's decay a bit at a time, beginning at one end, like stalks of celery being nibbled away. Kratzke saw the same pattern of decay when he added his purified enzyme to histone mRNA in a cell-free system.

Among the factors that influence the rate at which enzymes degrade mRNA's is the structure of each mRNA itself. Messenger RNA's can vary in size, but their basic organization is fairly standard. The molecules are conventionally described in terms of the direction in which they are translated into protein. The end closest to the first translated codon, or coding unit of three nucleotides, is designated 5′ and the other end is designated 3′. The regions between are known as the 5′ noncoding, or untranslated, segment, followed by the coding segment, the 3′ noncoding segment and, in many mRNA's, the polyadenylated, or poly(*A*), tail. The tail, consisting of about 200 adenine nucleotides, is added to the mRNA transcript as it is processed in the nucleus before being transported to the cytoplasm.

The region most important to protein synthesis is the coding segment. It is here that ribosomes do their work, beginning to translate the mRNA at a "start" signal at the 5′ end of the segment and completing the protein when they reach a "stop" signal at the 3′ boundary of the segment. The nontranslated segments can affect protein synthesis in indirect ways, but their functions are not completely understood. Each of the four segments of an mRNA, including the three untranslated ones, can also affect mRNA turnover.

Some of the most fascinating evidence for the influence of the first region—the 5′ noncoding segment—on mRNA turnover comes from the study of the cancer-causing role of the gene c-*myc*. This gene is a proto-oncogene: it encodes a protein that normally helps cells to replicate properly but can also transform healthy cells into cancerous ones if the protein is altered or produced in excess.

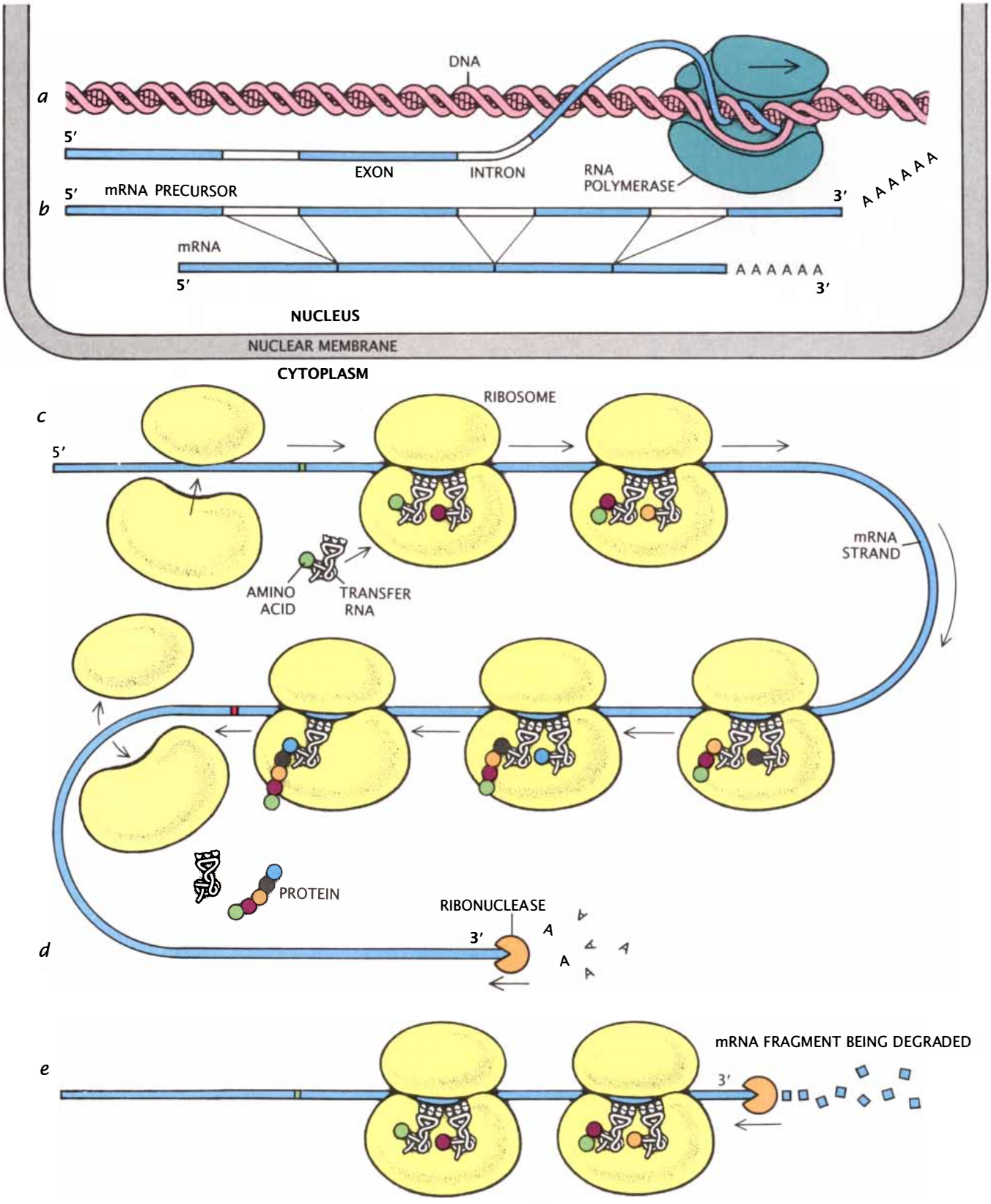
Christian Dani, Philippe Jeanteur and their colleagues at the University of Montpellier in France showed in 1984 that normal c-*myc* mRNA transcribed from a normal c-*myc* gene is relatively unstable, with a half-life of approximately 10 minutes. Then Marc Piechaczyk, Jean-Marie Blanchard and Jeanteur at Montpellier and, independently, Kenneth B. Marcu of the State University of New York at Stony Brook found that mRNA's transcribed from mutated c-*myc* genes in cancerous lymph-node cells were truncated: they had the normal c-*myc* coding, 3′ noncoding and poly(*A*) segments but not the usual 5′ noncoding segment. The resulting protein was normal, but the half-life of the truncated mRNA was prolonged; it was three to five times stabler than full-length mRNA.

Those results demonstrate unequivocally that the nucleotide sequences in the 5′ noncoding segment can in some cases influence mRNA turnover. In the case of the c-*myc* mRNA, the alteration of the 5′ noncoding segment and the concomitant prolongation of the mRNA half-life might have induced the lymph-node cells to overproduce the c-*myc* protein, replicate abnormally and become cancerous. The mechanism by which the 5′ noncoding segment affected the half-life of the c-*myc* mRNA is not clear, however.

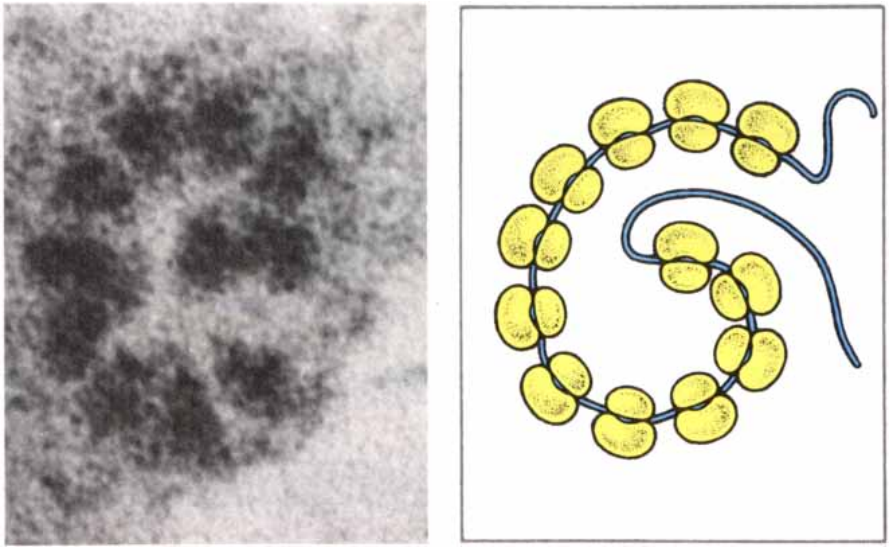
The influence of the second region—the coding segment—on mRNA turnover has also been demonstrated convincingly, this time in histone mRNA's. William F. Marzluff of Florida State University, Nathaniel Heintz of Rockefeller University and their colleagues altered histone genes so that when the mutated genes were inserted into cells and transcribed, mRNA's with a shortened or lengthened coding segment resulted. The workers then compared the half-lives of the mRNA's with the half-lives of normal histone-gene transcripts.

The normal histone mRNA's, their start and stop signals separated by a coding sequence of the normal length, formed normal polyribosomes that yielded normal histone proteins. In contrast, the mutated mRNA's, with repositioned stop signals, formed abnormally configured polyribosomes that were not able to synthesize normal histone molecules. (Although the ribosomes started translation at the right place, they stopped at a wrong one.) More important for the analysis of mRNA turnover, Marzluff and Heintz found that the half-lives of the mutated mRNA's were at least twice as long as the half-lives of normal histone mRNA transcripts.

Clearly, then, the structure of the coding segment, in this case its size,



LIFE CYCLE of an mRNA begins in the nucleus of the cell when a gene is transcribed from double-strand DNA into single-strand RNA (*a*). Enzymes, including an RNA polymerase, separate the DNA strands and form an RNA copy of one strand. The poly(*A*) tail is then added (in most cases) and stretches of nonessential nucleotides (introns) are excised (*b*), after which the mRNA migrates to the cytoplasm. There it joins organelles called ribosomes to form a protein-making factory called a polyribosome (*c*). One after another, each ribosome travels along the mRNA and interacts with transfer-RNA molecules to make a chain of amino acids in the sequence specified by the mRNA's coding region. After a time that varies with the species of mRNA and the conditions in the cell, the mRNA molecule is degraded by one or more enzymes known as ribonucleases. The poly(*A*) tail is probably degraded first (*d*), after which many mRNA's are further degraded in a 3′ to 5′ direction (*e*).



POLYRIBOSOME from the liver of a rat includes 12 ribosomes (*black spots*). The mRNA strand itself, which is indicated in the map (*right*), is not visible in the electron micrograph. Henry Pitot of the McArdle Laboratory provided the micrograph.

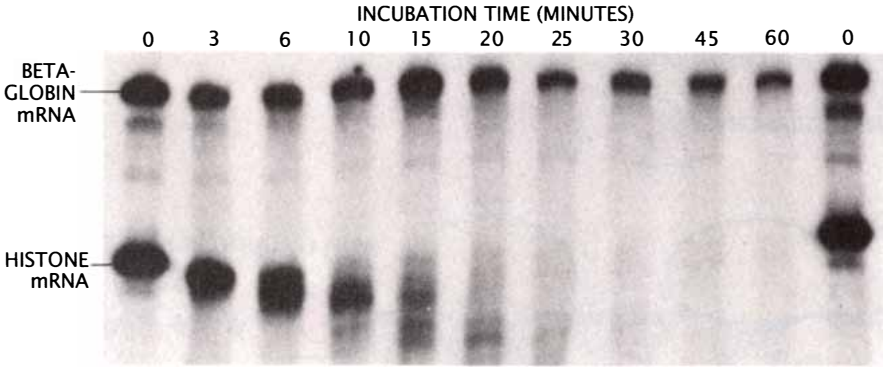
can influence the turnover of this mRNA, but again the mechanism is open to conjecture. One hypothesis suggests that a ribonuclease somehow recognizes whether the mRNA is in its normal "polyribosome context," that is, whether the polyribosomes are attached to the transcript in their usual number and positions. Moving the stop signal would surely affect both factors. According to this hypothesis, the enzyme degrades histone mRNA at the proper rate only if it somehow discerns that both the number and the relative positions of the ribosomes attached to the transcript are "correct." Exactly how a ribonuclease might accomplish such recognition cannot be explained until it is known what polyribosomes look like in three dimensions.

Two kinds of experiments have demonstrated that the next segment, the 3′ noncoding segment, also affects the turnover of certain mRNA's. One approach compares the half-lives of closely related, naturally occurring mRNA species in the same cell type. For example, Aixa Alfonso-Pizarro, a former graduate student of mine, and I measured the half-lives of beta-globin and delta-globin mRNA in human bone-marrow cells. These mRNA's code for the protein components of two types of hemoglobin, and their coding segments and 5′ noncoding segments are quite similar. We found that delta-globin mRNA is degraded some four times faster than beta-globin mRNA. Since the two molecules differ mainly in their 3′ noncoding segment, we suspect the difference in the mRNA-turnover rates is related to differences in the structures of those regions.

Several investigators, including Janet L. and Gary S. Stein of the University of Massachusetts, Richard Treisman of the Medical Research Council Laboratory of Molecular Biology in Cambridge, England, and Gray Shaw and Robert Kamen of the Genetics Institute, Inc., in Cambridge, Mass., compared degradation rates by a more direct approach, involving the construction of chimeric, or hybrid, genes. When these altered genes were inserted into the nucleus of cells, they produced chimeric mRNA's that were normal except that the usual nucleotide sequences of their 3′ noncoding segments were replaced by sequences from the 3′ noncoding segments of other mRNA's. The rate at which each chimera was degraded depended on the source of the transplanted segment. If the 3′ noncoding segment came from a stable mRNA, the chimeric mRNA was stable; if the segment came from an unstable mRNA, the stability of the molecule was reduced.

A tantalizing clue has emerged recently that might help to explain how this part of an mRNA molecule can influence the molecule's turnover. Shaw and Kamen have found that the stability of some mRNA's is affected by the percentage of adenine and uracil nucleotides in the 3′ noncoding segment. The higher the percentage is, the faster the mRNA's are degraded, suggesting that some ribonucleases (or other binding proteins that facilitate ribonuclease activity) recognize regions that are rich in adenine and uracil. In possible support of this notion, these and other workers have discovered that a specific sequence of the two bases is common in the 3′ noncoding segment of many short-lived mRNA's, including the ones encoding several growth factors and proto-oncogene proteins.

The final segment of the mRNA molecule, the poly(*A*) tail, has been shown definitively to influence mRNA turnover as well. Georges Huez of the Free University of Brussels, Uri Nudel of the



DEGRADATION of two mRNA's in a cell-free system is traced in gels. The mRNA's were incubated with ribonucleases contained in an extract from lysed cells; before incubation (time zero) and after successively longer incubation periods, the mRNA's were separated by gel electrophoresis, which sorts molecules by size. The position of each band indicates the number of nucleotides remaining (shorter molecules travel faster through the gels, here from top to bottom), and the darkness of each band indicates the amount of mRNA still present. The virtual disappearance of histone mRNA after about 20 minutes shows that it was degraded faster than beta-globin mRNA. As time went by, the histone mRNA migrated progressively farther through the gel because it was shortened bit by bit as a result of ribonuclease digestion. The same pattern of histone-mRNA lability and globin-mRNA stability is seen in intact cells, which confirms the value of cell-free systems for studying mRNA turnover.

Weizmann Institute and their co-workers removed this segment from an mRNA known to be stable, injected the modified transcript into cells and found that most of it was degraded after four to eight hours. In contrast, the unmodified mRNA, with its intact poly(*A*) tract, was stable for a minimum of 60 hours. This finding, together with work done in other laboratories, suggests that the poly(*A*) tail may have to be removed before an mRNA molecule can be degraded.

Actually a complex formed by the poly(*A*) tail and a protein that binds to it with high specificity, rather than the tail by itself, may account for the tail's stabilizing effect. The poly(*A*)-binding protein (PABP) was discovered by Günter Blobel of Rockefeller University and independently by Brawerman; it binds at least 100 times more tightly to the poly(*A*) tail than it does to other mRNA segments. Indeed, mRNA's probably do not exist as free mRNA but as ribonucleoprotein complexes: mRNA bound by PABP and other proteins.

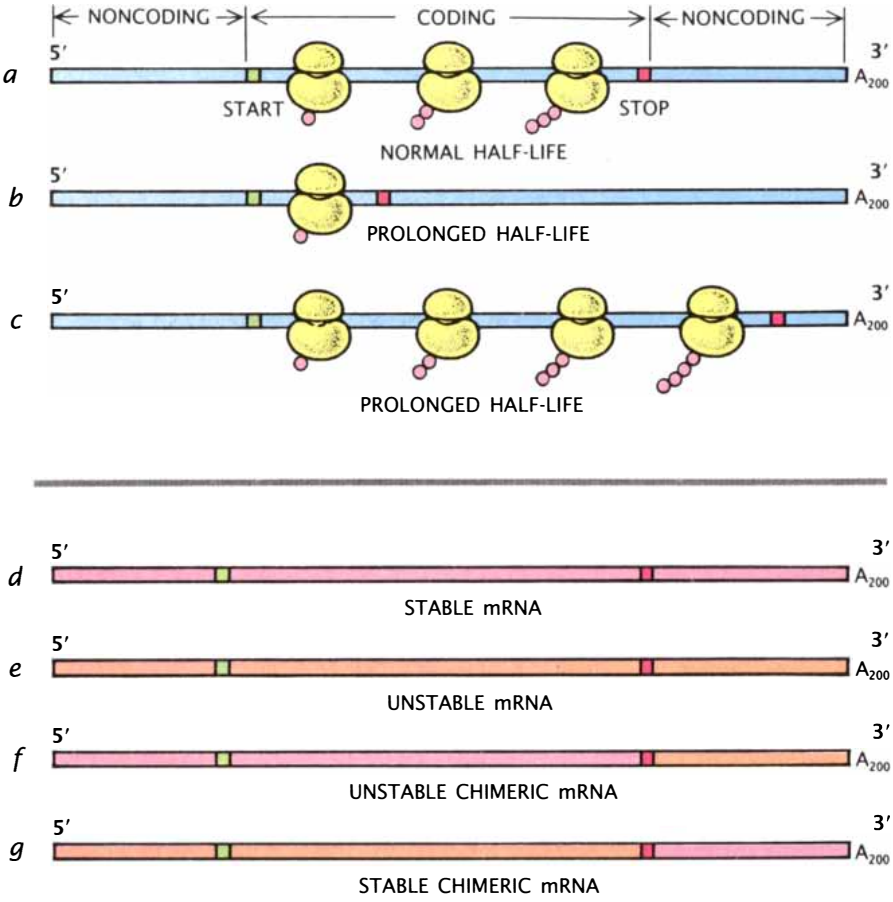
Two graduate students in my laboratory, Philip L. Bernstein and Stuart W. Peltz, and a postdoctoral fellow, Gary Brewer, employed our cell-free mRNA-turnover system to demonstrate the potential importance of PABP to mRNA stability. They began their three-stage investigation by identifying an mRNA that had a poly(*A*) segment and was highly stable in its native form as a ribonucleoprotein complex. They then separated the binding protein from the complex and found that the PABP-free mRNA was degraded rapidly, indicating that even though the nucleotide sequence of the mRNA was unchanged, the molecule became unstable in the absence of PABP. Next they reconstituted the ribonucleoprotein complex by mixing the PABP-free mRNA with purified PABP that was provided by Alan B. Sachs and Roger D. Kornberg of Stanford University. Once the PABP became bound to the poly(*A*) segment, the reconstituted complex stabilized.

PABP apparently helps to stabilize mRNA by buffering the poly(*A*) segment against attack by ribonucleases. But it is important to learn just why PABP does not protect every polyadenylated mRNA to an equal extent. One hypothesis suggests that the other three segments of the mRNA molecule affect how tightly PABP binds to the poly(*A*) tract and hence how efficiently the binding protein protects the mRNA from ribonuclease attack.

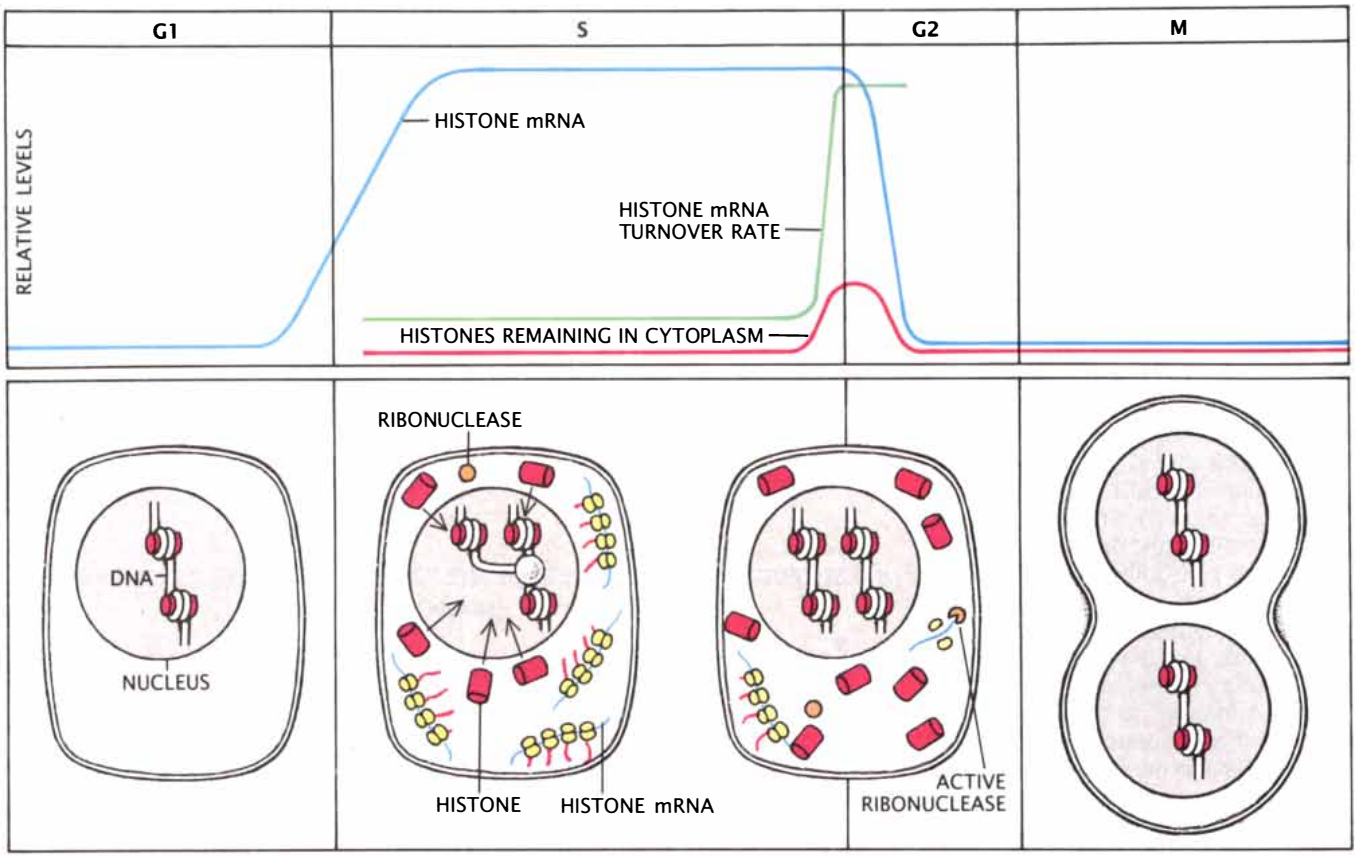
Not only can any of the four segments of an mRNA molecule (or, more properly, a ribonucleoprotein complex) potentially influence the molecule's stability, but also it appears that two or possibly more segments within the same mRNA can influence turnover. For example, Michael D. Cole of Princeton University found that deleting parts of the 3′ noncoding segment of the c-*myc* mRNA has the same effect as deleting the 5′ noncoding segment, namely prolongation of the half-life. This finding alone raises several questions: How do the 5′ and 3′ segments interact to determine the overall half-life of c-*myc* mRNA? Does the mRNA fold, juxtaposing its segments so that they interact, or do the segments function independently? Do nucleotide sequences in the coding segment also affect the turnover of c-*myc* mRNA? To what extent do the various segments of other mRNA's interact with one another?

Some way to map the three-dimensional structure of mRNA's, both on and off polyribosomes, would certainly help to answer such questions and to explain why different mRNA molecules have different turnover rates. Yet even then the analysis of mRNA turnover would be incomplete: an understanding of the broad range of extrinsic signals that can directly or indirectly influence mRNA degradation is also needed. Several extrinsic influences, such as the viruses and hormones alluded to above, have already been identified.

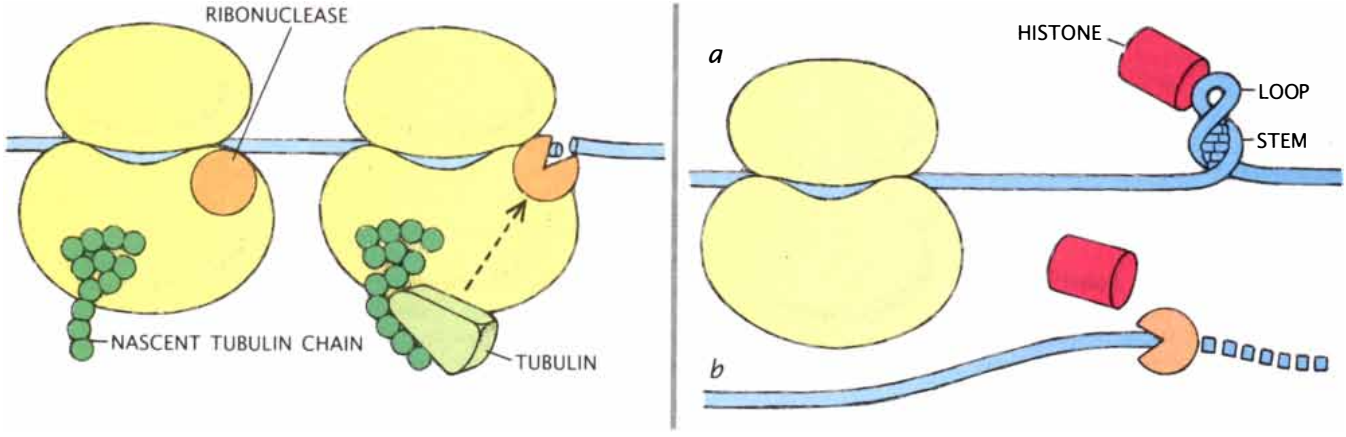
Among the viruses that can dramatically affect mRNA turnover is herpes simplex, which causes cold sores and some genital infections. Like many other viruses, it captures the protein-making machinery of infected cells to replicate itself. Several investigators,



VARIED APPROACHES have demonstrated the influence of individual mRNA segments on mRNA turnover. As is shown at the top, workers have demonstrated the importance of the coding region by moving the stop signal (*red*) of normal mRNA's (*a*), thereby shortening (*b*) or lengthening (*c*) the coding region. Both changes significantly prolonged the half-life of the mRNA, perhaps by altering the number and spacing of ribosomes attached to the region. Investigators have also studied the regulatory role of each mRNA segment by combining segments (*bottom*) from stable (*d*) and unstable (*e*) molecules into "chimeric" mRNA's. For instance, replacement of the 3′ noncoding region of a normally stable mRNA with that from an unstable mRNA made the molecule unstable (*f*). Similarly, insertion of the 3′ noncoding segment from a stable molecule into a basically unstable mRNA prolonged the mRNA's half-life (*g*).



PROPOSED MODEL to explain changes in the turnover rate of the mRNA's coding for histone (which organizes chromosomal DNA) at different stages of the cell-replication cycle suggests that excess histone regulates the destruction of its own mRNA. As is shown by a graph (*top*) and schematic diagrams (*bottom*), in the G1 stage, before DNA is replicated for cell division, little histone protein is needed and little of its mRNA is made. In the next stage (S) DNA is synthesized and histone is required; hence the histone gene is turned on, histone mRNA is synthesized, the histone-mRNA level rises and the mRNA's direct the production of abundant histone protein. The histone level in the cytoplasm remains fairly constant, however, because the protein is taken up by the nucleus as quickly as it is made. Near the end of the synthesis stage, when DNA duplication is almost complete, the nucleus reduces its uptake of histone, which collects in the cytoplasm. The autoregulatory model suggests that the excess histone binds specifically to, and alters the structure of, its own mRNA, thereby making the mRNA highly accessible to one or more ribonucleases and increasing its turnover. The result is a drop in the levels of histone mRNA and protein in the G2 stage. The low levels continue as the cell goes through mitosis (M) and the new cells enter G1.



MECHANISMS by which the proteins tubulin and histone may trigger the destruction of their own mRNA have been proposed. Don W. Cleveland and his colleagues at the Johns Hopkins University School of Medicine suggest that free tubulin binds to the first four amino acids of a nascent chain of tubulin emerging from a ribosome (*left*). Such binding may send a signal through the ribosome to an associated ribonuclease, which is then activated to attack the ribosome-bound mRNA. In histone mRNA (*right*), which has no poly(*A*) tail, part of the 3′ noncoding region normally forms a loop and a double-strand stem that partially protects the mRNA from attack by ribonucleases (*a*). The author and his colleagues suggest that binding by free histone may somehow "melt" the double strand (*b*), straightening the mRNA and making it more accessible to ribonucleases.

including Michael L. Fenwick of the University of Oxford, Yutaka Nishioka and Saul J. Silverstein of Columbia and Niza Frenkel, G. Sullivan Read and Ann D. Kwong, then at the University of Chicago, have demonstrated that the virus appropriates the cell's protein-synthesis machinery by eliminating competing demands on the machinery. It does so by accelerating the degradation of virtually every cellular mRNA, even ones that are normally very stable, so that enough ribosomes will be free to associate with viral mRNA's and produce viral proteins. Perhaps the virus destabilizes mRNA's by activating ribonucleases.

One hormone that influences mRNA turnover is estrogen. David J. Shapiro and Martin L. Brock of the University of Illinois at Urbana-Champaign demonstrated that this hormone, which is known primarily for its role in the female reproductive cycle, stimulates the production in amphibians of vitellogenin (a protein made in the liver). It does so both by increasing the rate at which vitellogenin mRNA is synthesized and by stabilizing the mRNA. The half-life of vitellogenin mRNA in estrogen-treated cells is three weeks, compared with 16 hours for the same mRNA in untreated cells. No one knows how estrogen specifically stabilizes the vitellogenin mRNA, but one theory suggests that the hormone binds to—and thereby interferes with—a ribonuclease responsible for degrading that transcript. Other hormones known to regulate mRNA turnover include growth hormone, which stabilizes the mRNA of the pituitary hormone prolactin (responsible for stimulating lactation), and cortisone, which increases the stability of growth-hormone mRNA.

Not all the extrinsic molecules that regulate the turnover of mRNA are viral products or hormones. For instance, glucose, a sugar that circulates in the bloodstream, helps to stabilize insulin mRNA in pancreatic cells. The resulting increase in insulin production helps to prevent glucose levels from rising too high, because insulin facilitates the transport of glucose from the blood into cells.

In an unknown number of cases, a protein may actually regulate the level of its own mRNA, a process known as autoregulation. The DNA-binding histones, for example, are thought to increase the turnover of their mRNA transcripts at a stage in the cell-replication cycle when the cell does not need to have large quantities of histone available.

Many groups have observed that histone-mRNA molecules are abundant only during the synthesis (S) phase of the cell cycle, when DNA is duplicated in preparation for mitosis (cell division) and large amounts of histone are needed to bind to the newly made DNA. Histone mRNA's appear in the cytoplasm in measurable quantities at or just before the start of DNA replication. By the middle of the synthesis phase they are at least 50 times more abundant than they were in the immediately preceding resting phase, known as the G1, or gap, phase. Once DNA replication slows and stops, the cell no longer needs new histones; hence the transcription of histone genes into mRNA slows and the turnover of histone mRNA accelerates.

The autoregulation hypothesis suggests that the rapid elimination of histone mRNA is triggered to a great degree by free, unused histones remaining in the cytoplasm at the end of the DNA-synthesis phase. Peltz and I tested this idea with our cell-free system. We found that histone-mRNA turnover quadrupled when we added histones to polyribosomes isolated from cells. The added histones had no effect on other mRNA molecules, and other proteins had no effect on histone mRNA's, indicating that histones can indeed induce rapid and highly specific destruction of their own mRNA. We know free histone proteins do not themselves degrade mRNA, and we have suggested that they bind to the 3′ end of their mRNA—the part known to degrade first. In so doing they may somehow make the 3′ region more accessible to one or possibly more ribonucleases.

Tubulin is another protein that regulates its own mRNA, according to studies led independently by Sheldon Penman of the Massachusetts Institute of Technology, Marc W. Kirschner of the University of California at San Francisco and Don W. Cleveland of the Johns Hopkins University School of Medicine. Tubulins are housekeeping proteins that aggregate to form microtubules, intracellular fibers important to both the structural integrity of cells and the separation of chromosomes during mitosis.

Cleveland and his co-workers also suggest how such tubulin-specific autoregulation might be accomplished. Apparently free tubulin interacts with a unit of four amino acids at the tip of a nascent tubulin protein as the amino acid chain emerges from an actively translating ribosome. The interaction may somehow send a signal to the ribosome, activating an associated ribonuclease, which would then degrade the mRNA molecule.

My laboratory and others have identified some of the factors that influence mRNA turnover, but in most cases we do not know how they exert their effects. What kinds of molecules convey the signals sent by external factors, such as hormones and growth factors, and do these intermediaries act on the mRNA molecule itself or on ribonucleases? How many enzymes are necessary to degrade each mRNA molecule? Those are some of the biggest questions investigators face right now.

Once such questions and others are answered, those of us interested in the role of mRNA turnover in protein regulation can explore new ways to alter the growth and metabolism of various cells by altering the turnover of selected mRNA's, first in the laboratory and then in whole organisms. Eventually we might even hope to apply what we learn in the laboratory to therapies for some of the more challenging diseases of our time. I can imagine, for instance, that certain viral infections may someday be treated with agents that prevent the viruses from degrading cellular mRNA's. We might even one day devise a way to destabilize specifically the mRNA's transcribed from certain oncogenes and in so doing to halt the abnormal proliferation of cells and slow or stop the growth of cancers.

FURTHER READING

ACCUMULATION OF RARE AND MODERATELY ABUNDANT mRNAs IN MOUSE L-CELLS IS MAINLY POST-TRANSCRIPTIONALLY REGULATED. Mauro Carneiro and Ueli Schibler in *Journal of Molecular Biology,* Vol. 178, No. 4, pages 869–880; October 5, 1984.

A CONSERVED AU SEQUENCE FROM THE 3′ UNTRANSLATED REGION OF GM-CSF mRNA MEDIATES SELECTIVE mRNA DEGRADATION. Gray Shaw and Robert Kamen in *Cell,* Vol. 46, No. 5, pages 659–667; August 29, 1986.

REGULATION OF MESSENGER RNA STABILITY IN EUKARYOTIC CELLS. David J. Shapiro, John E. Blume and David A. Nielsen in *BioEssays,* Vol. 6, No. 5, pages 221–226; May, 1987.

AUTOGENOUS REGULATION OF HISTONE mRNA DECAY BY HISTONE PROTEINS IN A CELL-FREE SYSTEM. Stuart W. Peltz and Jeffrey Ross in *Molecular and Cellular Biology,* Vol. 7, No. 12, pages 4345–4356; December, 1987.

MULTIPLE REGULATORY STEPS CONTROL HISTONE mRNA CONCENTRATIONS. William F. Marzluff and Niranjan B. Pandey in *Trends in Biochemical Sciences,* Vol. 13, pages 49–52; February, 1988.

Quantum Interference and the Aharonov-Bohm Effect

These counterintuitive effects play important roles in the theory of electromagnetic interactions, in solid-state physics and possibly in the development of new microelectronic devices

by Yoseph Imry and Richard A. Webb

Although quantum theory is about a century old, its capacity to produce counterintuitive insights into the nature of matter remains undiminished. One such surprise began with a provocative experiment proposed by Yakir Aharonov and David Bohm in 1959. Imagine, they said, a magnet shielded in such way that it could not exert a force on another magnet nearby. In other words, no conventional manifestation of the shielded magnet's field could be detected. Yet if a beam of electrons were to travel through the vicinity of this shielded magnet, Aharonov and Bohm predicted, the phase of the electron wave function would change. (In quantum theory an electron can be described sometimes as a wave and sometimes as a particle.)

How could one account for the phase change of an electron wave function? Aharonov and Bohm predicted that this effect was due to a physical entity more fundamental than electric and magnetic fields: a potential, whose rate of change over space and time yields the electric and magnetic fields. After three decades the Aharonov-Bohm effect has been demonstrated conclusively in experiments done on electrons traveling through a vacuum, and in the past four years the effect has been observed in very small conducting wires at low temperatures.

YOSEPH IMRY and RICHARD A. WEBB are leaders in theoretical and experimental work on the Aharonov-Bohm effect. Imry is professor of physics at the Weizmann Institute of Science in Israel. He received his Ph.D. from Weizmann in 1966. Webb is manager of Low Temperature Quantum Phenomena at the IBM Thomas J. Watson Research Center in Yorktown Heights, N.Y. He got his Ph.D. from the University of California, San Diego, in 1973.

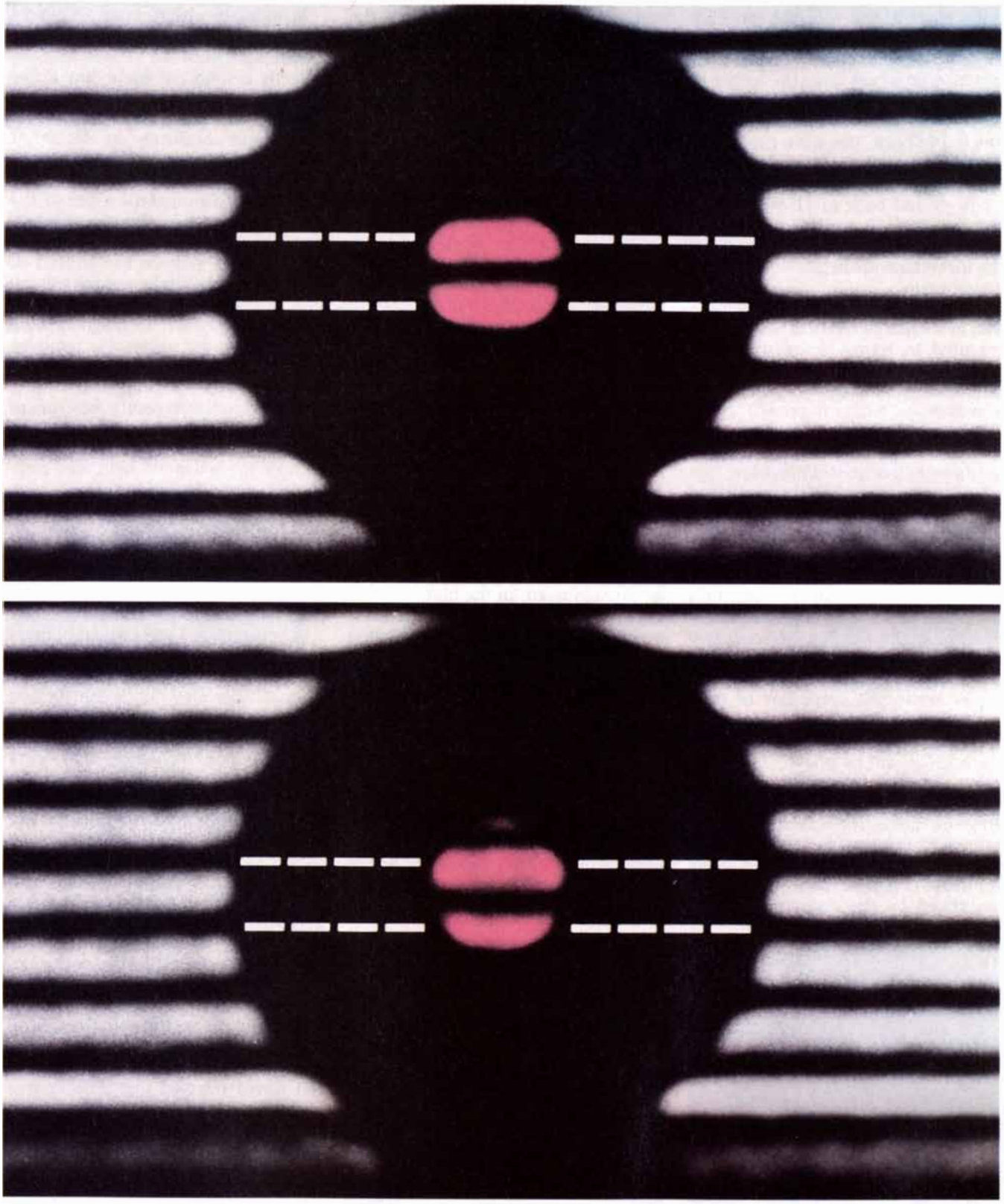
The Aharonov-Bohm effect has had considerable influence on how physicists think about electrodynamics. It has long been known that a positive charge passing close to but in no way touching a stationary, negative charge will nonetheless accelerate and change direction. In order to explain such a phenomenon, known as action at a distance, Michael Faraday proposed in 1846 that charges encounter fields that exert electric and magnetic forces. From what was known at the time, fields described the dynamics of charges completely. When the theories of relativity and quantum mechanics were introduced, the potentials, not the electric and magnetic fields, appeared in the equations of quantum mechanics, and the equations of relativity simplified into a compact mathematical form if the fields were expressed in terms of potentials. The experiments suggested by Aharonov and Bohm revealed the physical significance of potentials: a charged particle that passes close to but in no manner encounters a magnetic or electric field will nonetheless change its dynamics in a subtle but measurable way. The consequence of the Aharonov-Bohm effect is that the potentials, not the fields, act directly on charges.

Physicists have been exploring the effect's broad implications in areas ranging from the quantized Hall effect through superconductivity to superstring theory. In the years ahead the Aharonov-Bohm effect may even have a profound impact in electronics. By the end of the century manufacturers hope to be able to produce silicon chips one centimeter square that contain as many as 100 million components. If that number is to be reached or exceeded, a different set of physical principles must be exploited. The Aharonov-Bohm effect may point a way to such a technology.

A two-slit interference experiment provides an elegant demonstration of the wavelike nature of electrons and provides a basis for understanding the Aharonov-Bohm effect. In such an experiment a particle generator emits a beam of electrons all of which have the same energy. The beam is directed at a plate that absorbs electrons. Two narrow, closely spaced vertical slits pierce the plate. They lie just to the left and right of where the beam impinges. Centered behind the slits is a film that records a bright spot every time an electron hits it. After the particle generator has emitted many electrons, a sequence of light and dark bands parallel to the slits emerges on the film. In the center is a bright band that diffuses out on each side into two dark bands. They in turn are flanked by light bands, and so on [*see illustration on page 59*].

How could the electrons produce this pattern? If electrons acted like bullets, they would ricochet off the slits or pass straight through them. The film would, then, record a concentration of impacts directly behind the slits; there would be relatively few hits off to one side or the other. Clearly such an effect cannot account for the complex pattern that is observed.

A better approach (and one that agrees with the behavior quantum theory ascribes to electrons) would be to assume that the particles behave like waves. The feature that characterizes waves as they travel through space and time is an amplitude that varies periodically from a maximum to a minimum and back again. The instantaneous variation in amplitude and



INTERFERENCE PATTERNS demonstrate the Aharonov-Bohm effect. Part of an electron beam passes through a toroidal magnet (*black ring at top*) coated with niobium. The other part of the beam passes outside the toroid. Together the beams cause an interference pattern (*colored region at top*). The background interference pattern outside the ring results from interference among electrons that do not go through the toroid. The interference pattern framed by the toroid is shifted with respect to the background even though the electrons were shielded from the magnetic field. The shielding occurred because the niobium was cooled below 9.1 degrees Kelvin and became superconducting. The shift, as predicted by Yakir Aharonov and David Bohm, is the result of interaction among the electron waves and the vector potential, which is present even in the absence of the magnetic field. When the niobium coating is heated above 9.1 degrees K (*bottom*), it ceases to be superconducting, the magnetic field contained within the toroidal magnet changes and the interference pattern shifts abruptly (*colored region at bottom*). Akira Tonomura and his colleagues carried out this experiment in 1986 at Hitachi Ltd. in Tokyo.

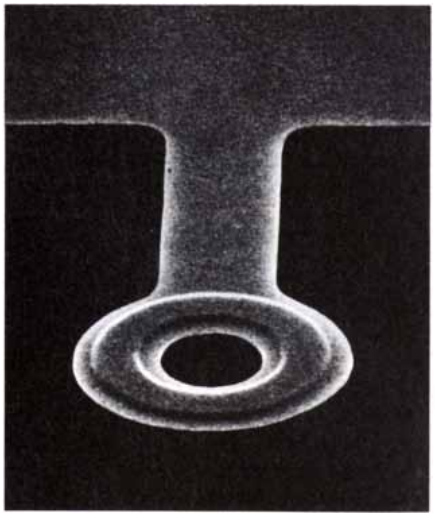
other characteristics of the wave are conveniently described by a mathematical wave function. Consider, for instance, an ocean wave the height of which varies from one meter above the average surface to one meter below it and back. The wave can be described by a cosine function, since the value of the cosine changes from +1 to −1 and back to +1 as its angle changes from 0 to 180 to 360 degrees. The angle that corresponds to the instantaneous height is called the phase angle.

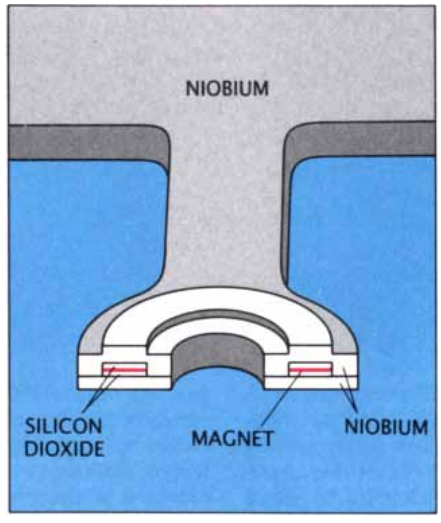
The mathematical wave function that describes an electron wave is represented in terms of its maximum amplitude and phase angle. The amplitude of an electron wave describes a probability, which is related to the fact that the position and velocity of a particle can be determined to within only a certain degree of precision. Specifically, the square of the maximum amplitude of the electron's wave function is the probability of finding the electron at a particular location at a particular time.

The phase angle of an electron's wave function is especially useful for describing the relation between two waves. If two waves are "in phase" at a particular location or time, the two waves are in the same part of their cycle: both have reached maximum or minimum amplitude. If two waves are "completely out of phase," one wave has reached a maximum while the other is at a minimum. The phase angle of an electron wave can also be expressed in terms of more intuitive physical quantities. In simple cases the phase is related to the momentum multiplied by the distance the electron wave has traveled and also to the energy multiplied by the time.

These concepts provide an adequate explanation for the pattern yielded by the two-slit experiment. Since the particle generator emits electrons having the same energy and momentum, the electron wave functions have the same phase at a given distance from the generator—a condition known as coherence. As an electron wave penetrates the two slits, it divides into partial waves. Since partial waves travel the same distance to each slit, the partial wave emerging from the lefthand slit has the same phase as the partial wave emerging from the righthand one. Thus at a point on the film that is equidistant from the two slits, the left and right partial waves will be in phase. Hence the waves reinforce each other and produce a bright band in the middle of the film. It is also fair to say that the bright bands represent the fact that electrons have two times as much chance of striking there as they do at an average point on the film.

To the left of the bright band, however, the right partial wave must travel a greater distance than the left partial wave. Consequently, at some points to the left, the two waves will be completely out of phase and will cancel each other. Hence a dark band will form (because electrons have almost no chance of striking there). At a point still farther to the left on the film, the right wave travels such a distance that it is exactly one full cycle behind the lefthand wave. Once again the waves are in phase and create another bright (high probability) band.





TOROIDAL MAGNET and niobium film employed in the experiments of Tonomura are depicted in the photograph at the left and in the illustration at the right. The magnet, which is five microns across, consists of an alloy containing 83 percent nickel and 17 percent iron. A silicon dioxide coating insulates the magnet from the niobium.

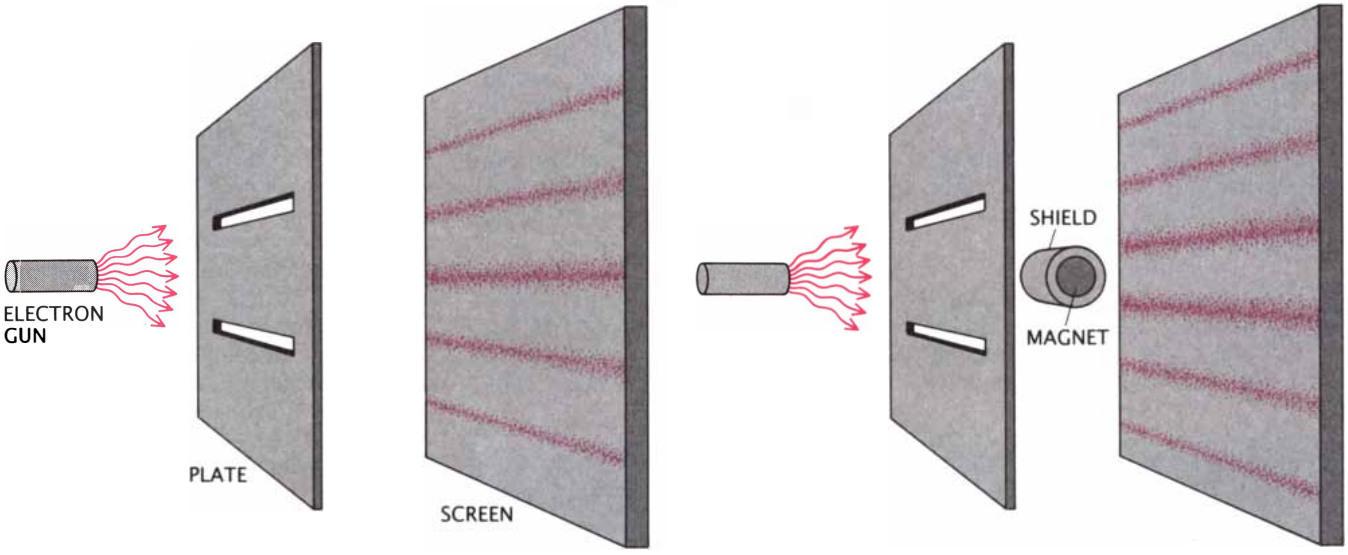
In order to observe the Aharonov-Bohm effect, the two-slit interference experiment must be altered slightly. Directly behind the plate and in between the slits is placed a very long solenoid that has a magnetic field inside it and absolutely no electric or magnetic field outside it. When a beam of electrons now penetrates the two slits and goes around the solenoid, the film records a new interference pattern. Compared with the original pattern, the new pattern has shifted so that previously bright regions will appear darker and dark regions will appear brighter. When the magnetic field contained in the solenoid is removed from the experiment, the interference pattern returns to its original form.

In this new interference experiment the phases of the left and right partial waves apparently changed even though the magnetic field was completely confined inside the solenoid. The change in phase of an electron wave function in a region where no magnetic field exists is one manifestation of the Aharonov-Bohm effect.

The effect revealed that the phase change of a wave function must be related to some physical entity present outside a confined magnetic field. Aharonov and Bohm derived from the fundamental equations of quantum mechanics that the phase change is due to an entity that exists anywhere in and around a magnetic field called the magnetic vector potential. Although the vector potential is a vector field in the sense that it has a magnitude and a direction at every point in space and can change with time, the vector potential can be measured directly only by observing changes in phase of wave functions. The phase shifts caused by the vector potential can account for all measurable magnetic effects on charged particles.

How did the vector potential act on the phase of an electron in the two-slit experiment? As the left and right partial waves traveled in the force-free region near the solenoid, the vector potential changed the momentum of the left partial wave with respect to the right partial wave without changing the kinetic energy. Since the phase of a wave function is related to its momentum, the left partial wave changes phase in relation to the right partial wave.

The magnetic vector potential and the Aharonov-Bohm effect have counterparts in electric interactions. They are the electric scalar potential and the electrostatic Aharonov-Bohm effect. The electric scalar potential is not



TWO-SLIT interference experiment demonstrates the wave behavior of electrons. A particle generator emits a beam of electrons that travels toward two slits in a plate. The electron wave functions pass through the slits, creating an interference pattern (*left*). When a magnet introduces a vector potential field (*right*), the pattern shifts owing to the Aharonov-Bohm effect.

a vector field; it simply has a magnitude at every point in space. Although the absolute magnitude of the potential cannot be determined, the difference in potential between two points is the energy necessary to move a unit of charge from one point to the other. This potential difference is commonly measured in volts.

Like the magnetic vector potential, the electric scalar potential can also cause a phase shift of an electron wave function. The electrostatic Aharonov-Bohm effect is then the phase shift of an electron wave function due to the electric scalar potential in a region where no electric field exists.

The electrostatic Aharonov-Bohm effect can be explained in terms of a thought experiment [*see upper illustration on next page*]. An electron beam is split into two partial waves. Each partial wave is directed into a hollow, metallic cylinder. After the partial waves enter each cylinder, a potential difference is applied between the two cylinders. Before the waves leave, the potential is removed. In this way the waves do not feel an electric force. The total energy difference between the two waves, however, is changed by the charge of the electron multiplied by the difference in potential between the two cylinders. Since the phase of an electron wave function is related to total energy and travel time, the phase of one electron wave is changed with respect to the other.

Soon after Aharonov and Bohm predicted the effects of the potentials on the phase of charged particles, experiments were begun. Robert G. Chambers of the University of Bristol did the first one in 1960. A coherent beam of electrons was generated in an electron microscope and split in two by an aluminum-coated quartz fiber 1.5 microns in diameter. An interference pattern resulted that resembled the pattern produced in the two-slit experiment. When a magnetized iron filament one micron in diameter was placed directly behind the quartz fiber, the pattern shifted. Chambers argued that the magnetic field produced by the filament in the region where the electrons traveled was much too small to explain the magnitude of the observed shift. He concluded that the vector potential must have caused a change in the phase of the electron wave function as predicted by Aharonov and Bohm.

A number of investigators challenged Chambers' conclusion. Since he did not completely confine the magnetic field to a small region of space, some of them maintained that the force exerted by the magnetic field on the electron obscured the contribution of the vector potential to the shift in the interference pattern.

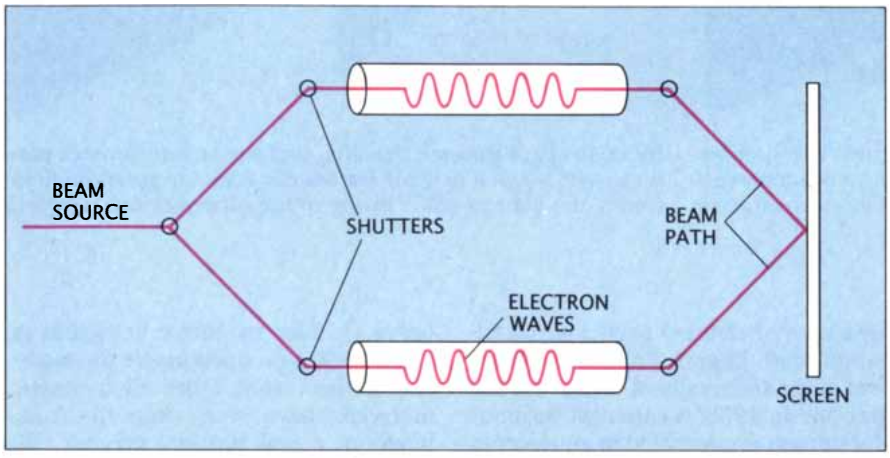
In 1986 Akira Tonomura and his colleagues at Hitachi Ltd. in Tokyo solved the problem of magnetic field confinement, following the suggestion of Charles Kuper of Technion–Israel Institute of Technology. Tonomura knew that a homogeneous magnetic material in the shape of a toroid has a circular magnetic field that is completely contained inside the material of the toroid. Since all magnetic materials have some imperfections, however, a real toroidal magnet will always have some small unconfined magnetic field. This so-called leakage field can be confined by coating the magnet with a superconducting material. Therefore Tonomura constructed a toroidal magnet and coated it with niobium, which superconducts at temperatures below 9.1 degrees Kelvin. This arrangement ensured that less than 1 percent of the field inside the toroidal magnet penetrated into the central hole.

A beam of electrons was generated in a vacuum and split in two so that one beam passed through the hole of the toroid and the other passed outside. Together the beams caused an interference pattern on a film directly behind the hole. A reference interference pattern was created simultaneously. When the niobium was cooled well below 9.1 degrees K and became superconducting, the magnetic field was confined to a specific strength by the laws of superconductivity, and so the vector potential in the hole could attain only a specific strength. As a result the vector potential changed the phase of the electrons traveling through the toroid relative to those traveling outside. This meant that (depending on the particular experimental arrangement) in some cases the interference pattern behind the hole exactly matched the background pattern; in other cases the pattern behind

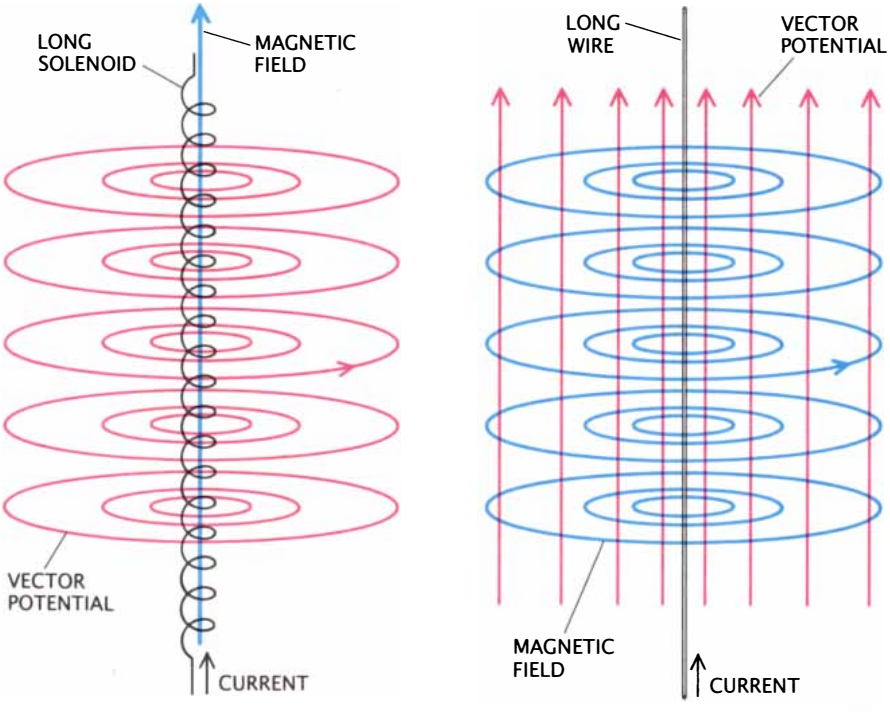
the hole exactly mismatched the reference, that is, the dark bands of one existed alongside the bright bands of the other. Whichever the case, this gave a unique verification of the role of the vector potential in changing the phases of electron waves in a region where no magnetic field exists [*see illustration on page 57*].

When electrons travel through a vacuum, the Aharonov-Bohm effect can be observed because the phase of the electron wave function remains well defined as the wave splits and interferes. The effect is harder to observe in solids, because electrons scatter off various imperfections in the crystal lattice.

Although every solid exhibits some form of scattering, techniques have been developed over the past decade that reduce scattering to the point where electrons travel much as they do through a vacuum [see "Ballistic Electrons in Semiconductors," by Mordehai Heiblum and Lester F. Eastman; SCIENTIFIC AMERICAN, February, 1987].



ELECTROSTATIC Aharonov-Bohm effect can be observed by splitting an electron beam and directing it toward two hollow, metallic cylinders that shield electrons from electric forces. As electrons pass through the cylinders, a scalar potential difference (voltage) is applied between the cylinders. The interference pattern observed on the screen is shifted by an amount directly related to the scalar potential.



VECTOR POTENTIAL FIELD (*red lines*) is compared with the magnetic field (*blue lines*) for a long solenoid (*left*) and a long wire (*right*). Each line represents its respective field at a given strength. The circulation of the vector potential field around a curve is equal to the magnetic field multiplied by the area bound by that curve.

It was, however, the appreciation of two types of scattering in solid conductors—inelastic and elastic—that led to the first discoveries of the Aharonov-Bohm effect and other quantum-interference effects in solid materials.

Inelastic scattering occurs when atoms that make up a solid conductor exchange energy with the electron. Strictly speaking, inelastic scattering alters the wave functions of the atoms making up the solid, that is, scattering induces a change in the quantum state of the environment in which the electron moves. For instance, the electron can absorb energy from or give energy to the vibrations of atoms in a crystal lattice. One key to reducing inelastic scattering is to limit the energy available for such interactions. If enough energy is removed from the crystal lattice and the electron system so that they are essentially quiescent, inelastic scattering will be scarce. The way one can remove this energy is to cool the wire to low temperatures. At quite attainable temperatures of a few degrees Kelvin, electrons in many metals can move across several thousand atoms (a distance of approximately one micron) without undergoing inelastic scattering.

Cooling a solid conductor to low temperatures has another beneficial effect. The range of energies with which electrons travel through a solid decreases as the temperature decreases. At temperatures low enough to make inelastic scattering improbable, the range of energies is so narrowly defined that all electrons traveling through the wire have effectively the same energy. This makes all conducting electrons in the solid interfere in essentially the same way.

Elastic scattering takes place when an electron encounters a static potential, such as an impurity or a defect in the crystal lattice. A static potential changes the phase of the electron wave function in a well-defined manner but not its total energy. Although a random distribution of static potentials in a solid will lead to a random change in phase, the change will be the same for every electron that travels through the solid at a particular energy. As temperature approaches absolute zero, it turns out that an electron wave should encounter only elastic scattering, which leads to a random but constant phase change and does not obscure electron-interference effects in a solid conductor. That was the key to observing quantum-interference effects in solids.

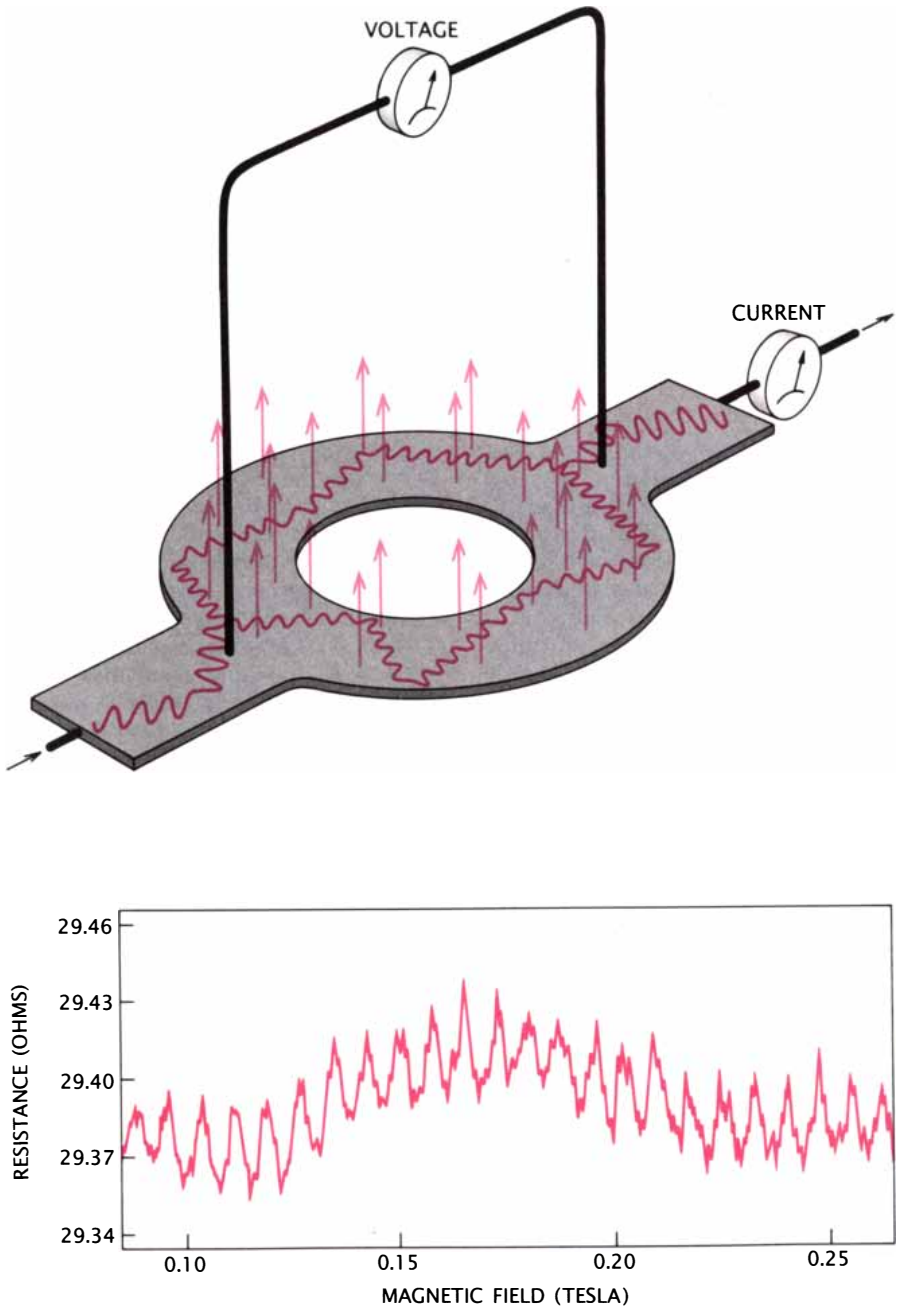
In real experimental systems, however, solid conductors cooled to low

temperatures still exhibit some degree of inelastic scattering that will introduce some uncertainty in the phase of the electron wave function. As the size of the solid conductor decreases, the number of phase-randomizing events decreases. To observe quantum interference the conductor must be sufficiently small to essentially eliminate inelastic scattering. Experiments have shown that, although a metal wire .03 micron thick, .03 micron wide and one micron long contains nearly 100 million atoms, the phase of an electron wave function traveling through the wire will typically be maintained at temperatures below one degree K.

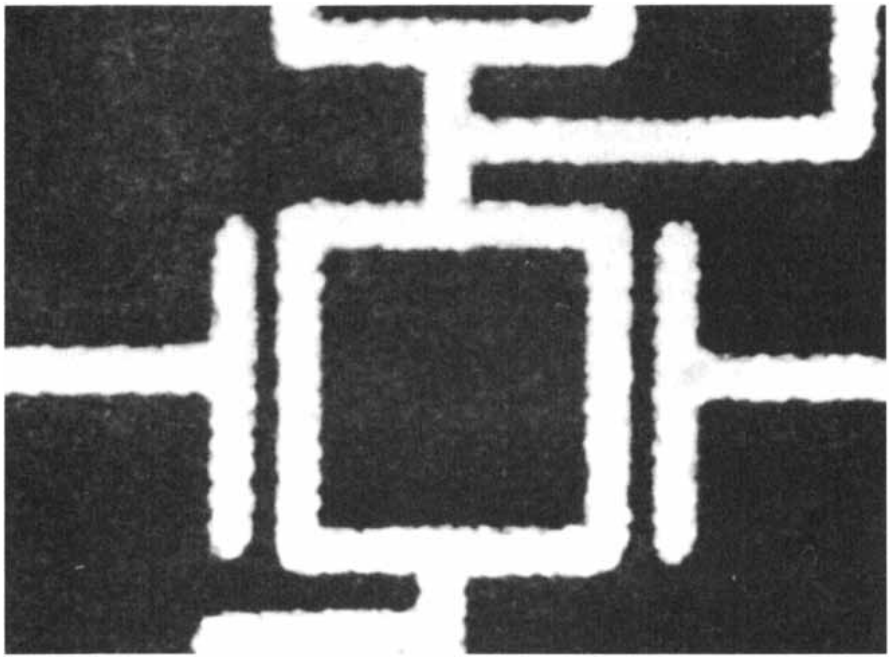
In order to measure electron-interference effects in solid conductors, one must translate the mechanics of electron waves into physical quantities that can be measured easily. When an electron wave travels through a small wire at low temperatures, part of the wave scatters from one end to the other while other parts scatter back to their point of origin. A measure of the difficulty an electron wave has traveling from one end of a wire to the other is electrical resistance; conversely, a measure of the ease with which the wave function moves is the wire's conductance. More than 25 years ago Rolf Landauer of the IBM Thomas J. Watson Research Center in Yorktown Heights, N.Y., developed a theoretical framework expressing the conductance in terms of the probability that an electron wave will be transmitted through the wire. His work shows that the conductance is approximately proportional to the transmission probability divided by a fundamental quantum unit of resistance: 25,812 ohms. This value is equal to Planck's constant divided by the charge of an electron squared.

One factor that contributes to transmission probability and conductance is wave-function interference. Markus Büttiker, Landauer and one of us (Imry) did theoretical work on metallic rings without leads, which demonstrated that elastic scattering did not destroy quantum-interference effects. Then Yuval Gefen, Mark Ya. Azbel and one of us (Imry) predicted in 1984 that, as a result of the Aharonov-Bohm effect, the electrical resistance of a metal ring would oscillate periodically as a magnetic field applied to the center of the ring varied smoothly. When the electron wave functions traveling in two different sections of the ring reinforce each other, the transmission probability and thus the conductance should increase. When the electron wave functions cancel, the transmission probability and the conductance should decrease. Hence the conductance or resistance of a wire should oscillate between these two extremes. In 1981 Boris L. Al'tshuler, Arkady Aronov and Boris Spivak of the Leningrad Institute of Nuclear Physics made a related prediction, and Yuri V. Sharvin and his son at the Institute for Physical Problems in Moscow confirmed it experimentally.

One of us (Webb), working with Sean Washburn, Corwin P. Umbach and Robert B. Laibowitz of the Thomas J. Watson Research Center first demonstrated the Aharonov-Bohm effect in small metallic rings in 1985. The group fabricated a gold ring on a silicon wafer. The ring had an inside diameter of .78 micron and an outside diameter of .86 micron. A current was applied through an input lead attached to one side of the loop and collected at an output lead on the opposite side of



RING measures the Aharonov-Bohm effect in solid conductors (*top*). Electron waves enter from the left and scatter through the ring, which has been cooled to low temperatures. A vector potential field due to a magnetic field (*arrows*) shifts the phase of the electron wave function and changes the ring's electrical resistance, which is determined by measuring the voltage and the current. The Aharonov-Bohm effect accounts for the oscillation in the electrical resistance of the ring (*bottom*).



SWITCHING DEVICE can be based on the electrostatic Aharonov-Bohm effect. An antimony loop .8 micron on a side is flanked by two bars. By applying a potential difference (voltage) to either of the bars or to both, the wave functions of the electrons that travel through the loop change phase, so that the output voltage is altered.

the ring [*see illustration on preceding page*]. Additional wires were attached near the loop to each current lead to measure the voltage drop across the ring. The voltage divided by the current yielded the resistance of the ring. A magnetic field applied perpendicularly created a magnetic vector potential that circulated in the plane of the sample.

The workers observed that the electrical resistance of the ring oscillated periodically as the magnetic field increased. This agreed with what is known about the Aharonov-Bohm effect and potentials. Electron waves that traveled around the gold ring in a clockwise direction interfered with the electron waves that traveled in the opposite direction. As the magnetic field (and the vector potential) was increased, the waves traveling clockwise shifted phase in relation to the waves traveling counterclockwise. As the phase was shifted through a full cycle by the vector potential, the resistance of the ring fluctuated. The average period of oscillation in terms of the magnetic field was .0076 tesla. This quantity multiplied by the average area enclosed by the ring yields a fundamental, quantum-mechanical value equal to Planck's constant divided by the charge of an electron, as predicted theoretically.

The magnitude of the resistance oscillation in this case was quite small: about .1 percent of the total resistance of the ring. Daniel E. Prober of Yale University, Supriyo Datta of Purdue University and their colleagues quickly confirmed the resistance oscillations in other metals and in semiconductors. More recently, experiments by numerous other groups have demonstrated oscillations as large as 50 percent of the total resistance. Furthermore, oscillations in the conductance of these samples are independent of the average resistance and are roughly equal to the charge of an electron squared divided by Planck's constant. Such universality (that is, resistance oscillations independent of the material and its elastic-scattering impurities) was first predicted by Al'tshuler and shortly thereafter by Patrick A. Lee of the Massachusetts Institute of Technology and A. Douglas Stone of Yale University.

The observation of the Aharonov-Bohm effect has opened up an entire new field of research in which the quantum nature of the electron moving in a solid can be studied in the domain that lies between atoms and macroscopic objects. Such "mesoscopic" systems, which are much larger than an atom or a molecule, can be manipulated and measured by macroscopic means, yet they still play by the rules of the game of microscopic physics. These systems directly display the unusual effects of quantum mechanics in, for example, ordinary electrical measurements. It is as though one could measure the resistance of the electrons orbiting an atom. These systems will help to answer fundamental questions such as how large a system must be to behave macroscopically.

Interesting in their own right, the Aharonov-Bohm effect and quantum interference may play a particularly important role in the future of electronics. Since the discovery of the transistor the dimensions of electronic devices have steadily decreased to the point where fewer than 1,000 atoms make up the width of a wire. At the same time the power per unit area that computer chips dissipate in the form of heat has increased. Unless new devices are developed that perform reliably and consume less power, a limit on the number of components per chip will be reached. Ultimately this would limit the operating speed of electronic devices.

Recent research on quantum-interference effects indicates that new electronic devices that dissipate extremely small amounts of power could be developed. An experimental prototype of one such device has already been tested in a low-temperature environment. The device controls resistance and voltage by employing a potential to manipulate the wave characteristics of an electron. In the near future, as the size of electronic components continues to decrease, devices could be constructed that maintain the quantum-mechanical behavior of electrons at much higher temperatures. We find it remarkable that the Aharonov-Bohm effect and other quantum-interference effects, which developed from the abstract foundations of quantum mechanics, have found their way to experiments on down-to-earth samples.

FURTHER READING

SIGNIFICANCE OF ELECTROMAGNETIC POTENTIALS IN QUANTUM THEORY. Y. Aharonov and D. Bohm in *The Physical Review,* Vol. 115, No. 3, pages 485–491; August, 1959.

SINGLE VALUEDNESS OF WAVE FUNCTIONS. E. Merzbacher in *American Journal of Physics,* Vol. 30, No. 4, pages 237–247; April, 1962.

THE PHYSICS OF MESOSCOPIC SYSTEMS. Yoseph Imry in *Directions in Condensed Matter Physics,* edited by G. Grinstein and E. Mazenko. World Scientific Publishing Co., 1986.

QUANTUM INTERFERENCE FLUCTUATIONS IN DISORDERED METALS. Richard A. Webb and Sean Washburn in *Physics Today,* Vol. 41, No. 12, pages 46–53; December, 1988.

Create a new world where lightweight plastics can outfly metals.

Aerospace designers are limited by their materials, not their dreams.

At BASF, we looked at the design limitations of metals and saw the need for a radically new generation of materials. The result: strong, lightweight, carbon fiber reinforced plastics. These Advanced Composite Materials will enable future designs to carry more, faster, farther.

In one industry after another, from aerospace to automotive, our broad-based technologies help us create new worlds by seeing in new ways.

The Spirit of Innovation

BASF

THE BMW 5-SERIES.
THE PERFECT RECONCILIATION
BETWEEN LOVE OF FAMILY
AND PASSION FOR DRIVING.

The typical BMW engineer is a family man with a genuine passion for building and driving extraordinary automobiles.

So it's not surprising that BMW's 5-Series is the family-size four-door that "effectively recalibrates the performance sedan benchmark" (AutoWeek Magazine). This athletically-shaped BMW coddles the senses with rich leather seating, and ride comfort deemed the new standard in its class. And it reassures the mind with superior anti-lock disc brakes and a unitized steel body that is

designed to absorb 35 percent more impact energy than Federal law requires.

But what this car does best is exalt the spirit. The sweet song of its in-line 6-cylinder engine, the feel of the road transmitted through its precise power steering, the pull of gravity as its 50/50 balance and patented independent suspension mate your tires to the curving pavement—all this must be experienced.

The experience is yours by contacting your BMW dealer and test driving the 5-Series sedan.

The car you need, reconciled with the car that you really want.

THE ULTIMATE DRIVING MACHINE.®

Astrocytes

Named for their starlike shape, these brain cells were once thought to be little more than passive support elements. Recent findings, however, implicate them in brain function, development and disease

by Harold K. Kimelberg and Michael D. Norenberg

Attempts to explain how the human brain works are usually formulated in terms of the activities of neurons; by conveying complex electrical signals, these excitable cells give rise to the brain's capacity to process information. It may come as a surprise, then, for the reader to learn that up to half of the brain's volume does not consist of excitable cells at all but of nonexcitable ones. The largest class of nonexcitable cells is that of the neuroglia, or "nerve glue," a name coined from the Greek. And among the most significant of the glial cells are those known as astroglia, or as they are more commonly termed, astrocytes.

Named for their starry shape, astrocytes have been known since the late 19th century. For most of that time it has been assumed that they serve as little more than passive physical support elements for the neurons. In little more than a decade, however, that point of view has given way as advances in methods for identifying and growing astrocytes have made it possible to understand their function better. It is now clear that astrocytes have key roles in normal physiology, in brain development and in the pathology of the nervous system. It has become evident that the brain will never be fully understood without an understanding of the many roles of astrocytes.

HAROLD K. KIMELBERG and MICHAEL D. NORENBERG are both investigators whose work centers on the astrocyte. Kimelberg is research professor of surgery in the division of neurosurgery at Albany Medical College. At that institution he is also professor of biochemistry and pharmacology/toxicology and director of the neuroscience program. When he moved to Albany in 1974, he began devoting his research exclusively to astrocytes. Norenberg is professor of pathology and director of neuropathology at the University of Miami School of Medicine. His primary scientific interest lies in degenerative brain diseases due to metabolic conditions and the role of astrocytes in such disorders.

Identifying the Astrocyte

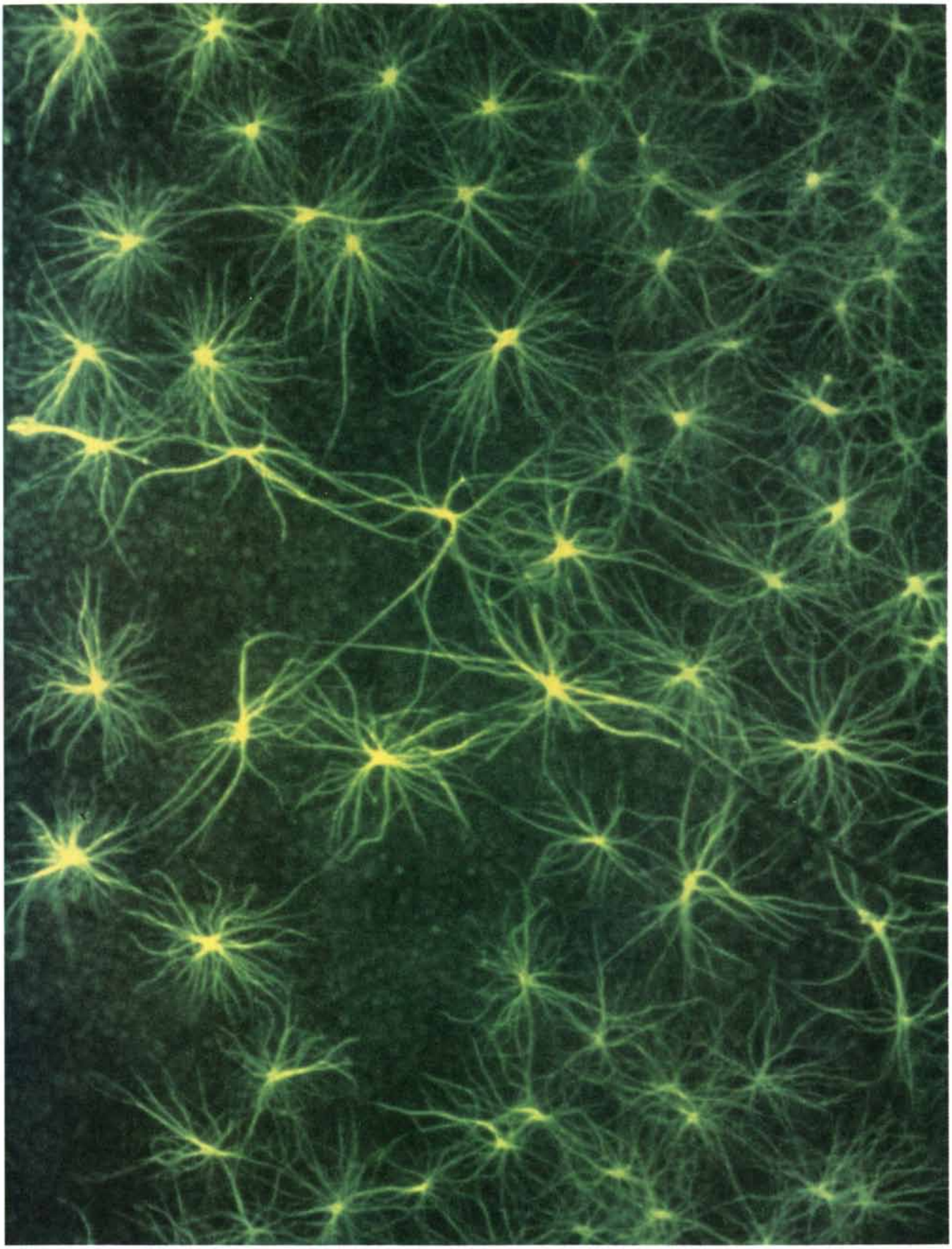
The term neuroglia was first used by the great German pathologist Rudolf Virchow in 1846. Virchow coined the term to describe regions between neurons that he considered analogous to the cementing connective tissue seen in other bodily organs. The finding that such regions consist of a distinct class of nonneuronal cells was made possible by the development of staining techniques based on metallic impregnation introduced by the famed Italian cytologist Camillo Golgi in the mid-1870's. The pioneering Spanish neuroanatomists Santiago Ramón y Cajal and his student Pio del Río Hortega carried those staining techniques further and developed a system for classifying glial cells that has persisted largely intact until today.

In this system the overwhelming majority of the cells in the neuroglia are classified as macroglia. (A small minority, thought to be derived from nonnervous tissue, are called microglia.) The macroglia are in turn subdivided into astrocytes and oligodendrocytes. The oligodendrocytes are known to form the myelin that wraps the axons of neurons in the white matter of the brain (and indeed gives the white matter its characteristic coloring, which distinguishes it from the gray matter). Astrocytes come in two main forms: fibrous astrocytes are found in the white matter and protoplasmic astrocytes in the gray matter. (Other subtypes also exist.) Although the types of astrocytes differ anatomically, the functional differences between them are not understood, which is not surprising, because the overall functions of astroglial cells are only now coming into view.

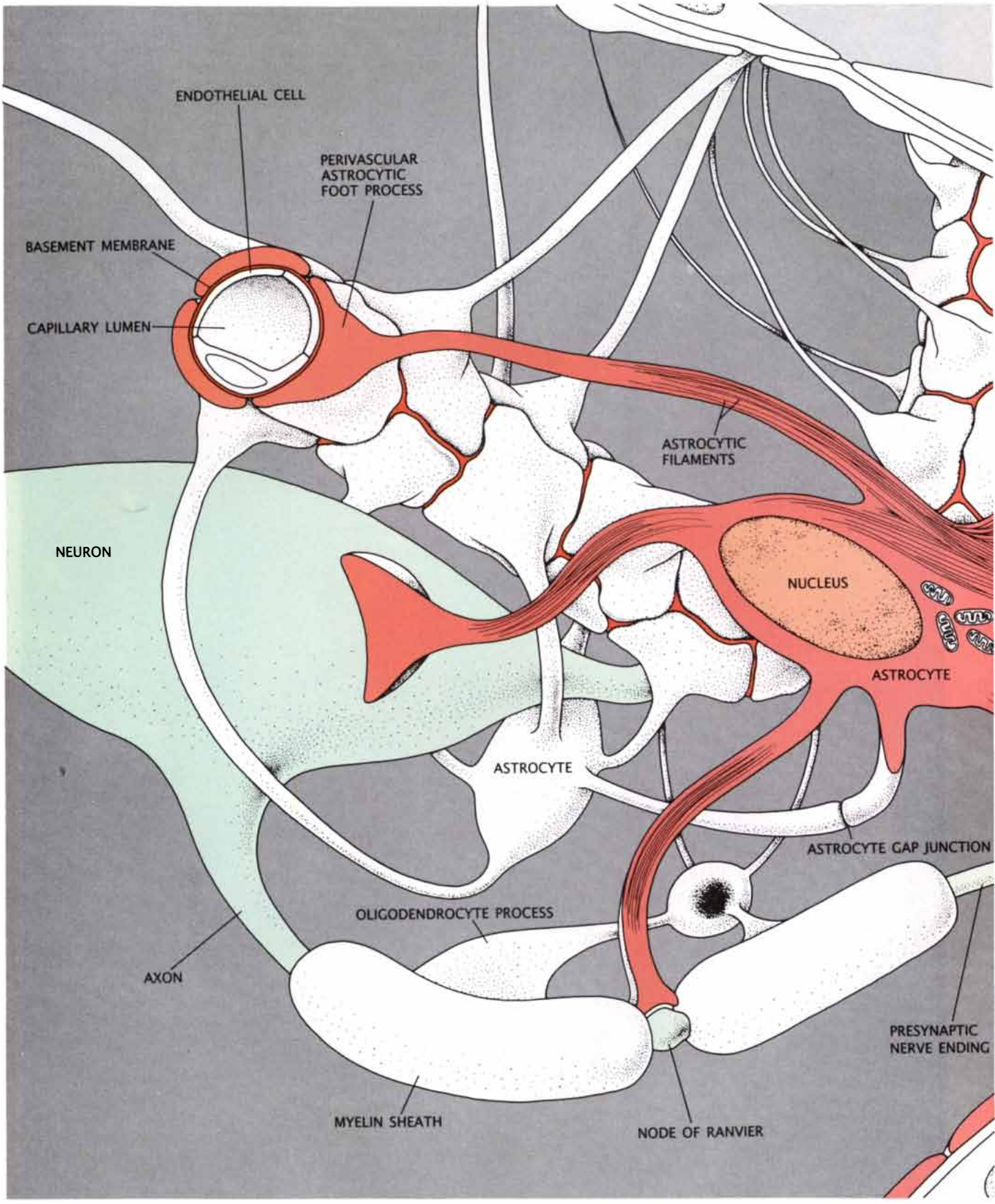
If the functions of astrocytes have begun to come clear in the past 15 years, one reason is a significant advance that took place at the beginning of the 1970's: the discovery of GFAP. GFAP (glial fibrillary acidic protein) is one component of the intermediate glial filaments found in the cytoplasm of the astrocyte; those filaments in turn form part of the cytoskeleton that endows the astrocyte with overall shape (as it does most cells). First isolated by Lawrence F. Eng and Amico Bignami of Stanford University, GFAP is found only in astrocytes and has therefore proved to be an invaluable marker for identifying those cells in tissue samples and in culture. Indeed, many of the data on astrocyte function presented in this article come from work with pure cultures of astrocytes that have been certified as such by the presence of GFAP.

Such work suggests that astrocytes have a variety of active roles in maintaining normal brain physiology. For example, astrocytes are now known to occupy a critical position in the metabolism of glutamate and gamma-aminobutyric acid (GABA), which are respectively important excitatory and inhibitory neurotransmitters. In order for the neuronal network to function smoothly, these transmitters, after being released into the synaptic cleft (the gap between neurons), must be removed. It is now clear that some of the removed transmitter molecules are carried into astrocytes. In the astrocyte both GABA and glutamate can be metabolized to form the amino acid glutamine. Subsequently glutamine can serve as raw material for making more of the transmitters.

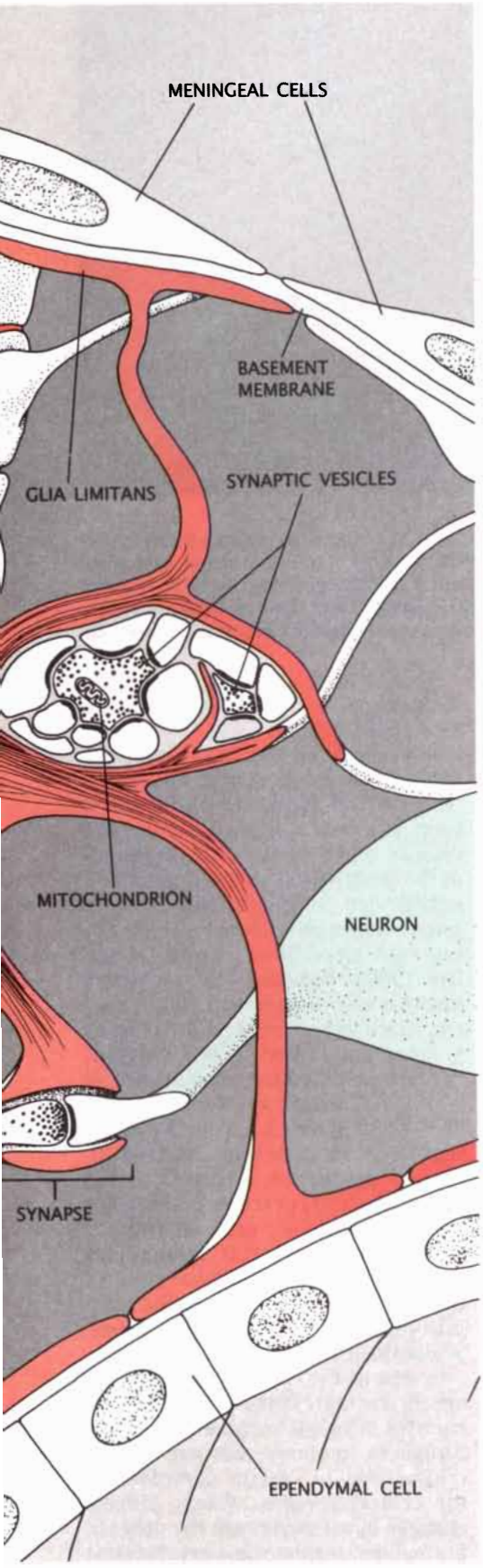
Beginning in the 1950's on the basis of concepts developed by Heinrich Waelsch and Soll Berl of Columbia University's College of Physicians and Surgeons, it was shown that one fraction of glutamate in the brain was converted into glutamine, whereas another (larger) fraction was consumed in unrelated metabolic activities. Based on indirect observations, it



ASTROCYTES are shown in tissue from the retina of a cat. The astrocytes are greenish because they have been made to fluoresce. The fluorescent dye was linked to an antibody to GFAP, a protein found only in astrocytes. The radiating fibers enable astrocytes to interact with neurons, brain capillaries and other cells. The micrograph was made by Andreas Karschin, Heinz Wässle and Jutta Schnitzer of the Max Planck Institute for Brain Research. The magnification is about 300 diameters.



COMPLEX RELATIONS between astrocytes and other cells are shown in a highly schematized view of a portion of brain tissue. Astrocytic foot processes surround capillaries in the brain; they seem to have a role in inducing the "blood-brain barrier" that prevents water-soluble substances from freely diffusing into the brain from the capillaries. Astrocytes also make various types of contacts with neurons. They are present near certain synapses (the gaps where neurotransmitters are released); they take up and metabolize certain neurotransmitters. Astrocytes also reach to the cells that line the ventri-



cles (the fluid-filled cavities at the center of the brain) and the meninges (the membranous coverings of the brain). Oligodendrocytes extend fibers that wrap certain axons with sheaths of myelin.

was suggested that glutamine might be synthesized in the glial cells. Direct confirmation of that concept was provided by one of us (Norenberg) and his co-workers, who showed that glutamine synthetase, the glutamine-making enzyme, is found exclusively in astrocytes. GABA and glutamate are efficiently taken up into the astrocytes and then metabolized to glutamine. Once formed, the glutamine is conveyed to the neurons, where it is probably the chief precursor for GABA and glutamate in the neurons that utilize these transmitters.

Balancing the Ions

If neurotransmitters are crucial to the function of the brain, the ionic composition of the region around the neurons is equally critical, and here too astrocytes have an important role. The reason ions are so significant is that the balance and flow of these electrically charged molecules across the neuronal membrane is responsible for generating the action potential that conveys the nerve impulse. In particular, the flows of (positively charged) sodium and potassium ions are responsible for the generation of the action potential, and the levels of these ions—particularly potassium—must be very finely calibrated around the neurons.

In much of the rest of the body external ions reside in a fluid-filled space outside the cell. The fact that early electron micrographs revealed astrocytic processes (fibers) filling all the available space around the neurons led some workers to speculate that astrocytes served the same function in the brain that the extracellular space did in the rest of the body. But the hypothesis did not hold up. In the mid-1960's Stephen W. Kuffler, John G. Nicholls and Richard K. Orkand of the Harvard Medical School examined the glial cells in the optic nerve of amphibians and found that these cells do not have the high sodium content found in the extracellular space. Instead, like neurons, they have a high internal concentration of potassium.

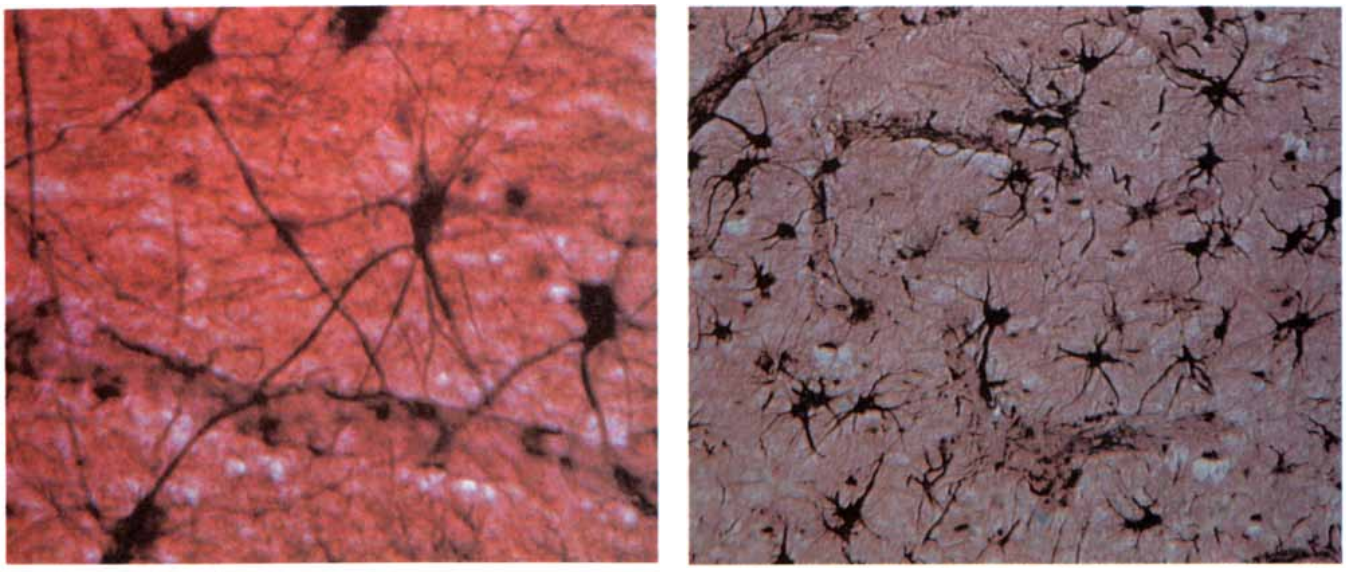
These and other observations led to the concept of "potassium spatial buffering" by astrocytes. In this concept excess potassium in the extracellular space produced by neuronal activity is carried away from the neurons by the action of astrocytes. The driving force for the flow is a localized increase in extracellular potassium due to the neuronal activity. As a result the astrocytic membrane in the region of the local increase becomes more electrically positive. The difference in electric potential between that region and distant ones yields a current flow within the glial cell that carries the charged potassium ion away from its site of release.

Findings supporting the notion of potassium spatial buffering have recently been published by Eric A. Newman of the Eye Research Institute of the Retina Foundation of Boston for a specialized type of astrocyte called the Müller cell, which is found in the retina. In addition, freeze-fracture studies have shown a high density of structures in localized regions of the surface membrane of astrocytes that may be potassium channels. The actual function of these intriguing geometric structures, however, has not yet been established.

In maintaining the correct balance of potassium, the astrocytes are serving to maintain the proper environment for neurons. As we noted in the introduction to this article, the operation of the brain has generally been conceived of in terms of the operation of the neurons, and their properties have generally been thought of as occupying a category distinct from that of the lowly astrocytes. Much recent work shows, however, that there is actually a fair amount of overlap in the properties of the two cell types.

The complexity of the neuronal responses in the brain is thought to be due in large part to the action of neurotransmitters on receptors, which are complex proteins embedded in the surface membranes of cells. When transmitters bind to specific receptors in the neuronal membrane, they cause a wide variety of different effects because the receptors are effectively linked to ion channels in the membrane and to the release of molecules called second messengers within the cell. Of great interest in relation to astrocytes is the fact that many receptors and second messengers first found in neurons have now been identified in astroglial cells.

For example, the most prevalent receptor so far found on astrocytes is the beta-adrenergic type that binds the neurotransmitter norepinephrine, triggering a large increase in the cell of a second messenger called cyclic adenosine monophosphate (cAMP). For glial cells this phenomenon was first described (in glial cultures) by Alfred G. Gilman of the University of Virginia and Bruce Schrier of the National Institutes of Health in 1972; the presence of the beta-adrenergic receptors has now been localized to individual astrocytes by Ken D. McCarthy and his



METALLIC STAINS provided the first means of identifying individual astrocytes in the light microscope. Both panels show astrocytes marked by a gold-sublimate stain. The one at the left depicts normal astrocytes, their fibers terminating on a brain blood vessel that extends from left to right. The one at the right shows postmortem brain tissue from a neurosyphilis patient. There are more astrocytes than usual, and their form is aberrant. The micrographs, made by Lowell W. Lapham of the University of Rochester Medical Center, have a magnification of about 300 diameters (*left*) and 200 diameters (*right*).

colleagues at the University of North Carolina at Chapel Hill. Other studies show that astrocytes have receptors for most neurotransmitters. What they are doing there is not yet known, but the presence of receptors on astrocytes implies that those cells, far from being passive architectural elements, respond to changing brain conditions in as complex a way as neurons do.

Developmental Scaffolds

The contributions of astrocytes to normal brain physiology have been patiently pieced together for more than a decade. Now another area, which has developed over the same period, is generating even more excitement among those who work on astrocytes: brain development. Beginning about 20 years ago it was shown that certain astrocytes, known as radial glia, have a key role in the embryonic development of the brain, as the celebrated Swiss embryologist Wilhelm His first suggested a century ago. These astroglial cells act as a scaffold, enabling neurons to migrate from their points of origin in the developing nervous system to their final destination in the maturing brain.

This effect was first shown for a type of neuron called the granule cell in the monkey cerebellum by Pasko Rakic of the Harvard Medical School. Granule cells migrate inward from the outer layers of the developing cerebellum, supported by long fibers emanating from certain radial glial cells called Bergmann glia. In an important series of experiments Jerry Silver of the Case Western Reserve University School of Medicine has shown that in the mouse brain a "sling" made of astrocytes supports the growth of axons from one side of the brain to the other, forming the corpus callosum (the major tract of nerve fibers connecting the two hemispheres of the brain). In animals lacking the glial sling, the corpus callosum does not form. If a piece of plastic filter material on which astrocytes are growing is implanted, however, the axons will regrow. These findings may ultimately have some significance for the regrowth of cut or broken axons in the brain or spinal cord.

Another proposed developmental function for the astrocyte is in the induction of the blood-brain barrier. In most of the body the endothelial cells that form the lining of the smallest blood vessels, the capillaries, are only loosely joined, allowing smaller water-soluble substances that are in solution in the blood to pass into the tissues. In the brain, however, the capillary endothelial cells are joined by so-called tight junctions and as a result are virtually impermeable to substances soluble in water. The barrier formed in this way is of great significance in maintaining a constant chemical milieu within the brain, but it also is a hindrance in therapy for brain diseases [see "The Blood-Brain Barrier," by Gary W. Goldstein and A. Lorris Betz; SCIENTIFIC AMERICAN, September, 1986].

Now, the capillaries in the brain are almost completely surrounded by the "end feet" of astrocytic fibers, which form a jigsaw pattern covering the vessel's outer surface. The geometry of the capillary-astrocyte combination initially led to speculation that the astrocytes might themselves constitute the blood-brain barrier. In the late 1960's, however, elegant experiments employing electron-dense tracers, which were carried out by Thomas S. Reese and Milton W. Brightman of the National Institute of Neurological and Communicative Disorders and Stroke and Morris J. Karnovsky of the Harvard Medical School, demonstrated that the barrier is formed by the endothelial cells of the vessel wall. But astrocytes have not been entirely deprived of their role: it is thought that they induce the endothelial cells of the brain capillary to form the tight junctions and enzymes that are characteristic of the blood-brain barrier.

In view of the critical functions that are increasingly being ascribed to astrocytes, it is not surprising that disturbances in those cells can have a crucial role in certain disorders of the central nervous system. Indeed, changes in astrocytes are the only abnormalities seen in some neurological disorders; in others such changes precede alterations in other elements of the nervous system, principally the neurons. Injury to the brain often results in reactive gliosis: an increase

in the size and number of the astrocytes. Reactive astrocytes have more processes and glial filaments than ordinary ones and also have increased metabolic activity. The result of this reaction is often the formation of a glial scar made up of these reactive astrocytes.

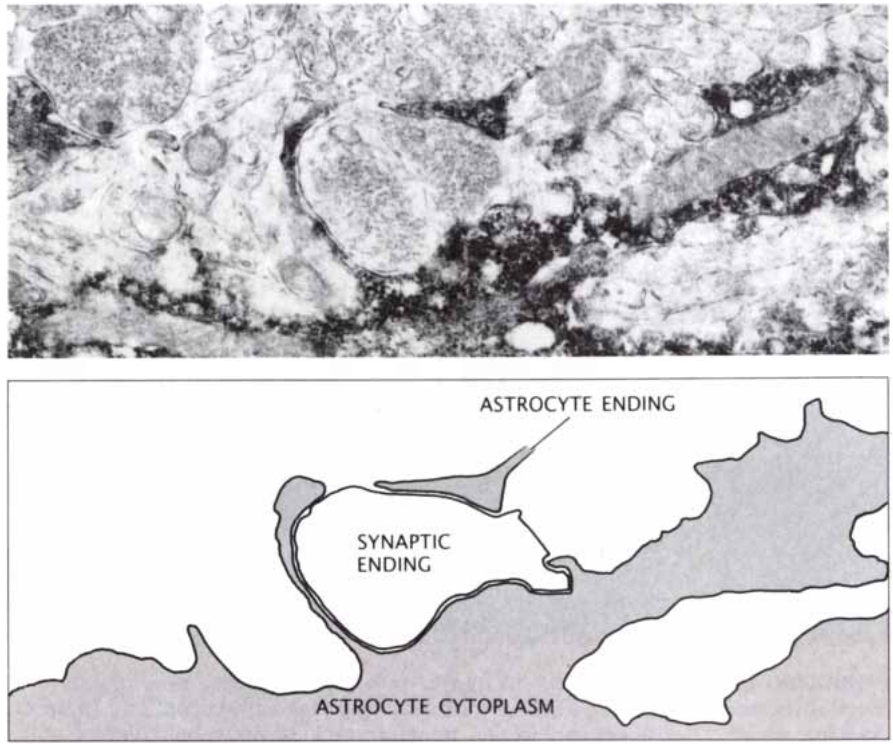
The Astrocytic Scar

The process of reactive gliosis has generally been viewed as the laying down of inert scar tissue. For many years most students of brain injury thought the astrocytic scar impeded neurons from reextending their axons and perhaps also prevented the oligodendrocytes from remyelinating axons that had lost their wrappings (as happens in such diseases as multiple sclerosis). It now seems possible that—at least in the early stages—reactive astrocytes actually have a restorative role.

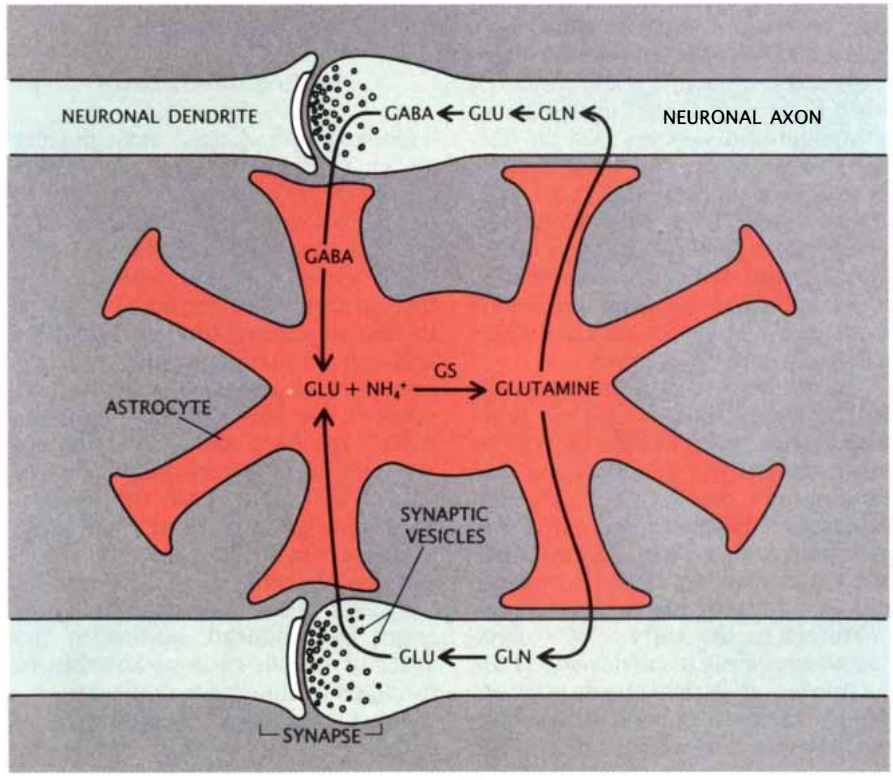
Astrocytes can release a number of growth factors for neurons. Some of these, such as nerve growth factor, may stimulate the neuron as a whole as well as promoting axonal growth. Others, including laminin, fibronectin and other extracellular matrix constituents, may specifically promote and accelerate the growth of neuronal processes. Furthermore, as we described above, during embryonic development astrocytes provide a scaffolding along which neurons migrate to find their proper places. It is quite possible that in reaction to injury the astrocytes may recapitulate this role, thereby making it possible for disrupted nervous-system connections to be restored. This function would clearly be aided by the release of the various factors that promote nerve-cell growth.

But if astrocytes do have a healing function in the injured brain, they can also have a variety of deleterious effects if their own operation is disrupted. Because the glial scar is characteristic of epilepsy, there has long been conjecture that astrocytes play a part in the disorder. Daniel A. Pollen and Michael C. Trachtenberg of the Massachusetts General Hospital suggested that a defect in potassium buffering might lower the threshold for excitation of the nerve cell and thereby yield the neuronal hyperactivity seen in epilepsy. Although their hypothesis is still controversial, it is certainly plausible.

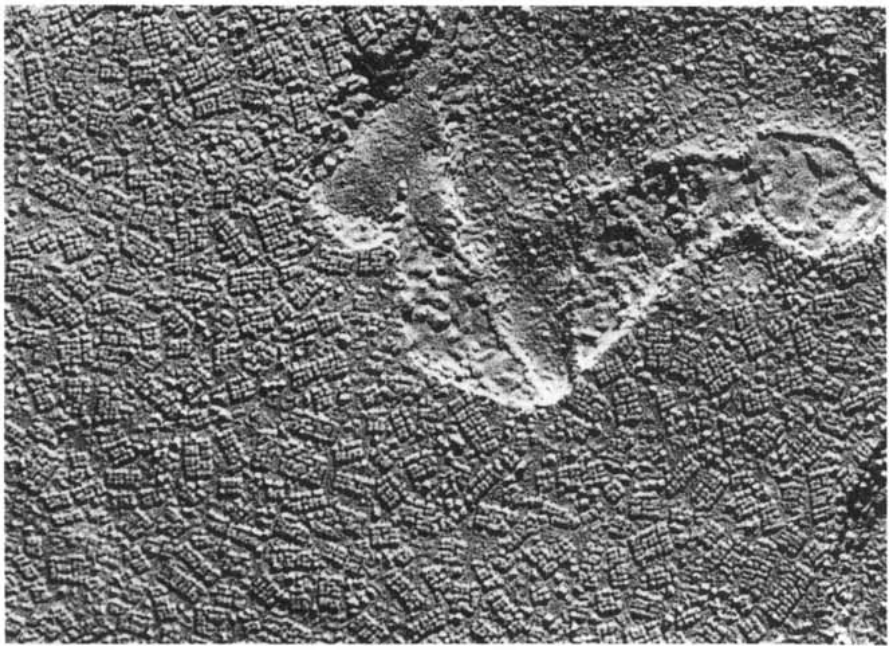
Some further confirmation of the involvement of astrocytes in epilepsy stems from the fact that both glutamate and GABA are commonly impli-



SYNAPTIC NERVE ENDING is surrounded by astrocytic fibers containing the enzyme glutamine synthetase, whose presence is indicated by a black reaction product. The enzyme converts glutamate, a neurotransmitter, into glutamine. Demonstrations such as this one, carried out by one of the authors (Norenberg), showed that the conversion is carried out in the astrocyte. The magnification is 21,000 diameters.



NEUROTRANSMITTER METABOLISM takes place partly in astrocytes, as is indicated in the diagram. Both glutamate (*Glu*), an excitatory transmitter, and an inhibitory transmitter called GABA are converted into glutamine (*Gln*) in the astrocyte by glutamine synthetase (*GS*). Glutamine returns to the neurons, where it is used to make new neurotransmitter molecules. These reactions consume ammonia (NH_4^+) and therefore serve to keep the level of that toxic chemical low in brain tissue.



INTRIGUING GEOMETRIC ASSEMBLIES appear in a freeze-fracture micrograph of a piece of the outer membrane of an astrocyte surrounding a brain capillary. Indirect evidence suggests these assemblies may be potassium channels that enable the astrocyte to regulate that ion's level in the brain. The micrograph has a magnification of about 120,000 diameters; it was made by Dennis M. D. Landis and Thomas S. Reese of the National Institute of Neurological and Communicative Disorders and Stroke.

cated in the development of epilepsy. Povel Krogsgaard-Larsen of the Royal Danish School of Pharmacy in Copenhagen and his colleagues have recently reported some intriguing results of work done with mice susceptible to sound-induced seizures that are similar to the seizures of epilepsy. The Danish investigators find that a substance called THPO, which selectively blocks the uptake of GABA by astrocytes but not by neurons, can protect the mice from seizures; analogous molecules that block GABA uptake by neurons are not protective.

The evidence that astrocytes take part in the genesis of epilepsy is intriguing but indirect. In the case of two other conditions—Parkinsonism and Huntington's disease—it seems that substances produced by astrocytes may play a direct role. Parkinsonism is a motor disorder whose symptoms can include trembling and rigidity. In California in the early 1980's some intravenous-drug users developed the symptoms of Parkinsonism after injecting themselves with drugs contaminated by a substance called MPTP. It was subsquently shown that the toxic agent was actually a metabolic product of MPTP called MPP^+; it was the MPP^+ that was killing a certain group of neurons and causing Parkinsonism. The modification of MPTP to MPP^+ is carried out by an enzyme—one of the class known as monoamine oxidases—that has now been shown to be present in astrocytes.

Huntington's Disease

Huntington's disease may present an analogous picture, as Robert Schwarcz of the Maryland Psychiatric Research Center near Baltimore and his colleagues have suggested. A substance called quinolinic acid is a normal by-product of the metabolism of the amino acid tryptophan. Quinolinic acid can act inappropriately on a certain class of glutamate receptors on neurons, thereby killing the neurons. It has now been shown that the enzyme that synthesizes quinolinic acid (known as 3-HAO) is present predominantly, or perhaps even exclusively, in astrocytes. If there were an abnormality of that enzyme in the astrocyte, a surplus of the toxic quinolinic acid might be produced, leading to the death of specific neurons and thus to the signs of Huntington's disease.

The participation of astrocytes in pathology is by no means limited to the generation of disease. Indeed, it now seems that astrocytes are important for the immune response in the brain. The absence of a lymphatic system in the brain, along with the presence of the blood-brain barrier (which excludes many immunologic cells and agents), led many investigators to believe the brain is an organ insulated from the effects of the immune system. The work of Adriano Fontana of the University Hospital in Zurich and his colleagues has radically altered this view by showing that astrocytes function (at least in culture) as accessory cells that mediate immune reactions in the brain.

In this aspect of their function astrocytes serve as antigen-presenting cells (APC's). Antigen-presenting cells are well known outside the brain and include such cells as monocytes, macrophages and the Langerhans cells of the skin. The combination of an antigen (a foreign protein) and particular endogenous molecules as presented by the APC triggers other cells, generally *T* lymphocytes, to respond and destroy the invading organism. The endogenous molecules presented with the antigen are encoded by a part of the cell's genetic repertory referred to as the Major Histocompatibility Complex (MHC).

MHC molecules come in two forms, designated class I and class II. It has now been shown that both types can be found on astrocytes, giving them what appears to be the full range of function of the APC's. Class I MHC proteins are required for cells called killer *T* cells to find and destroy their targets; it has been known for some time that these molecules can be induced on astroglial cells. More recently Paul Massa and Volker ter Meulen of the University of Würzburg and Fontana have shown that astrocytes from certain species are capable of generating class II MHC molecules and that this capacity may be related to multiple sclerosis and other disorders involving the immune system.

Multiple sclerosis has long been suspected to be the result of an attack by the patient's own immune system on certain myelinated axons. A laboratory model of the disorder is provided by experimental allergic encephalomyelitis (EAE), which can be induced by injecting homogenized brain tissue or purified myelin basic protein from one animal into another. Presumably the injected tissue triggers an immune response that ultimately leads to the destruction of healthy tissues. Intriguingly, the reaction is specific to species and even to strains within a single species: Lewis rats are susceptible, whereas Brown Norway rats are not.

Differences in the capacity for developing EAE may in turn reflect variation among astrocytes in the strains. Massa and his co-workers have shown that astrocytes from Lewis rats can

Tandy® Computers: The broadest line of PCs in America.

The Tandy 1000 TL

286 power with MS-DOS® and DeskMate® built in.

The Tandy 1000 TL is a powerful computer for personal and business use. Its 80286 microprocessor gives you extraordinary speed and processing power. Plus, the 1000 TL comes with MS-DOS and the DeskMate Graphical User Interface built in, so you can be up and running in seconds, using plain-English commands.

DeskMate features ten applications that let you write reports and letters, prepare budgets, file, draw colorful pictures, create and play back songs and more. Plus, there's PC-Link®, an online information service.

You also get the latest in computer-audio technology. When you use DeskMate's sound editor, you can record and edit voice, music or any analog source onto diskettes.

The 640K Tandy 1000 TL comes with a 3½" disk drive and has room for an additional 3½" and 5¼" drive. A parallel printer adapter, RS-232 serial port, two joystick ports, a clock/calendar and five expansion slots are all standard. You also get a 101-key enhanced keyboard for the ultimate combination of power, ease of use and affordability.

The new generation Tandy 1000 TL. From the best-selling family of PC Compatibles made in America.

Send me a 1989 RSC-20 Computer Catalog

Mail to: Radio Shack, Dept. 89-A-922
300 One Tandy Center, Fort Worth, TX 76102

Name _____

Address _____

City _____

State _____ ZIP _____

Phone _____

Tandy Computers: Because there is no better value™.

MS-DOS/Reg. TM Microsoft Corp. PC-Link/Reg. TM Quantum Computer Services.

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

generate the MHC class II proteins; those from Brown Norway rats cannot. Thus the ability of astrocytes to make class II MHC molecules may be crucial to at least one abnormal immune response. Support for this concept comes from studies in which antibodies to the class II molecules greatly lessened the structural damage caused by EAE. If the capacity of the astrocytes to generate MHC proteins is essential to an autoimmune response such as EAE, it might also be expected to be implicated in responses of the brain to foreign antigens.

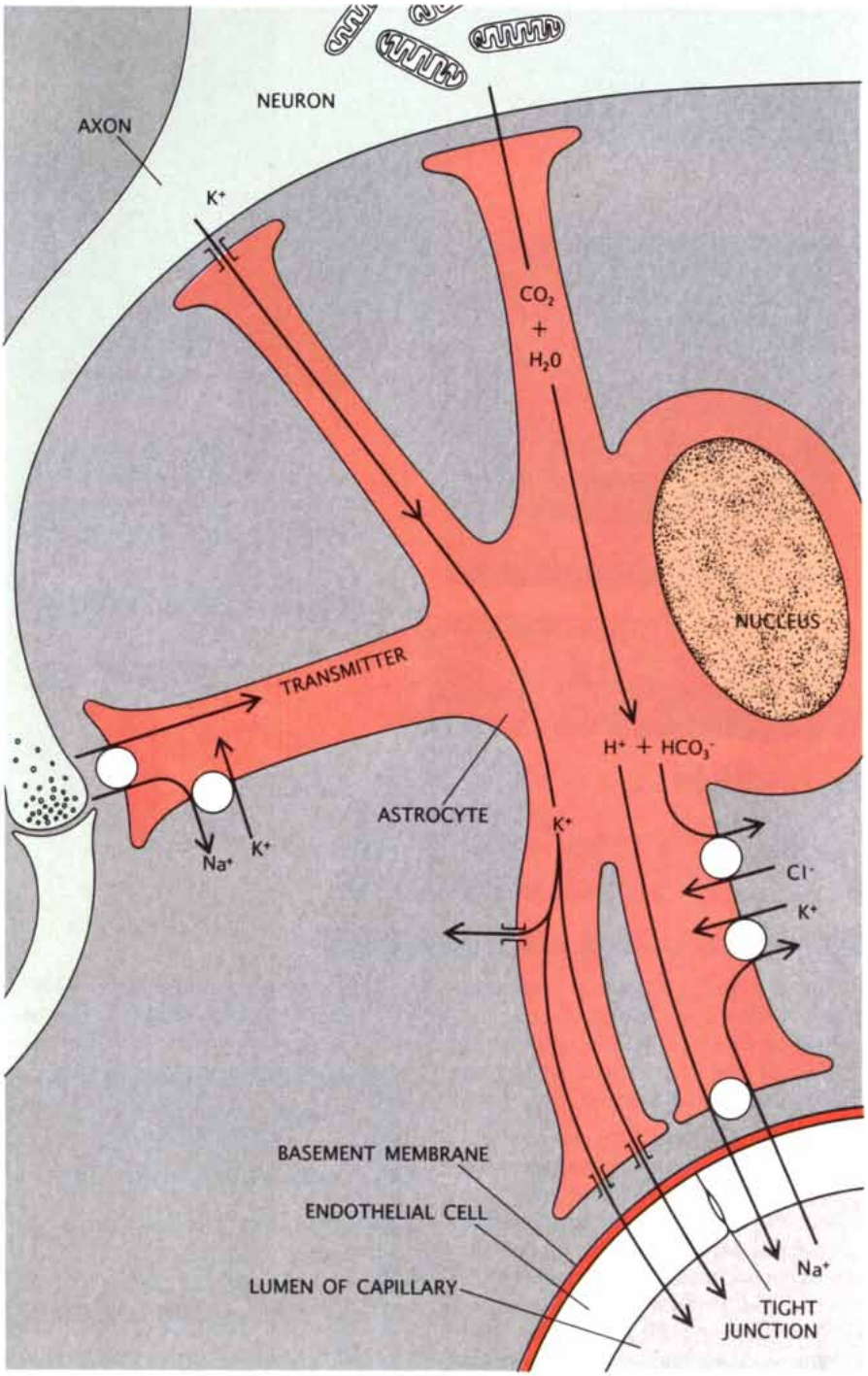
Swelling of the Astrocytes

These findings may ultimately play an important role in the understanding of multiple sclerosis. Meanwhile attempts to understand why swollen, watery astrocytes are usually seen as an early response to brain injury may have a more immediate clinical payoff. For some years a group at the Albany Medical College that includes Kevin D. Barron, Robert S. Bourke and Louis R. Nelson along with one of us (Kimelberg) has been trying to clarify the role of astrocytic swelling in one type of brain injury. Our group found that in laboratory animals astrocytic swelling and the mortality associated with brain trauma can be reduced by application of a nondiuretic substance derived from a group of compounds known as loop diuretics, so named because they act at the loop of Henle in the kidney; the nondiuretic derivative was synthesized by Edward J. Cragoe, Jr., of the Merck Sharp & Dohme Research Laboratories.

How might the derivative work? The answer appears to involve systems for transporting ions across the astrocyte membrane. We noted above that there are channels for potassium in the astrocyte membrane, but those channels clearly do not exhaust the complexity of the transport systems there. Other systems exist, some of which transport several ions in a linked fashion. There is, for example, a system for the coupled exchange-transport of sodium ions and protons (hydrogen ions) as well as one for the exchange of chloride and bicarbonate ions.

Now, the suggestion had been made by the Albany group that astroglial swelling might be due to increased activity of these two exchange-transport systems, leading to accumulation of sodium and chloride ions in the astrocytes. In the process of osmosis, water flows across the cell membrane to eliminate any net concentration gradient of charged or uncharged substances that cannot freely cross the membrane. If sodium and chloride were to build up within the astrocyte, "osmotically obligated" water would also accumulate there, leading to astrocytic swelling. Some support for this mechanism has been provided by Ricardo Garay of the Hôpital Necker in Paris, who found that one action of the diuretic derivative was indeed to inhibit the ion-transport system that controls the exchange of chloride and bicarbonate ions.

The work done on astroglial swelling shows that astrocytes take a fundamental part in pathologies that stem directly from the brain. They may have an equally important role even when the primary disorder lies outside the central nervous system. Patients with

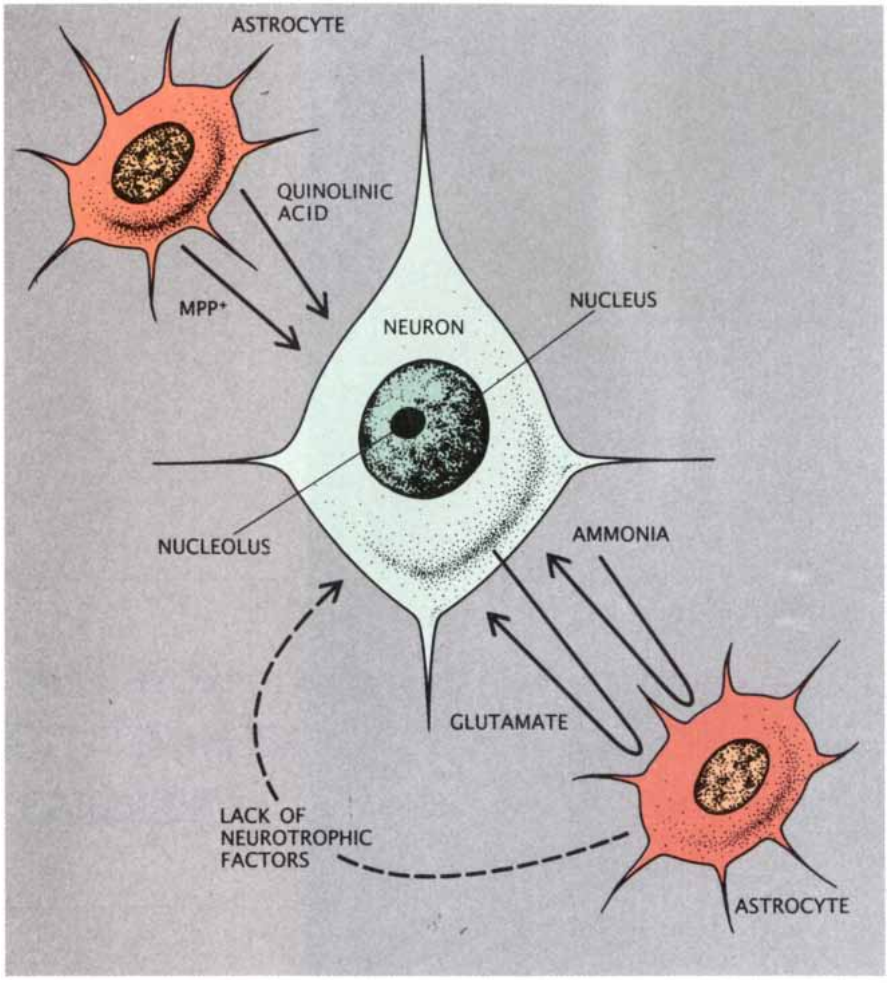


ION-TRANSPORT PROCESSES thought to be present in astrocytes are shown diagrammatically. A variety of processes are present, enabling the astrocyte to exchange substances such as sodium (Na^+), potassium (K^+), bicarbonate (HCO_3^-), chloride (Cl^-) and protons (H^+) with neighboring cells, including neurons; carbon dioxide, in contrast to ions and charged molecules, diffuses freely into and out of cells.

BEAT THE COMPETITION OVERSEAS.

There are those who still see flying Concorde as an indulgence. But at twice the speed of sound, it may be more appropriate to view it as simply a terrific head start. One that saves you 3 hours from New York, 3 hours and 25 minutes from Washington, and 2 hours and 15 minutes from Miami. Think of all those hours as billable, and you'll realize Concorde does more than let you beat your competition on the way to Britain. It also lets you beat them on the bottom line.

BRITISH AIRWAYS
The world's favourite airline.®



ASTROCYTES CAN DAMAGE NEURONS by both indirect and direct means. An artificial metabolic product of astrocytes (MPP^+) and a natural one (quinolinic acid) have been implicated in causing Parkinsonism and Huntington's disease respectively. Conversely, if astrocyte metabolism is disrupted, potentially toxic substances such as ammonia and glutamate may build up or the astrocyte may fail to release neurotrophic factors required for the proper maintenance and growth of nerve cells.

severe cirrhosis of the liver often develop neurological symptoms ranging from subtle personality changes to stupor and coma; collectively these alterations are known as hepatic encephalopathy, an example of which is Reye's syndrome. The precise mechanisms by which liver disease adversely affects cerebral function are not yet known, but it is thought the process is mediated by toxins that circulate in the blood because the liver—the body's main detoxifying organ—is impaired. The toxins can then cross the blood-brain barrier and interfere with brain metabolism.

Among the toxins that have been proposed for this role are ammonia as well as substances called short-chain fatty acids and mercaptans. All three seem to act on astrocytes. In many patients who die of hepatic encephalopathy the only observable brain pathology is the appearance of aberrant astrocytes with large nuclei and few fibers. These alterations are known as Alzheimer's Type II changes after the German neurologist Alois Alzheimer (for whom Alzheimer's disease is also named).

A Vicious Circle

The notion that the astrocytes are principally affected is supported by the fact that the toxins mentioned above can act on these cells. What is more, the presence of ammonia in particular may set in motion a vicious circle. The cycle described above in which glutamine is synthesized from glutamate in the astroglial cells consumes ammonia; indeed, that set of synthetic reactions is the brain's primary protection against the accumulation of the toxic chemical. If the astrocytes are injured, ammonia will begin to build up in the brain tissues, making the injury worse. Hepatic encephalopathy, then, appears to be a disorder in which the astrocytes are initially affected and may be primarily responsible for the derangement of neurological function.

Although most of the experimental work we have described focuses on physiological function, it is even possible that astrocytes are involved in psychiatric conditions; that notion has received attention through the studies of Fritz A. Henn of the State University of New York at Stony Brook and Leif Hertz of the University of Saskatchewan and their colleagues. Many of the psychoactive drugs currently being prescribed (including antianxiety medications and antidepressants) either interact with receptors that are now known to be present on astrocytes or influence the metabolic processes that go on within those cells. Although the participation of astrocytes in psychiatric conditions can for the moment only be thought of as highly speculative, it should be a fruitful area for further research.

The discussion of psychiatric disorders—which clearly involve emotions and ideas—brings us full circle. We began by pointing out that most attempts to understand consciousness have been focused on the neurons. Yet some attention is now shifting to the glial cells and in particular the astrocytes. Clearly the old dogma that these are essentially inert cells functioning in physical support roles is fading. Ongoing investigation has uncovered many new and unexpected roles for these cells and as a result the aesthetically pleasing forms of astrocytes, which have been known and admired by neuroanatomists since the time of Ramón y Cajal, can now begin to be coupled to function. The time has come to consider astrocytes equal partners with neurons in both the normal and the abnormal brain.

FURTHER READING

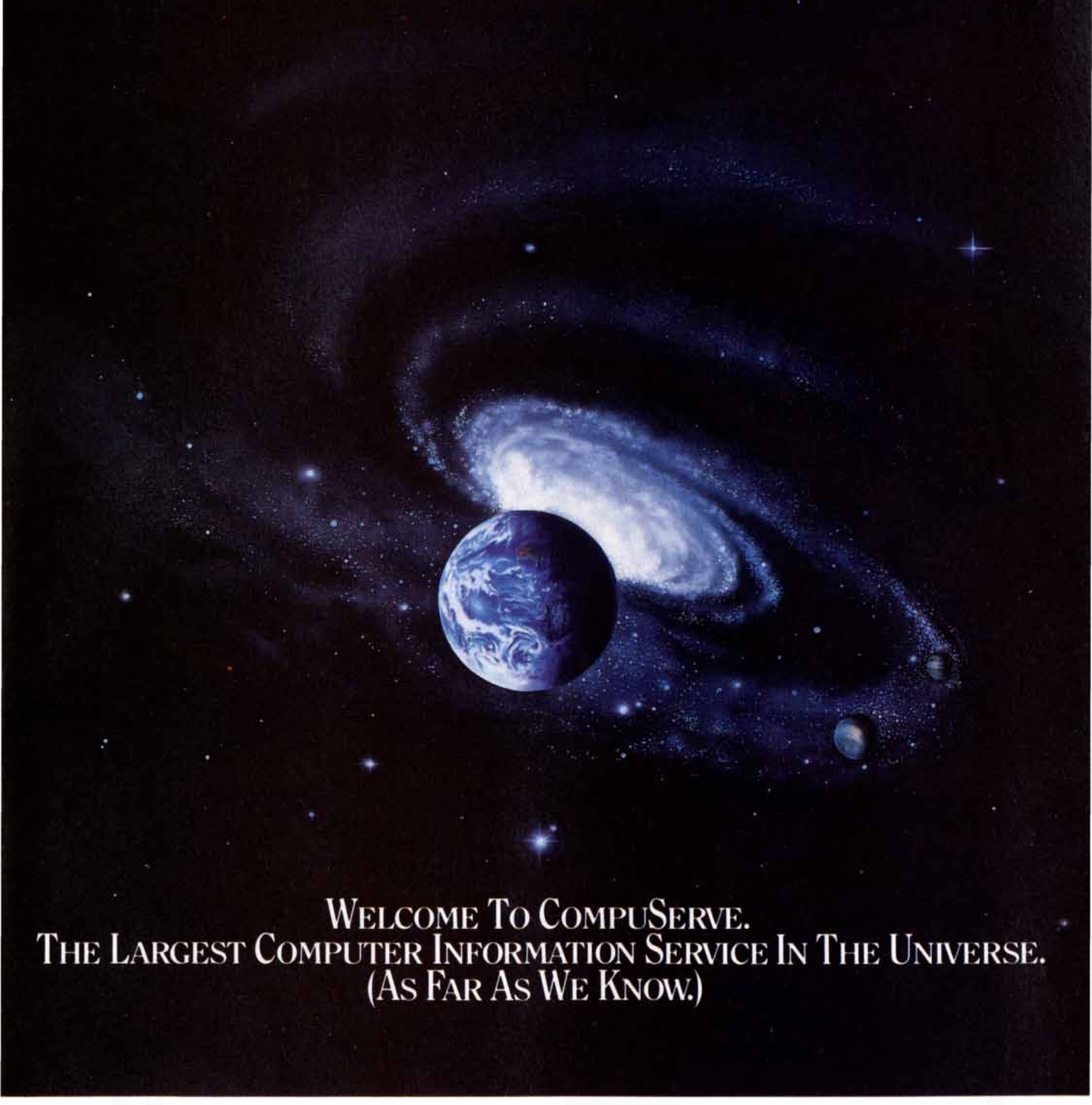
ASTROCYTES. L. Hertz in *Handbook of Neurochemistry,* edited by Abel Lajtha. Plenum Press, 1982.

PRIMARY ASTROCYTE CULTURES—A KEY TO ASTROCYTE FUNCTION. H. K. Kimelberg in *Cellular and Molecular Neurobiology,* Vol. 3, No. 3, pages 1–16; March, 1983.

ASTROCYTES. Edited by Sergei Fedoroff and Antonia Vernadakis. Academic Press, 1986.

BIOCHEMICAL PATHOLOGY OF ASTROCYTES. Edited by M. D. Norenberg, L. Hertz and A. Schousboe. Alan R. Liss, Inc., 1988.

GLIAL CELL RECEPTORS. Edited by H. K. Kimelberg. Raven Press, 1988.



More people use CompuServe than any other online computer information service on Earth.

Half a million members access hundreds of information, communications, and entertainment services online.

Thousands with similar interests and special expertise converge regularly and exchange ideas on an ever-expanding list of special-interest Forums.

They have access to a combination of more financial information, a greater variety of shopping services, and deeper research resources than any other online computer service. Anywhere.

Of course, it's conceivable that there's a service like ours somewhere that could give you more for your money. But you may have to travel a few light-years to find it.

Instead, all you have to do is visit your nearest computer dealer today. To order CompuServe direct, or for more information, write: CompuServe, 5000 Arlington Centre Blvd., Columbus, OH 43220. Or call, 800 848-8199.

CompuServe®

An Alliance for Progress?

European technology companies hope for a boost from 1992

The rallying cry in many European countries these days is "1992." In that year trade barriers between European countries are supposed to crumble, leaving a single market of 320 million consumers spanning the 12 countries of the European Community (EC). Companies will be unleashed to compete on an equal footing throughout the unified market; in the expanded marketplace, Europe's high-technology companies are expected to grow strong enough to rival the best in Japan and the U.S.

At least that is the way it is supposed to work. From the vantage of technology manufacturers in particular, however, building a single market is a painstaking process. Walter Gründsteidl, the adviser on international affairs at Philips, sees 1992 as "a symbol, a significant step forward. But it's not like switching on a light." The disarray in Europe's telecommunications industry provides a telling example of just how stony the path to 1992 is. "No one can pretend that the present situation in European telecommunications is clearly outlined," says Gerhard Zeidler, general manager of Standard Elektrik Lorenz, a French-German telecommunications company.

The telecommunications industry is potentially among the greatest beneficiaries of the unified marketplace. In 1987 the Commission of the European Communities (the body that in 1985 laid the groundwork for unification) issued a special Green Report on the industry. The report noted that while no single European country amounted to more than 6 percent of the world market for telecommunications, the EC as a whole represented 20 percent. (The U.S. has 35 percent of the market and Japan 11 percent.)

Europe has long hoped to remold its fragmented market into one in which manufacturers could gain meaningful economies of scale. The Treaty of Rome, which marked the EC's first efforts at unification, was signed in 1957, but until four years ago its 12 signatories had taken only tentative steps toward a unified market. (The EC's members are Belgium, Denmark, Eire, France, Greece, Italy, Luxembourg, the Netherlands, Portugal, Spain, the United Kingdom and West Germany.)

Emerging Europe, precious blood, solar sailing

As a consequence telecommunications technology developed separately in each country. Although transnational standards bodies did exist, each country's "post, telegraph and telephone" (PTT) agency established the local standards and protocols, defined what minimum services would be provided and bought most of the equipment—from switching systems to telephones. The fragmentation has meant that switching systems routed telephone calls from one part of Europe to another slowly and with difficulty at best. At worst, incompatible systems developed. Mobile telephones, for instance, could be used only within the range of a specific national or local network.

"Europe turned in on itself" during the 1970's, observes Margaret Sharp of the Science Policy Research Unit of the University of Sussex. In telecommunications and in other industries, a government agency bought its equipment from a domestic supplier, thereby fostering a "national champion"—a local company that would be supported regardless of how inefficient it became. "In Britain we produced our own digital switching system that was five years behind those available for export sale. That's the kind of thing that won't happen next time," says John A. Kay of the London Business School.

In order to gain the economies of scale a unified market will make possible, one producer must be able to sell the same product throughout the EC. At the level of end-user equipment, such as telephone sets, this means broadening each country's standards to accommodate the standards of others, says G. M. S. Kestens, deputy secretary-general of CENELEC, a European standards body devoted to electrotechnology. But standards for complex products such as central-office switches cannot be adapted easily. According to a report from the European Institute of Business Administration (INSEAD), manufacturers of central-office switches will gain economies of scale only when the current generation of technology is replaced.

The EC is therefore putting heavy emphasis on newer communications technologies, including satellites and the Integrated Services Digital Network (ISDN). By enabling users to transmit voice, data and video signals rapidly over a single line, the ISDN should make possible a wide range of new services, including voice and electronic mail, color facsimiles and even high-definition television. To speed standards development for telecommunications, the EC recently established a new standards institute.

The commission will also have to tackle the politically sticky issue of how to standardize the procurement policies of the PTT's. Currently 10 percent of the public telecommunications contracts are open to bids by companies outside the national borders. Widening that margin to 100 percent, a goal of the commission, may initially spell plant closings and higher unemployment in some areas. Moreover, the PTT's themselves are important employers, providing jobs for almost one million people in the EC. (The commission nonetheless predicts a net gain in employment, due in part to the creation of new ISDN services.)

As many companies prepare to take advantage of the expanded marketplace, they are looking for partners or consortiums to share research costs. Although existing consortiums, such as the program on research and development in advanced communications technology for Europe (RACE), target "precompetitive," or basic, research, companies are increasingly seeking to share development costs as well.

Merging with companies in other countries is becoming a popular route toward gaining a better foothold in various national markets as well as increasing the internal resources of the company. Last year the French telecommunications company Alcatel teamed with ITT's European telecommunications division. (Alcatel also owns Standard Elektrik Lorenz.) Ericsson, a Swedish company, has bought a piece of the French public-exchange business. Britain's electronics conglomerates GEC and Plessey merged

their telecommunications divisions last year in a joint venture called GPT. Now the German company Siemens is hoping to gain a stake in GPT.

Siemens and GEC have also sought to take over Plessey outright. Plessey's lines of business include defense electronics, semiconductors and medical technology as well as telecommunications, but it is in the last realm that the greatest benefits for the firms are likely to accrue, says James Eastlake, a senior analyst at Dataquest-U.K. With large firms such as a Siemens-GEC-Plessey conglomerate competing in an expanded marketplace, Sharp says, Europe can achieve the economic success the U.S. and Japan have enjoyed.

An intricate collection of financial, monetary and social questions must be addressed if 1992 is to fulfill its promise. Should there be a pan-European central bank? How should the differing safety standards be consolidated? Mergers become complicated as economic goals conflict with those of national security. For instance, in spite of what the telecommunications sector may gain from a Siemens-GEC-Plessey partnership, the U.K. Ministry of Defense reportedly opposed the takeover as late as February for fear that it would narrow the field of domestic military contractors.

Regardless of the obstacles to creating a common market, the forecast of the commission rings loud in Europe: as a result of the unified market, aggregate gross national product should rise by 4.5 to 7 percent, prices should fall by 4.5 to 6 percent and between two and five million jobs should be created. At the same time, Gründsteidl says, European industry will gain "a platform comparable to that of our competitors." —*Elizabeth Corcoran*

The Burden of Proof

Donated blood runs a costly gauntlet of tests

Before a unit of blood reaches a patient, it is put through a series of tests for infectious diseases. As the list of tests grows and the costs mount, blood-bank officials are left wondering whether blood must be screened for all rare blood-borne diseases. What is more, for those companies that make the test kits, the guidelines blood banks eventually establish may well affect research directions and product planning.

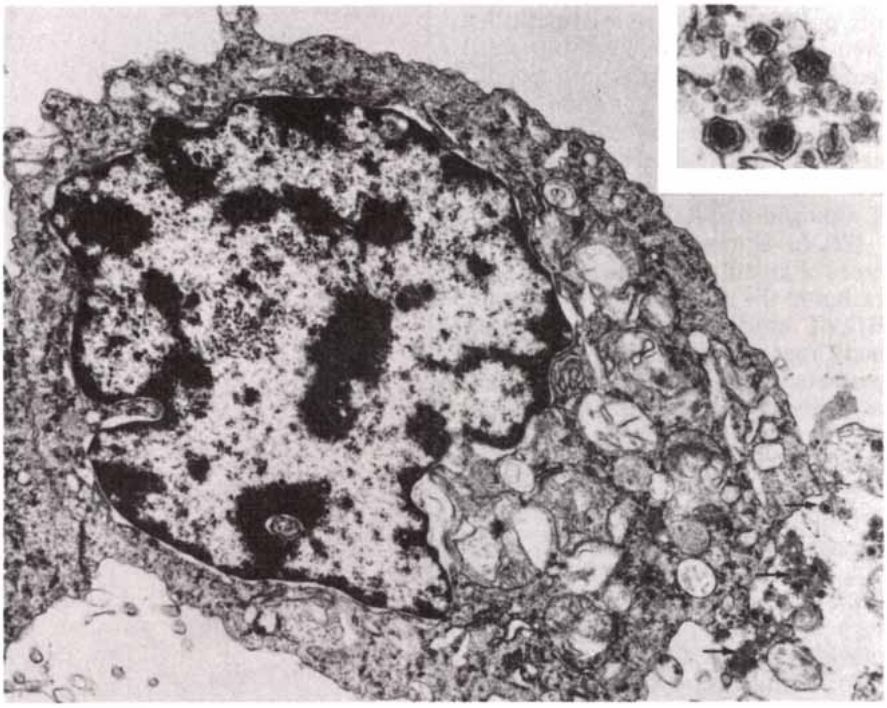
Early this year the majority of blood banks throughout the U.S. began testing for human *T*-cell lymphotropic virus, or HTLV-I, a rare retrovirus associated with life-threatening adult *T*-cell leukemia and with crippling tropical spastic paraparesis. The move raises to five the number of infectious diseases for which blood banks typically test. (The others are syphilis, hepatitis B, AIDS and non-A, non-B hepatitis.)

HTLV-I is now most prevalent in areas of Japan, the Caribbean and Africa. But in the absence of screening for HTLV-I, officials from the American Red Cross predict that at least 2,800 patients of the three to four million who receive blood transfusions in the U.S. every year would become infected. Moreover, since HTLV-I has a latency period of as much as 30 years, officials speculate that it could become widespread if it is not curbed early. "The only prudent action is to intervene while collecting data on the virus," says S. Gerald Sandler, associate vice-president of the American Red Cross.

For manufacturers, producing kits to test the more than 12 million units of blood collected every year in the U.S. is big business. Du Pont and Cellular Products, the two companies that worked with the Red Cross on its initial study of the prevalence of HTLV-I, subsequently received the U.S. Food and Drug Administration's approval to market tests for the virus. Trailing not far behind was Abbott Laboratories. As it became clear that blood banks were eager to test for HTLV-I, Abbott refocused its development work.

At the same time, however, many investigators worry about the costs of adding more tests for infectious diseases in the future. Currently tests are instituted ad hoc, says Thomas F. Zuck, director of the Hoxworth Blood Center at the University of Cincinnati Medical Center. Within the next year several additional tests may become available, including a more specific test for non-A, non-B hepatitis and an HIV-antigen test aimed at finding any blood samples that are infected with HIV but do not contain antibodies to the virus. The known array of diseases transmitted by blood transfusions is likely to grow as well. "It seems that it's only a matter of time until we find another rare disease spread by blood transfusion," Sandler says.

The cost of adding more blood tests builds in several ways. Most officials estimate that it costs about $5 per unit of blood for each test—or from $50 to $75 million in the U.S. every year. There is an even more critical drawback, however: the addition of tests depletes the supply of healthy blood by increasing the number of "false positive" results. Although the



WHITE BLOOD CELL sheds particles of HTLV-I virus* (lower right); *the particles are shown at a higher magnification in the inset. Photographs courtesy of Bernard J. Poiesz of the State University of New York Health Science Center at Syracuse.

Red Cross estimates that fewer than 1 percent of blood donors are infected with non-A, non-B hepatitis, some 5 to 6 percent of all donors fail current tests or produce inconclusive results. Finally, even if the tests improve, Zuck says, anything that adds to the complexity of the work done by blood-center technicians is likely to increase the chance of mistakes. "Any step to go to a zero-risk blood supply is flawed because it's futile," Zuck adds.

Recognizing that blood banks hope to find alternatives to instituting additional tests, companies are pursuing a variety of strategies. Cellular Products, which is devoted to diagnostics tests, is awaiting the outcome of a Patent and Trademark Office review of a combination HIV and HTLV-I test that would help to simplify the task of screening; Du Pont has already bought the rights to market the test. Du Pont, as well as investigators sponsored by the National Heart, Lung, and Blood Institute, is exploring means of inactivating a range of viruses without damaging the blood. Abbott is looking for ways to automate testing procedures. Other workers hope that growth factors, such as erythropoietin, might increase red-cell production in some patients and eliminate the need for transfusions.

Meanwhile the blood banks are organizing themselves to cope with the challenges. Late this spring the American Association of Blood Banks plans to convene investigators to consider blood-transfusion issues. But until guidelines are laid it will be difficult to avoid adding new tests, Sandler says, if they offer a reasonable promise of making transfusions safer. —*E.C.*

PROTOTYPE SOLAR SAIL, made of Mylar, might someday soar in space powered only by photons. The circular bus at the center would carry instruments. Photograph courtesy of Robert L. Staehle of the World Space Foundation.

On Gossamer Wings

This anniversary celebration might surprise even Columbus

Christopher Columbus would have understood: new technology, an epic voyage, a search for financial backers. Welcome to 1992, an auspicious year for many reasons, not the least of which is the 500th anniversary of Columbus' voyage to the New World. In charge of the U.S. celebration is the Quincentenary Jubilee Commission, established by former President Reagan. How better to celebrate exploration than by funding a mission as daring as Columbus' own?

Last year, as the commission was planning an array of exhibitions and conferences to commemorate Columbus, a local businessman and space enthusiast, Klaus P. Heiss, suggested a solar-sailing race. Although the race is still far from reality, the technology of solar sailing is not as far out as it sounds.

As currently envisioned, a solar sail is a huge, lightweight spacecraft made of metallized Mylar or Kapton, more like a silvery kite than a Spanish galleon. Much as a conventional sail is propelled by wind, a solar sail is pushed by photons of light through the vacuum of space. The craft starts off slowly; in an orbit around the sun, every second a sail measuring 150,000 square meters might accelerate by only .5 millimeter per second. At that rate it would take the sail about eight hours to reach a speed of 50 kilometers per hour. But since a sail would be launched more than 1,000 kilometers from the earth, it would encounter almost no air resistance and so would gain momentum continuously.

The unmanned race Heiss proposed would have various stages. "Anyone who gets a sail up there would be a winner," says James W. Symington, a

former U.S. congressman who is on the commission's space-sail committee. Additional prizes would go to the fastest sails to reach the moon and, eventually, Mars.

The commission hopes to encourage student entries in particular, partly by helping to fund the construction and launch costs of the best designs. "The idea is to have [at least] one sail from Europe, where Columbus hailed from, one from Asia, where he thought he was going, and one from the Americas, where he ended up," Symington says. Not surprisingly, the commission intends to name the ships *Niña, Pinta* and *Santa Maria.*

At least one likely American entry is already on the drawing board. In the mid-1970's the National Aeronautics and Space Administration devoted more than $8 million and two years to designing a sail that could rendezvous with Halley's comet. Even though NASA eventually scuttled both the solar sail and the Halley mission, engineers at the Jet Propulsion Laboratory who had done much of the work were not ready to give up the ship. In 1979 they called Robert L. Staehle, a JPL manager who had just started the World Space Foundation, to pursue a variety of space projects. Soon the foundation devoted its efforts to building a solar sail. Working evenings and weekends, the motley crew of volunteers designed a 900-square-meter sail. Although this sail would be too small to carry much payload, Staehle calculates that larger sails could economically ferry supplies to other planets.

Heiss also reports that solar-sailing teams are being formed in Canada, France, India, Japan and the Soviet Union. The race has attracted the attention of other space enthusiasts too. The American Institute of Aeronautics and Astronautics, for instance, has agreed to judge the technical merit of the designs. The Explorers Club, a New York–based society that counts many astronauts among its members, may provide a winner's cup. "We hope the president's space council will endorse it," Symington adds.

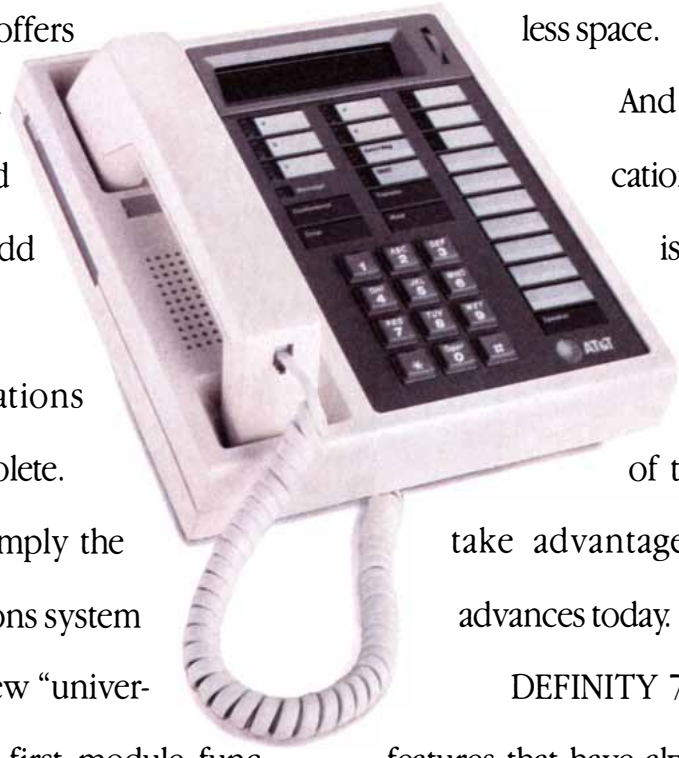
Nevertheless, as Columbus would have attested, attracting funding for the venture will be the greatest hurdle. Symington is hoping U.S. companies will contribute at least $15 million; European and Japanese companies will also be tapped to support sails. Heiss points out that, after all, "one entry in the America's Cup costs $15 million." Of course, since it would take between two and five years for a sail to reach Mars, television coverage of the event will probably be spotty. —*E.C.*

Hello D
goodbye ob

Imagine a communications system designed to expand right along with your business, no matter how large you grow or how complex you get. One that offers seamless communication from 40 to 30,000 lines and beyond. And allows you to add lines economically.

Imagine a communications system that may never be obsolete.

DEFINITY™ 75/85 is simply the most advanced communications system you can own. With AT&T's new "universal module" hardware, your first module functions as the system's base. After that, adding lines is simply a matter of adding modules.



AT&T's new high-density circuit packs allow you to add more phones and computers to your system, while requiring less power and less space.

And of course the communication system of the future is fully compatible with the telecommunication technology of the future, so you can take advantage of emerging ISDN advances today.

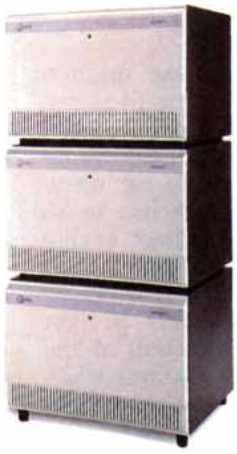
DEFINITY 75/85 also supports the features that have always made AT&T communications systems so popular. AUDIX voice mail, for example, offers you twenty-four-

EFINITY, solescence.

hour access for retrieving and leaving messages. And AT&T's System Management offers you a range of tools, including windowing capabilities, to administer and maintain DEFINITY 75/85.

All of which add up to increased business productivity.

With DEFINITY 75/85 you've got the very best in investment protection. And if you're a current System 85 or 75 customer, you can upgrade to the DEFINITY 75/85 line while protecting up to 90% of your installed communications investment.



The DEFINITY 75/85 was designed to evolve as your business requires. With it there'll be virtually no limit on how large you can grow and still communicate effectively.

To find out more about DEFINITY 75/85, contact your AT&T Account Executive, or call 1 800 247-1212, Ext. 190.

From equipment to networking, from computers to communications, AT&T is the right choice.

AT&T
The right choice.

Free-Electron Lasers

Electrons can impart energy to light waves with the help of a magnetic field. The resulting intense beam can probe crystalline structures—and perhaps destroy missiles in space

by Henry P. Freund and Robert K. Parker

In theory the free-electron laser is an extremely adaptable source of light. It is efficient, it can be tuned to virtually any wavelength, it operates at high power and, of course, it produces coherent radiation. Gas and solid-state lasers, in contrast, can generate light only at specific wavelengths corresponding to energy transitions in their lasing media. Dye lasers can be tuned over a narrow range but require a gas laser for optical pumping and can operate only at relatively low power levels. Furthermore, whereas conventional lasers typically convert only a few percent of their input energy into light, free-electron lasers have potential efficiencies as high as 65 percent. Free-electron lasers have been used in experiments ranging from solid-state physics to molecular biology, and versions are under development for a variety of strategic defense missions, including that of directed-energy weapons.

In practice, however, free-electron lasers have been largely confined to the laboratory. Most have been built around available electron accelerators. Although FELs have the potential to emit light at wavelengths that range from the microwave to the ultraviolet, investigators have encountered difficulties in getting them to lase at visible and shorter wavelengths. Only recently have free-electron lasers begun to come into their own; accelerators are being designed for their specific needs, and facilities are being set up so that workers in other disciplines can take advantage of this new source of intense light.

HENRY P. FREUND and ROBERT K. PARKER collaborate on free-electron laser research. Freund is a senior research physicist at the Science Applications International Corporation in Virginia. He received his Ph.D. in physics from the University of Maryland at College Park in 1976; since then he has worked on the generation of coherent radiation by both laboratory and astrophysical plasmas. Freund began intensive research on FELs in collaboration with physicists at the U.S. Naval Research Laboratory in 1980. Parker is head of the vacuum electronics branch of the NRL's electronic science and technology division. He has done research on the generation of coherent radiation in plasmas since joining the NRL in 1972. Parker received his Ph.D. in nuclear engineering (with a concentration in plasma physics) from the University of New Mexico in 1973.

In a free-electron laser high-energy electrons emit coherent radiation, as they do in a conventional laser, but (as the name suggests) the electrons travel in a beam through a vacuum instead of remaining attached to the atoms of a lasing medium. Because the electrons are free, the wavelength of the light they emit is not confined to a particular wavelength corresponding to a permitted transition between two energy levels of an atom. In quantum-mechanical terms, the electrons emit light by shifting from one continuum energy level to another; the process can be described, however, by classical electromagnetic theory alone.

The light is produced by an interaction among three elements: the electron beam, an electromagnetic wave traveling through the laser cavity in the same direction as the electrons, and a spatially periodic magnetic field produced by an assembly of magnets known as a wiggler. The magnetic field of the wiggler acts on the electrons so that they give up energy to the electromagnetic wave. The energy the electrons give up amplifies the wave, which is then emitted by the laser.

When a light wave moves through an undulatory magnetic field such as that produced by a wiggler, the spatial variations of the wiggler field combine with those of the light wave's electromagnetic field to produce a beat wave, which is essentially an interference pattern between the two. The beat wave has the same frequency as the light wave, but its wave number (a measure of the number of wavelengths in a given distance) is the sum of the wave numbers of the light beam and the wiggler field.

The beat wave has the same frequency as the light wave but a larger wave number (and thus a shorter wavelength); therefore it travels slower than the light wave and consequently is called a ponderomotive wave. Since the electromagnetic field of the ponderomotive wave is the combination of the light wave and the stationary field of the wiggler, it is the effective field that an electron sees when it passes through the free-electron laser. If an electron is moving at the same speed as the wave, it will see a constant field: that of the part of the wave with which it is traveling.

A good analogy to the interaction between electrons and the ponderomotive wave is the interaction between surfers and a wave approaching a beach. If the surfers remain stationary in the water, the incoming wave will merely lift them up briefly and then let them down to their previous level. But if the surfers "catch the wave" by paddling so as to match its speed, they will be able to gain significant momentum from the wave and be carried inshore. (In a free-electron laser the electrons amplify the wave, and so the situation is more analogous to the surfers "pushing" on the wave and increasing its height.)

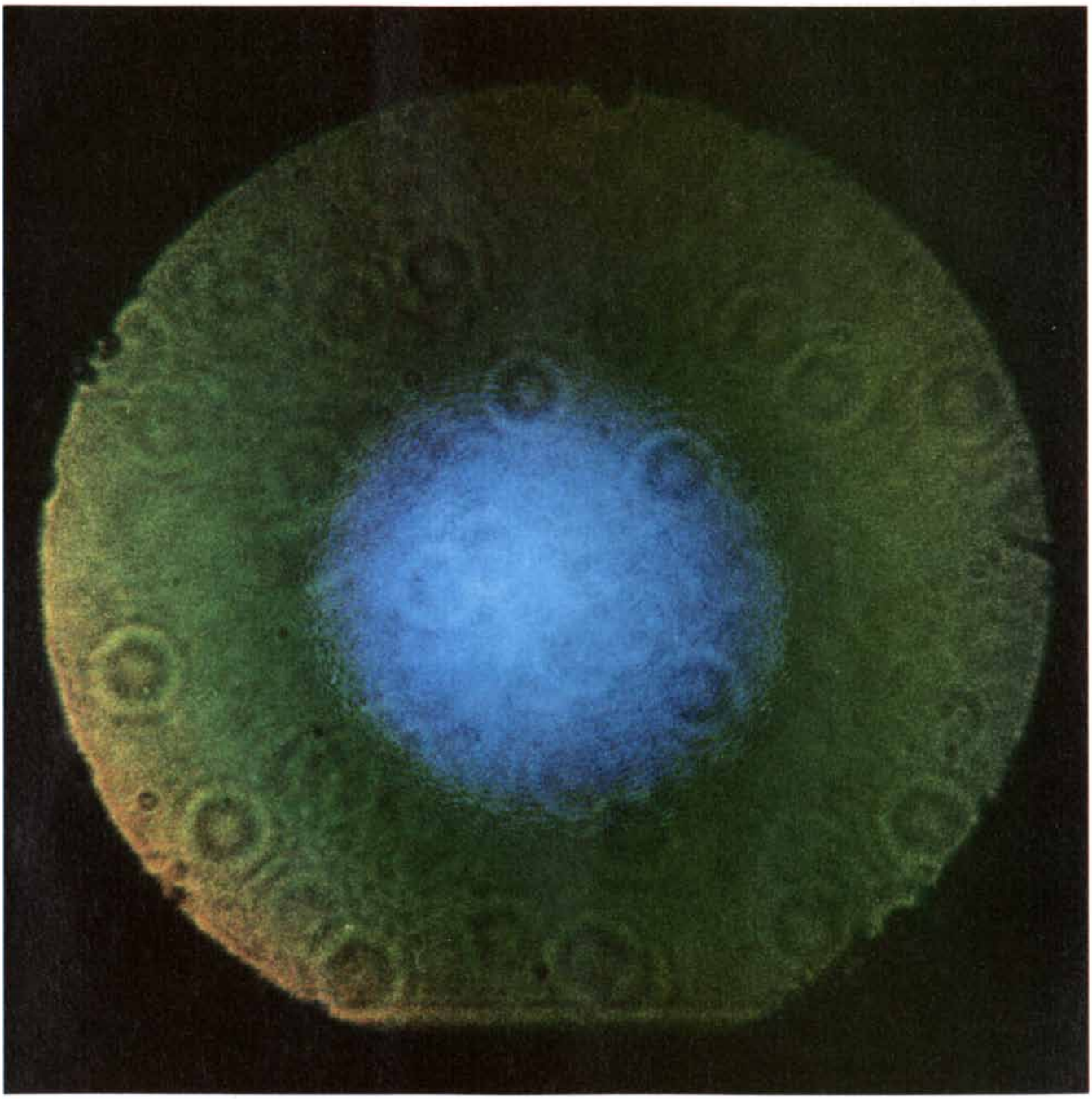
How do a transverse magnetic field and a forward-propagating electromagnetic wave, whose electric and magnetic fields are directed perpendicular to the direction of propagation, give rise to an axial force that can extract energy from the electron beam? An electron moving through a magnetic field experiences a force that acts at right angles to both the direction of the field and its direction of travel. When an electron enters the transverse field of the wiggler, it experiences a trans-

verse force and acquires a transverse velocity. The interaction of the transverse wiggler-induced velocity with the magnetic field of an electromagnetic wave induces a force perpendicular to both—in the axial direction. This is the ponderomotive force.

An electron moving faster than the ponderomotive wave will be traveling against the direction of the ponderomotive force and so will slow down. The total energy of the wave-particle system must be conserved, so that the energy lost by the electron is gained by the wave. As a result the light wave passing through the cavity is amplified by the energy from the electrons.

The amount of amplification depends on the difference between the electron velocity and the velocity of the ponderomotive wave as well as on the strength of the interaction between the electron and the wave. If the electrons are moving at almost the same speed as the wave, they will be able to give up only a little energy to it before they slow down and stop passing wave crests. On the other hand, if the electrons are moving either much faster or much slower than the ponderomotive wave, the interaction between the two will be slight.

As the electrons and the ponderomotive wave travel through the wiggler together, the electrons lose energy and slow down until they are no longer able to pass the crests of the ponderomotive wave. The wave continues to decelerate the electrons until



FREE-ELECTRON LASER at the University of Paris in Orsay produced this intense pulse of light. The intensity of the FEL's internal magnetic field and the energy of the electrons passing through it determine the light's wavelength; the blue color here is due to coherent laser emission and the green is incoherent spontaneous radiation. Because an FEL can be tuned to almost any wavelength, the devices are beginning to find use in such diverse fields as solid-state physics and surgery.

they are moving slower than the wave; then the next crest of the wave advances on the electrons, accelerating them again, so that they are trapped and bounce back and forth in the troughs of the ponderomotive wave. At this point amplification stops. The electron beam has been transformed from a high-energy beam to a lower-energy, bunched beam, where regions of high electron density alternate with regions of low density. To return to the surfing analogy, an ocean filled with a relatively uniform concentration of surfers has been replaced by one where most of the surfers have "caught" some particular wave.

The quality of the initial electron beam is crucial to the operation of this trapping mechanism. The lower the temperature of the beam—the narrower its velocity distribution—the better the performance of the free-electron laser. If the velocity spread of the beam is too large, many of the electrons in it will not give up a significant part of their energy to become trapped by the ponderomotive wave; instead they will simply speed through the wiggler with no net change in velocity. The acceptable velocity spread depends on the specific beam and wiggler parameters. The constraints become stricter as the operating wavelength decreases, and they pose severe difficulties to operation at ultraviolet wavelengths.

In its simplest mode of operation, an FEL acts as an amplifier that increases the power of an electromagnetic wave passing through the cavity. A free-electron laser can also serve other functions. It can operate as an oscillator: the electromagnetic wave is reflected by mirrors at the ends of the wiggler cavity, so that the radiation makes multiple passes through the system, receiving more energy from the electron beam on each pass. It can even function as a superradiant amplifier, in which the electron beam enhances random electromagnetic waves ("shot noise") traveling through the wiggler cavity.

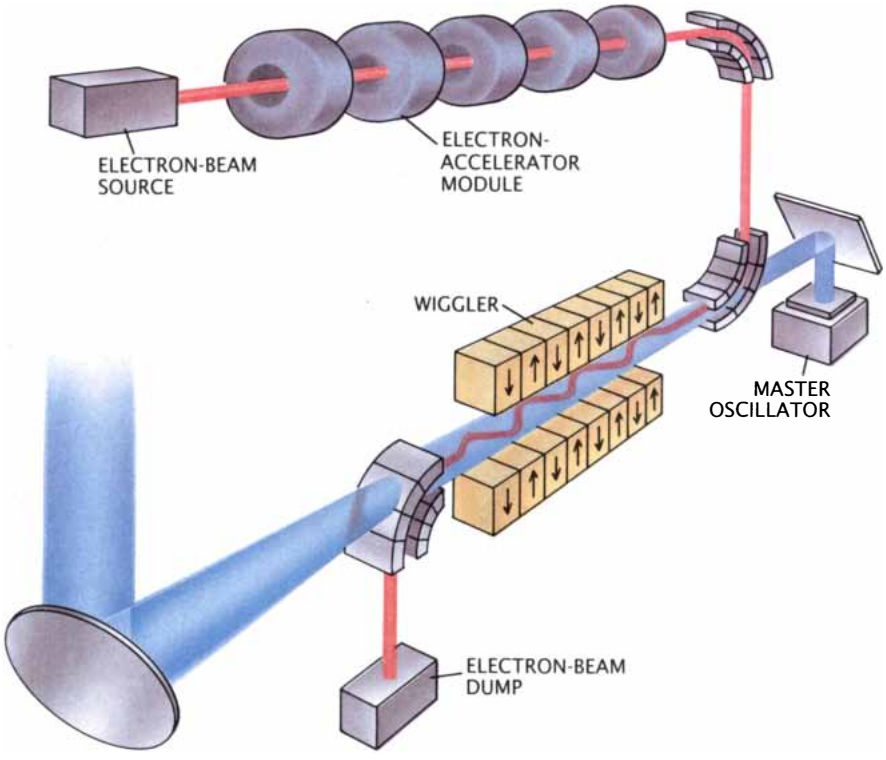
Although the principle by which the free-electron laser works is relatively simple, making the principle work in practice has been arduous. In 1951 Hans Motz of Stanford University first calculated the emission spectrum of an electron beam in an undulating magnetic field. He and his collaborators initially produced incoherent blue-green light; they were later able to obtain coherent amplification at millimeter wavelengths. Coherent emission at visible wavelengths eluded Motz and his co-workers because of the quality of their electron beam.

In 1957 Robert M. Phillips of the General Electric Company independently discovered the application of undulating magnetic fields to microwave amplification. He developed and refined his Ubitron (for Undulating Beam Interaction) over the next seven years, pioneering many design concepts still in use. One system generated 150 kilowatts of coherent microwave radiation at a wavelength of five millimeters. Phillips' timing was unfortunate: the electronics community was shifting from vacuum electronics to solid-state physics and quantum-mechanical devices, and GE terminated Ubitron development in 1964.

There was a resurgence of interest in the free-electron laser in 1975, when John M. J. Madey (who coined the term free-electron laser) and his co-workers at Stanford used a helical wiggler and an electron beam from a linear accelerator to amplify the output of a 10.6-micron CO_2 laser. Advances in electron-accelerator technology and wiggler design made Madey's success possible.

In parallel with the effort at Stanford, experimenters at several sites began work on microwave FELs, successors to the Ubitron. Those projects, at the Naval Research Laboratory, Columbia University and the Massachusetts Institute of Technology, were aimed at producing short pulses at high peak-power levels. Additional projects followed shortly thereafter at the École Polytechnique in France, TRW Inc. and the Lawrence Livermore National Laboratory. (Short pulses are more useful for many applications than longer pulses with the same average power because they deliver more photons to a target before the impinging beam has significantly changed the target's state.) Experimenters used intense electron beams with energies greater than one million electron volts (MeV) and currents of more than 1,000 amperes. Peak power ranged from two megawatts at a two-millimeter wavelength at Columbia to one gigawatt at eight millimeters at Livermore. The Livermore FEL converted 35 percent of the energy in its electron beam into microwave radiation by means of a nonuniform wiggler.

The first visible-light FEL was not built until 1983, at the ACO electron-



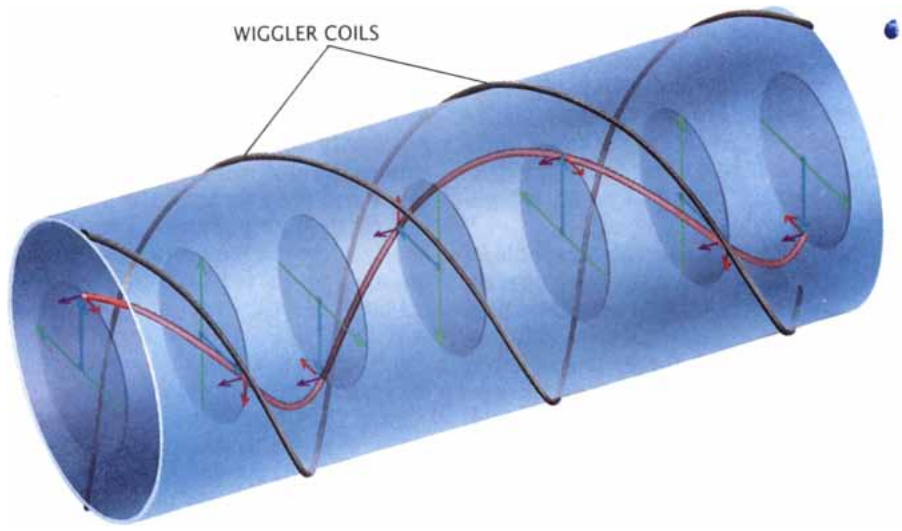
ELECTRONS in a typical FEL are brought up to speed in an accelerator and are then run through a cavity where they interact with an alternating magnetic field to amplify an incoming light beam. The magnetic field is produced by an array of opposing magnets called a wiggler; the wavelength of light that the electrons will amplify effectively depends on the electron velocity and the distance between succeeding poles of the wiggler. The magnitude of the transverse motion induced by the wiggler determines how much the electrons will amplify the light beam.

storage ring of the University of Paris at Orsay. Since then another visible-light FEL has been built around a storage ring at Novosibirsk in the U.S.S.R. Visible-light FELs have also been built at Stanford and at the Boeing Company's aerospace division in the U.S. using radio-frequency linear accelerators (rf linacs).

Today work continues on both optical- and millimeter-wave FELs. The primary goals of investigators are higher power, particularly average power, and shorter wavelength. Achieving these goals will require improvements in electron-accelerator design to provide more intense and more uniform beams, and in wiggler design to extract more energy from electrons and couple that energy to shorter wavelengths of light.

There is a limit to the efficiency that a straightforward FEL equipped with a uniform wiggler can attain. At best the electrons passing through such systems will transfer about 12 percent of their energy to a light wave passing through the cavity. After losing so much of their energy the electrons become trapped by the ponderomotive wave; they slow down to the point where they can no longer effectively transfer energy to the wave. In order to increase efficiency beyond that point some way must be found either to slow down the ponderomotive wave so as to match its speed to the electrons, or the electrons must somehow be speeded up to stay in step with the ponderomotive wave.

Typically this is accomplished by tapering either the amplitude or the period of the wiggler to maintain the forward velocity of the beam. When the electron beam first enters the wiggler, it is moving essentially in a straight line: all its velocity is in the axial direction. The magnetic field of the wiggler causes the beam to bend in the transverse direction, reducing the forward velocity and converting some of it into transverse velocity. The axial velocity, of course, is the component that has to match the velocity of the ponderomotive wave. Gradual reduction of either the strength of the wiggler field or its period from one end of the cavity to the other reduces the electrons' transverse velocity and converts it back into axial velocity, maintaining the forward travel of the electron beam even as the beam loses energy to the light wave it is amplifying. The electrons make smaller transverse excursions, and so they can maintain the axial velocity necessary to keep up with the progress of the ponderomotive wave even while their total velocity decreases.



HELICAL WIGGLER directs an electron beam on a spiral path through the FEL. The transverse and forward velocities of the electrons are constant. The magnetic field of a circularly polarized light wave traveling through the FEL cavity (*blue arrows*) with the electron beam, coupled with the transverse velocity of the electron beam (*red arrows*), produces a force perpendicular to both (*purple arrows*). This axial force slows down the electrons and in return imparts energy to the light wave.

In an FEL with a tapered wiggler, then, trapping the electron beam in the troughs of the ponderomotive wave is not the end of energy transfer. As the wiggler tapers, the trapped electrons can regain axial velocity, causing them to ride up on the crest of the ponderomotive wave and continue the amplification process. In the millimeter-wave FEL at Livermore a tapered wiggler extracted 35 percent of the energy of an electron beam and converted it into electromagnetic radiation, whereas a uniform wiggler in the same system extracted only 6 percent. Theoretical calculations indicate that a tapered wiggler may be able to convert as much as 65 percent of the energy of an electron beam into coherent radiation.

Another consideration in the choice of wiggler design is the tradeoff between a helical and a planar wiggler. Helical wigglers direct the electron beam in a spiral path rather than the sinusoidal one imposed by planar wigglers. The spiral path means that the transverse and axial velocities of the beam remain constant rather than oscillating as they do with the side-to-side motion induced by a planar wiggler. The helical wiggler can induce the same interaction between the electron beam and the ponderomotive wave with about 70 percent of the magnetic field required by a planar wiggler. On the other hand, planar wigglers are easier to adjust, simplifying experiments with different wiggler tapers.

Advanced wigglers can extract more energy from an electron beam, but the best FEL performance requires a high-quality beam as well. Experimenters continue to refine the electron accelerators that supply the beams needed for free-electron lasers; different kinds of accelerators lend themselves best to different kinds of FELs. For example, electron-storage rings produce high-quality, high-energy beams of low to moderate currents, and so they are best for short-wavelength, low-power lasers. FELs using high-energy electrons typically have low gain, because electrons traveling close to the speed of light have large effective masses and thus respond less strongly to the wiggler fields. The electron beam in a storage ring is composed of a series of pulses, each pulse picoseconds long; they circulate through the ring continuously, allowing the FEL to be operated as an oscillator. The FEL light produced by the electron beam from a storage ring consists of short pulses matching those of the beam. These pulses bounce back and forth in the cavity many times to build up their power level.

Radio-frequency linacs use a series of cavities that contain rapidly varying electromagnetic fields to accelerate electrons. They produce beams composed of a sequence of macropulses

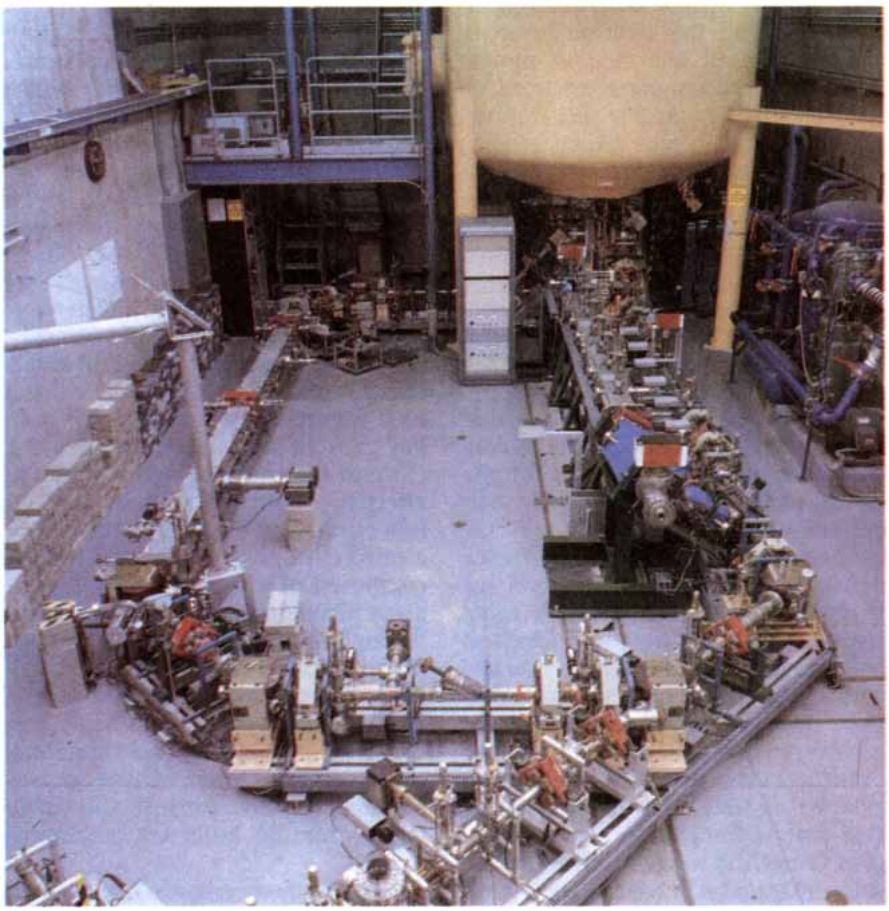
(typically about a microsecond long) each of which consists of a series of picosecond micropulses. These accelerators are now supplying electrons to FEL oscillators that produce visible light. FELs driven by rf linacs emit pulses much like those from FELs driven by storage rings, but at much higher powers.

Induction linacs are also capable of driving high-power FELs; they operate by rapidly changing the magnetic field strength in a cavity, thereby inducing an electromotive force. The electron beam acts as the analogue of the secondary coil in a magnetic transformer. Induction linacs produce beams with higher peak powers than radio-frequency linacs, but they must be fired repetitively to achieve high average power. Experimenters at Livermore have achieved 50-nanosecond pulses with currents of 10,000 amperes and energies of 50 MeV. The Livermore researchers are now developing linacs that could deliver such bursts of electrons 1,000 times a second.

Although their average power is less that that of a linac, electrostatic accelerators can deliver continuous electron beams suitable for FELs emitting beams from the microwave region through the visible. So can microtrons. A microtron is a device that consists of a single radio-frequency accelerating cavity coupled to a magnet that causes the electron beam to describe a circle and pass through the accelerating cavity many times. Electrostatic accelerators and microtrons may be most suited for research and biomedical applications, where the coherence and tunability of light from a free-electron laser is more important than extremely high power.

Even in their current immature state, free-electron lasers are beginning to find application in the research community as powerful sources of pulsed and continuous visible and infrared radiation. They show promise as experimental tools in disciplines as diverse as biomedical science and solid-state physics. The first experimental FEL facility was established in 1984 by Luis R. Elias of the University of California at Santa Barbara. A long-pulse 3-MeV electrostatic accelerator supplies the electrons. The system produces a peak power of 10 kilowatts at far-infrared wavelengths from 390 to 1,000 microns. In one of the first experiments conducted with the device, investigators studied the role of wavelength in light-induced mutations of DNA molecules. The Santa Barbara facility has also been exploited to study the structure of electron energy bands in semiconductors, Stark splitting of spectral lines in intense electric fields, and the linear and nonlinear excitation of phonons (quantized sound waves) in crystals.

More recently Madey has established a user facility at Stanford. This FEL is based on an rf linac; its light is tunable from the visible, at .5 micron, to the mid-infrared at 10 microns. The laser can emit trains of pulses several microseconds in duration, achieving peak powers as high as two megawatts. A third FEL facility is scheduled to open at Vanderbilt University in early 1990. Its laser is projected to be tunable from .2 micron, in the mid-ultraviolet, to 10 microns.



FEL FACILITIES such as this one at the University of California at Santa Barbara have been set up for investigators in many disciplines to use free-electron lasers. Early experiments explored the effect of wavelength on light-induced mutations of DNA.

Free-electron lasers are particularly suited to surgical applications. They perform both cutting and photocoagulation (cauterization). Cutting typically calls for infrared light at three microns, a wavelength that is strongly absorbed by water molecules in tissue but undergoes relatively little scattering. Coagulation, on the other hand, requires wavelengths of from one to 1.5 microns, which are strongly absorbed by water but are also strongly scattered, limiting their range to a thin layer of tissue. In principle an FEL could be tuned to short or long wavelengths as desired in the course of a surgical procedure.

The high power of the Stanford system makes it particularly suitable for studying surgical applications. The laser heats both soft tissue and bone very rapidly, creating a superheated plasma. (The typical cutting mechanism for surgical lasers is thermal ablation, or charring.) The combination of high power and short pulses results in smaller scars and faster healing compared with the effects of surgical lasers now in use.

One possible obstacle to widespread use, paradoxically, is the high power level of FEL pulses. Although they make for good surgery, such barrages of photons may be too much for existing optical fibers to handle. New fibers capable of withstanding higher power may be needed. A less tractable issue is the size and complexity of an FEL system. The high-energy electron beams needed to produce short-wavelength radiation also produce relative-

ly large X-ray and neutron fluxes. The combination of power supply, accelerator, wiggler, optics and shielding is quite bulky. FELs may therefore be limited to medical research centers or special-purpose facilities.

The unwieldiness of free-electron lasers is of less concern to solid-state physicists, for example, who can use the intense beams of monochromatic light to stimulate excited states in crystalline or amorphous solids. Tunability enables experimenters to match photon energy to many different transitions. FELs powered by radio-frequency linacs, microtrons or storage rings are particularly suitable for these kinds of experiments because they produce picosecond pulses of light that are ideal for exciting electrons and following in detail the subsequent decay of their energy states.

High power, coherence and extremely short pulses also make free-electron lasers valuable for any number of experiments in other areas that probe dynamic physical processes: the chemistry of combustion, high-resolution fluorescence spectroscopy and multiphoton ionization of liquids, to name a few. The high intensity of FEL light sources gives them an advantage over synchrotron light sources, which are tunable but incoherent.

In addition to their research applications, free-electron lasers also have applications in such areas as communications, radar and plasma heating. FELs might provide a source of high-power millimeter and microwave radiation for long-range, high-resolution radars. They are also being considered for heating magnetically confined plasmas to produce controlled thermonuclear fusion. Although it was originally thought that plasmas confined in a toroidal magnetic bottle, or tokamak, could achieve the densities and temperatures required for ignition by ohmic heating—the consequence of a current induced in the plasma by an external coil—it has now become clear that additional heating will be needed. Resonant absorption of millimeter-wavelength radiation has been proposed; a free-electron laser is one possible source.

The next major tokamak experiment planned in the U.S. is the Compact Ignition Torus, set to begin operation in 1996. It will require radiation of one-millimeter wavelength or less with an average power of roughly 20 megawatts over a pulse time of nearly three seconds. No existing radiation source can meet that requirement, but FELs have operated in this spectral region at such power levels for shorter pulse lengths. FEL pulse lengths could perhaps be scaled up to meet the needs of a fusion experiment.

A more controversial application of high-power, long-pulse free-electron lasers is strategic defense—including the downing of ballistic missiles. SDI planners envision a large-scale ground-based laser that would direct light toward a target by means of ground-based and orbiting mirrors. Designs based on both amplifier and oscillator FELs are being pursued.

In experiments done last year an FEL at Livermore amplified a 14-kilowatt input signal from a carbon dioxide laser at 10.6 microns to a level of approximately seven megawatts, a gain of 500 times. Boosting the input beam to five megawatts yielded a saturated power of 50 megawatts. In initial experiments investigators used a 15-meter planar wiggler with a uniform period and amplitude; they have since lengthened the wiggler to 25 meters, and work is under way on a tapered version to increase the extraction efficiency further.

Boeing has built an experimental FEL oscillator consisting of a five-meter-long planar wiggler and an advanced radio-frequency linear accelerator. The linac beam achieves energies as high as 120 MeV. The oscillator has lased in the red region of the visible spectrum at .62 micron; power levels were a billion times that of the normal spontaneous emission within the cavity. The average power over the course of a 100-microsecond pulse is about two kilowatts. The corresponding conversion efficiency is about 1 percent, but the peak power is a more respectable 40 megawatts. Even though oscillators typically generate short-wavelength harmonics that can damage the cavity mirrors, no degradation has been observed so far. As with the Livermore experiments, work is under way to modify the uniform wiggler to a tapered configuration to increase the oscillator's efficiency.

Unclassified figures indicate that pulses of visible or near-infrared light roughly one second long, delivering between 10 and 100 megajoules, are necessary to destroy a missile during its boost phase. That means lengthening the pulses or increasing the peak power levels of existing FELs by a factor of a million or more. Depending on laser efficiency and target hardness, a collection of ground-based free-electron lasers would call for somewhere between 400 megawatts and 20 gigawatts of power for several minutes during an attack. (For comparison, a large power plant generates about 1,000 megawatts.) For these and other reasons it is not clear whether it will be practical to scale up free-electron lasers to the power levels required. Current and future experiments will help to resolve this question.

In this application, as in many others, free-electron lasers have a long way to go before they reach technical maturity. The fundamental principles of the free-electron laser are currently well understood, and research directions have turned primarily toward evolutionary improvements in electron-beam sources and wiggler designs. It is important to recognize that the bulk of the work to date has employed accelerators not originally intended to power free-electron lasers; issues of beam quality crucial to FEL operation were not adequately addressed. The only accelerators so far designed specifically for FELs are the rf linac at Boeing, one induction linac at Livermore and the electrostatic accelerator at Santa Barbara.

As accelerators are designed specifically to drive free-electron lasers, their characteristics will influence other parts of the system, such as wiggler design. In addition to tapering, the most important direction in future wiggler design is the development of short-period wigglers. Such devices could generate short-wavelength light with relatively low-voltage electron beams. Lower voltage requirements would simplify accelerator design and lessen the need for shielding against X rays produced by the beam. That would in turn reduce the size and complexity of FELs, opening doors to widespread application of their intense, tunable light.

FURTHER READING

EXPERIMENTS ON RADIATION BY FAST ELECTRON BEAMS. H. Motz, W. Thon and R. N. Whitehurst in *Journal of Applied Physics,* Vol. 24, No. 7, pages 826–833; July, 1953.

FIRST OPERATION OF A FREE-ELECTRON LASER. D. A. G. Deacon, L. R. Elias, J. M. J. Madey, G. J. Ramian, H. A. Schwettman and T. I. Smith in *Physical Review Letters,* Vol. 38, No. 16, pages 892–894; April 18, 1977.

NEW SOURCES OF HIGH-POWER COHERENT RADIATION. Phillip Sprangle and Timothy Coffey in *Physics Today,* Vol. 37, No. 3, pages 44–51; March, 1984.

FREE-ELECTRON LASERS. Thomas C. Marshall. Macmillan Publishing Co., 1985.

HISTORY OF THE UBITRON. R. M. Phillips in *Nuclear Instruments & Methods in Physics Research,* Vol. A272, No. 1, pages 1–9; September, 1988.

Sensory Function in the Harbor Seal

Amphibious in its habits, the harbor seal divides its time between land and water. In the ocean the seal must navigate, locate prey and keep track of its pup; on land it gives birth, nurses its young and rests

by Deane Renouf

At night seals navigate through murky waters to find fish; during the day they often haul out on land, where they lie in the sun and, once a year, give birth to their young. Dividing time between land and water in such a fashion has its price: seals, like other members of the Pinnipedia (the order to which seals, sea lions and walruses belong), have had to adapt to two separate sets of physical challenges. Sound and light behave differently in air and in water, and sensory organs that are adapted for one habitat tend to function differently in the other. Consider the seal's eye. How, if it is designed to function in water, does it also function on land? On what types of special sensory receptors does the seal rely to find its way through turbid, choppy waters?

For the past 19 years I have been studying sensory function in the harbor seal, *Phoca vitulina*. The species is a good model for understanding how pinnipeds in general have adapted to an amphibious existence and is of special interest to me because it is found in coastal waters near my home in Newfoundland. Still, after many years of observing the animal—both in its natural habitat and in captivity—I continue to be puzzled by some of its sensory capabilities.

DEANE RENOUF is associate professor of biopsychology at Memorial University of Newfoundland, where she has been a member of the faculty since 1975. She received a B.A. at Memorial University and a Ph.D. in psychology from Dalhousie University in Nova Scotia. She and her husband, who is a general practitioner, have four children. They keep her busy when she is not observing seals and teaching animal behavior. Ironically, her recreational activity is a daily workout in the swimming pool.

Although it is a common species along the northern coasts of the Atlantic and Pacific oceans, *P. vitulina* is notoriously difficult to study under natural field conditions. It is extremely skittish and flees into the ocean at the slightest provocation. Such behavior makes quantitative studies nearly impossible and has discouraged many investigators. I have been lucky, however, because I have found a site on Miquelon Island (some 18 kilometers off the southeast tip of Newfoundland) that is uniquely suited for observing seal colonies at close range.

During the reproductive season—in late spring and early summer—when the tide is low, about 800 seals (both male and female) cluster on exposed sand flats near the center of a large lagoon called the Grand Barachois. Females give birth at this time and remain with their pups until they are weaned at about four weeks of age. The seal's daily activities are synchronized with the tide: when the tide is high, the seals (including mothers and pups) are forced off the sand flats and into the water, where they stay until the tide recedes and the flats are again exposed. My colleagues and I have constructed elevated observation blinds adjacent to the sand flats, from which we can watch the seals at close range without disturbing them. We discovered that if we enter the blinds during high tide, the animals will pay us little or no attention when they return at low tide. From this vantage we have been able to observe how the seals have adapted to several crucial environmental challenges.

John W. Lawson, one of my graduate students, was the first to document controlled labor in the harbor seal, a physiological adaptation that allows females to accelerate or delay a pup's birth according to environmental conditions. On three occasions when a female in the final stages of labor was disturbed by the arrival of a group of tourists, Lawson saw the emerging head of the pup disappear back into the birth canal and labor come to a halt, resuming only after the disturbance had passed. We suspect that the ability of a seal to control the timing of her pup's birth is an adaptation that minimizes the risk of predation and enables seals to synchronize their labor with the onset of low tide.

While we were observing the breeding colony from our elevated blinds I became interested in the close relation that exists between females and their pups. The seal's amphibious habits and skittish behavior make bonding between mothers and their offspring somewhat problematic. I have seen a pup less than 15 minutes old follow its mother into the ocean, where visibility is often low, the current is strong and the level of ambient noise (caused by wind, choppy water and turbulence) is high. If a mother and her pup become separated, the likelihood of reunion is slim; this prediction is underscored by the fact that in some colonies as many as 10 percent of the unweaned pups starve to death every year when they are separated from their mothers.

How—if the seals rush into the ocean at the slightest disturbance—does a newborn pup manage to stay with its mother? It appears that a number of factors are responsible. The harbor seal pup demonstrates a following response (something like imprinting in birds) within the first few minutes of life and will follow its mother wherever she goes. The relationship is reciprocal: the mother in turn monitors the whereabouts of her pup. Females track their pups visually

(in the water they can be seen stretching their heads backward to get an upside-down look at them) and also acoustically: a harbor seal pup vocalizes almost continuously when following its mother, emitting a call that is transmitted in air and underwater simultaneously; the call disappears from its repertoire soon after weaning.

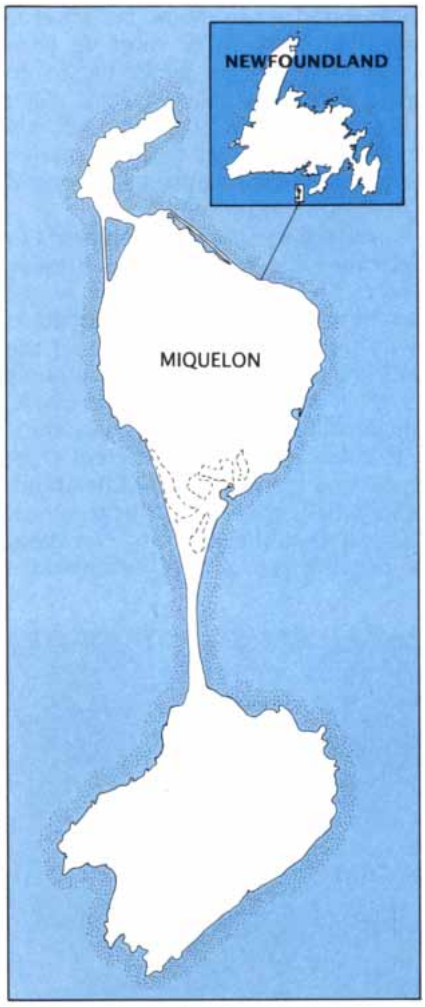
To test whether a mother can recognize her pup by its call, Elizabeth Perry, one of my graduate students, and I recorded the calls of newborn pups and analyzed them sonographically. We found that each pup's call has a unique frequency pattern and wondered whether these differences could be discerned by a female seal.

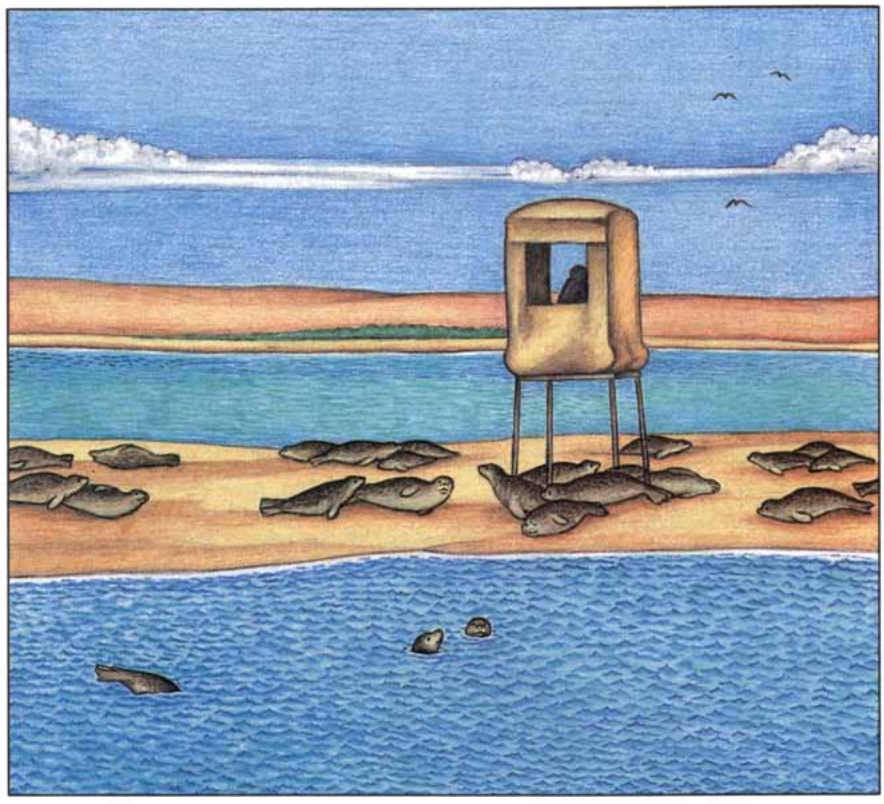
To answer that question we devised an experiment to test a female's ability to distinguish among calls. At the Ocean Sciences Centre of Memorial University we trained a captive seal to open the door of a specific feeder when she heard one call and to open the door of another feeder when she heard a different call. Every time she made a correct association we rewarded her with herring. Six different calls were presented in various combinations; after a brief training period she was able to distinguish among the calls at least 80 percent of the time, a finding that leads us to believe females can recognize their pups in the ocean by their vocal emissions.

Vocalization is clearly an important means by which a mother and pup stay together, and yet the harbor seal lives in an environment dominated by high noise levels. How can a mother hear the call of her pup when the background noise (above water) may reach 80 decibels or more (a level comparable to that generated in an urban setting by heavy traffic)? When the ambient noise levels are high, can seals detect sounds that for human beings and other animals are masked by background noise?

I initiated a series of experiments to determine the extent to which noise affects the auditory threshold of the harbor seal. I trained two animals to swim to a paddle on one side of the tank when they heard a tone (a short burst of approximately 25 decibels, which was presented simultaneously with either 60-, 70- or 80-decibel white noise) and to swim to a paddle on the other side if they did not hear a tone. I found that the auditory threshold of a harbor seal is raised when back-

HARBOR SEAL, *Phoca vitulina,* shown off the coast of Newfoundland, has well-developed whiskers, called vibrissae, and eyes that are adapted for seeing underwater. The species is common along northern coasts of the Atlantic and Pacific oceans.

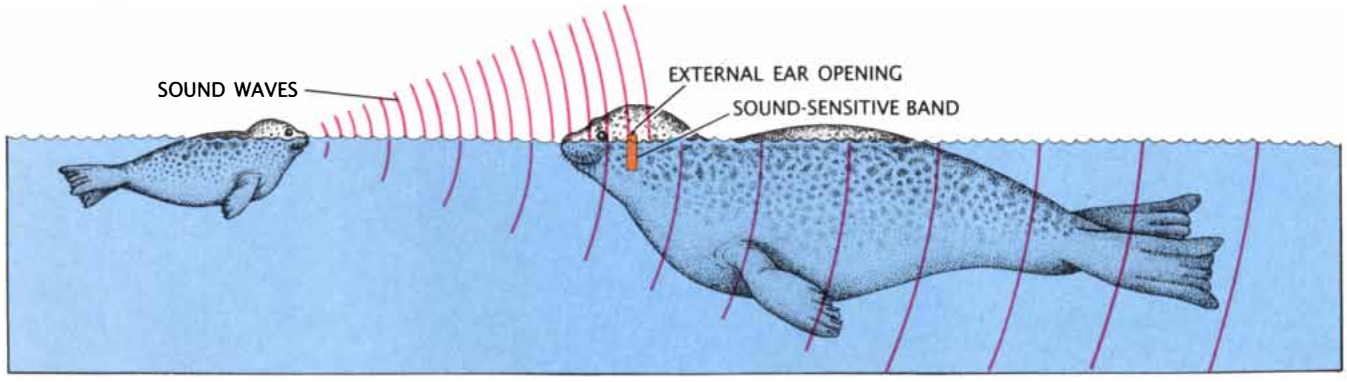




LARGE BREEDING COLONY of harbor seals can be found every spring and summer on Miquelon Island, about 18 kilometers off the southern coast of Newfoundland. The seals haul out on sand flats (*indicated on map by broken lines*) that are exposed during low tide. At high tide, when the flats are flooded, the seals retreat to the ocean. If the author and her colleagues enter the observation blind when the tide is high (*above*), they will be tolerated when the seals return at low tide, a stratagem that allows the author to observe the otherwise skittish seals at exceptionally close range.

ground noise levels are high and that the seal, like other mammals, has no special ability to compensate for noise. Extrapolating from these findings, I calculated that a pup can be heard only if it stays within about eight meters of its mother.

Within that radius of eight meters, however, females are adept at locating their young. I believe such an ability may be attributed to a unique quadraphonic hearing arrangement that enables them to determine from what direction a sound has come. On land sound reaches the seal's inner ear—as it does in most mammals—through the auditory meatus, or canal, and its direction is determined by the difference in arrival time at each ear. (In addition, certain sound characteristics such as volume and wave pattern are



FEMALE SEALS may possess a special quadraphonic hearing arrangement that enables them to locate their pups in the ocean. When a pup cries, sound waves enter the air and water simultaneously. The author believes a mother may rely on the difference in the arrival time, and also in the phase and intensity of the two sets of sound waves, to determine the direction of her pup's call. She can differentiate between the two sets because underwater sound waves, which travel about four times faster through water than through air, reach the seal's inner ear through a vertical band of sound-sensitive tissue (*color*), whereas aerial waves reach the inner ear through the external opening of the ear canal.

affected by arrival time.) In water, where sound travels about four times faster than it does in air, the difference in arrival time is much more difficult to detect.

Bertel Möhl of the Zoological Institute in Aarhus, Denmark, has shown, however, that in water sound is conducted to the seal's inner ear through a special vertical band of tissue that extends downward from the ear. When a mother's head is partially submerged, it is possible that sound passes through both the auditory meatus and the band of auditory tissue, enabling her to hear both the aerial and the underwater version of her pup's call. Because the call arrives at these receptors at slightly different times, she may be able to discern the direction from which it has come more precisely than if she relied only on the underwater or aerial versions of the call.

Harbor seals have interesting visual systems that reflect their amphibious habits. Behavioral studies by Ronald J. Schusterman and his colleagues at the University of California at Santa Cruz and anatomical studies by Glen Jamieson of the University of British Columbia and others show that the seal eye is remarkably well adapted for seeing both underwater and on land. The lens is large and spherical and its shape is suited for underwater acuity. The size and shape of the eye compensate for the fact that the refractive index of water is almost the same as that of the cornea. Consequently light waves entering a seal's eye in water do not refract, or bend, when they pass through the cornea as they do in air. Instead they are refracted only by the lens, which channels them to the retina, or focusing plane, at the back of the eye.

In contrast, human beings, whose eyes function best on land, where the cornea is refractive, see poorly underwater. Without the help of the cornea, light is refracted by the lens so that the visual image no longer forms on the retina and the image is therefore blurred. In seals the visual image forms on the retina and is in focus.

In air the seal's cornea is astigmatic: its curvature is distorted, particularly along the horizontal plane of the eye, and light waves are affected by the distortion as they pass through the eye. In water this astigmatism is of no importance because light there is not refracted as it passes through the cornea. On land the seal compensates by having a stenopaic (vertically contracting) pupil. Because the pupil closes down to a narrow vertical slit that is parallel to the axis of least astigmatism, the most astigmatic area of the cornea has little or no effect on the seal's vision. On foggy or dimly lit beaches the pupil does not contract and the seal has blurred vision. But when light levels are higher, as they usually are near the ocean or on ice, the pupil compensates for astigmatism and the seal's visual acuity in air should be comparable to that in water.

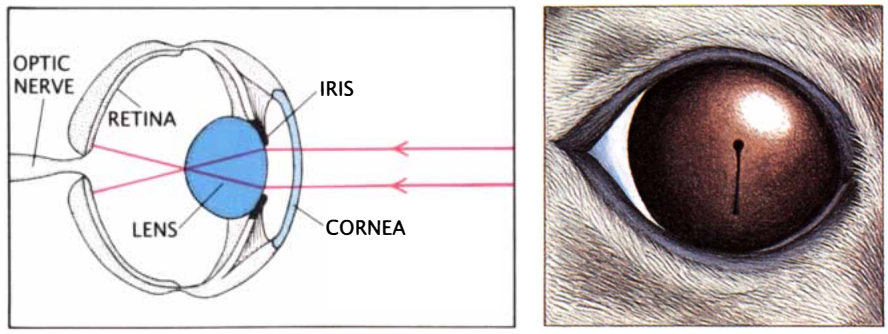
Underwater, harbor seals are extremely sensitive to low light levels; Douglas Wartzok of Purdue University has shown, for example, that on a moonlit night in clear water the seal can detect a moving object at depths as great as 466 meters.

How does the harbor seal, which spends much of its life in murky water where visibility is near zero and which feeds mostly at night, find the three kilograms or more of fish it must catch every day? In the 1960's Thomas C. Poulter of the Stanford Research Institute and others suggested that California sea lions (*Zalophus californianus*) can find and identify prey by echolocation. Echolocation, which was first discovered in bats and was later described in birds, porpoises and dolphins, is similar in principle to radar. Animals that echolocate emit a series of high-frequency sound pulses that reflect off anything they strike; the reflections in turn are processed by the animal's brain, where they form an image that effectively enables the animal to "see" in the dark [see "More about Bat 'Radar,'" by Donald R. Griffin; SCIENTIFIC AMERICAN, July, 1958].

Although no one has conclusively demonstrated that sea lions or other pinnipeds can echolocate, a growing amount of circumstantial evidence suggests that the harbor seal may indeed have that capability. Harbor seals emit click vocalizations: broad-frequency sounds that are produced in short, very fast bursts, most often at night. Recordings I have made of these vocalizations with special audio equipment reveal that many of the clicks are in the ultrasonic range (that is, above the upper limit of human hearing at 20 kilohertz). Working with captive seals, I have found that clicking increases when the seals are fed at night.

In 1968 Möhl was able to show that harbor seals can detect sound frequencies at least as high as 180 kHz and are most sensitive to frequencies of 32 kHz. (Human beings, in contrast, have a sensitivity range from .02 to 20 kHz.) Interestingly, some of the harbor seal clicks peaked in the range of 40 kHz, close to the seal's maximum sensitivity of 32 kHz. On land, where echolocation would be of little use to the seal (which feeds only in the water), the seal is unable to detect sound frequencies much greater than 16 kHz and emits no clicks, instead augmenting its vocabulary with low-frequency growls and snorts.

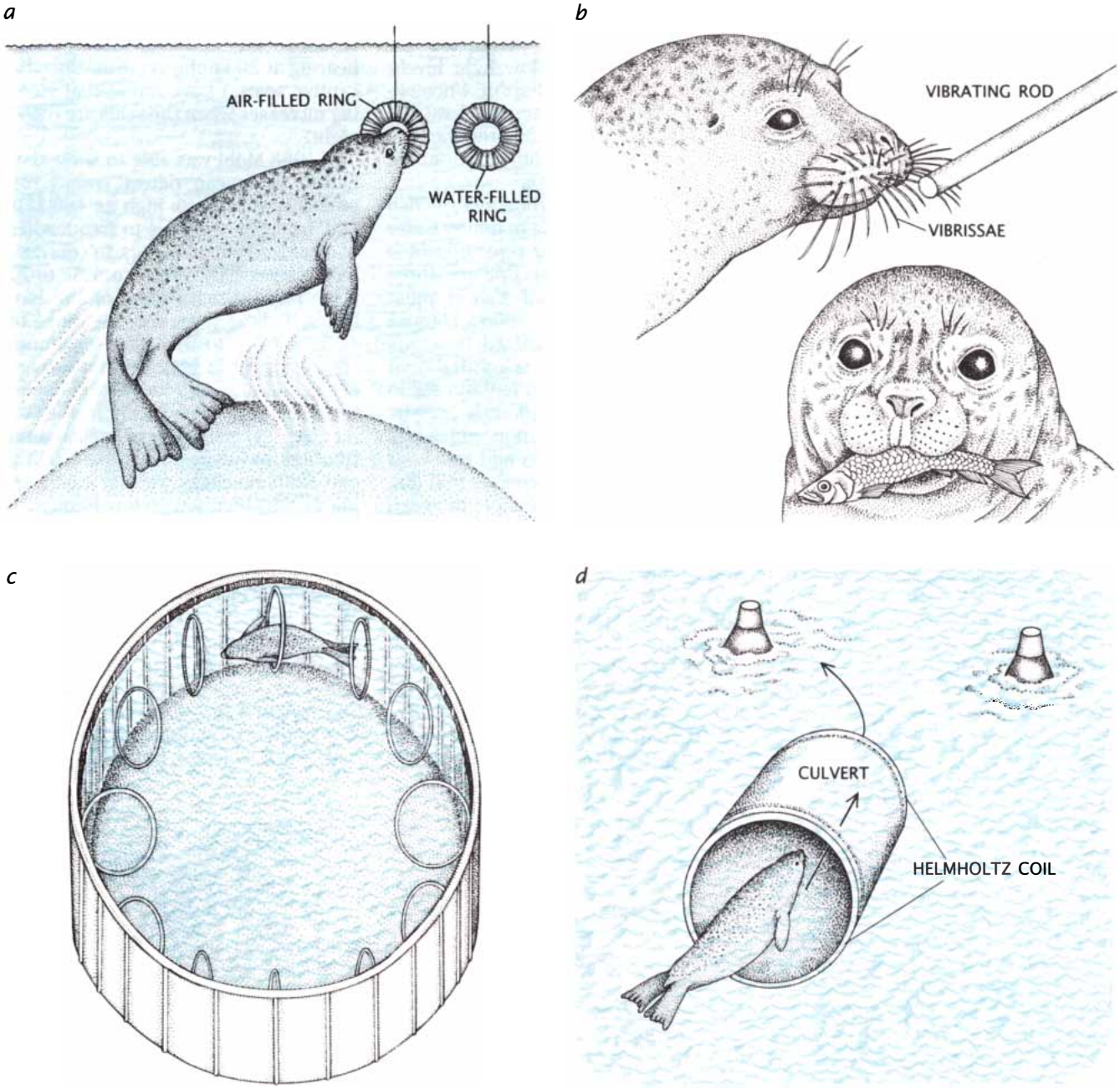
I devised an experiment with Benjamin Davis, one of my graduate students, to test echolocation in our captive seals. We wanted to see if they could distinguish between two doughnut-shaped rings that looked and weighed the same; one ring was filled with water and the other with air (and small weights), but they differed



EYE OF HARBOR SEAL is designed for maximum sensitivity and acuity on land and in the water. The spherical lens (*left*) compensates for the fact that a seal's cornea does not refract, or bend, light waves in water: light retains its original trajectory through the cornea but then bends as it passes through the highly curved lens, striking the retina at an angle that creates a focused image. On land, where the refractive power of the seal's cornea causes light waves to bend as they pass through it, the shape of the lens is disadvantageous because light is focused in front of the retina. The seal compensates by having a stenopaic pupil: when light levels are high, as they usually are near the shore, the pupil contracts to form a vertical slit (*right*). The contracted pupil increases visual acuity in much the same way as stopping down the shutter on a camera increases depth of field.

in their sound-reflecting characteristics. Because an object's density will affect the way sound waves reflect from it, we surmised that the only way the seals could differentiate between the rings was by echolocation.

We began the experiment by teaching a seal to retrieve just one of the rings at night: one of us would slip the ring into the water while the other distracted the seal at the opposite side of the training tank. The seal was then told to fetch; it quickly learned to do so, returning with the ring on its snout in an average of 34 seconds, whereupon we rewarded it with a piece of herring. Having determined that the seal could find a hidden ring and retrieve it without difficulty, we then tested the animal's ability to discriminate between the two rings. The same



VARIOUS FACETS OF SENSORY FUNCTION in the harbor seal were tested by the author and her graduate students in a series of experiments at the Ocean Sciences Centre of Memorial University. In the first experiment (*a*) they wanted to determine whether seals can find prey by echolocation. A seal was presented with two rings that were identical except in density. The seal was rewarded with herring if it could correctly distinguish between the rings. The results suggest that seals may echolocate. Next the author (*b*) tested the sensitivity of the seals' vibrissae by measuring their response to a vibrating rod. The vibrissae responded to high-frequency vibration, but clipping them had no long-lasting effect on the seals' ability to catch fish. Might the vibrissae act as a speedometer? An experiment was devised (*c*) to see whether a seal could swim at constant speed through a series of hoops. The results show that the seal could do so with or without its vibrissae. In the last experiment (*d*) the ability of seals to detect the earth's magnetic field was tested. The seals swam through a culvert in which the electric field could be shifted 60 degrees eastward by passing electricity through a Helmholtz coil. If the seal touched the left buoy when the coil was on and the right buoy when it was not, it was rewarded with herring. The results were inconclusive.

experimental procedure was followed except that when the seal returned with the air-filled ring, it was rewarded with a piece of herring; when it returned with the water-filled ring, a one-minute time-out (punishment for a hungry seal) was declared.

After 26 sessions the seal was able to correctly identify the air-filled ring from 75 to 80 percent of the time, a reasonable indication that it could distinguish between the two rings. We then removed the weights and added water to the air-filled ring, rendering it identical in every way with the other ring. When the experiment was repeated, the seal could no longer discriminate between them. Our results suggest that the seal can echolocate, but we are puzzled by one aspect of our study. Recordings showed that the seal vocalized very little during its search for the water-filled ring and that the clicks it emitted were intermittent and weak. Because of these findings, echolocation in the harbor seal remains unconfirmed.

Does something else explain the fact that harbor seals are extraordinarily adept at navigating and catching prey in their murky ocean habitat? For some years I have speculated that the seal's vibrissae, or whiskers, must be important sensory receptors. Vibrissae, which are present in almost all mammals (except human beings and a few other species), are unusually well developed in seals, sea lions and walruses. Research I have done with the help of my graduate student Fred Mills suggests that the vibrissae are highly sensitive to movement and thus might play a role in food capture.

We had four seals touch a small vibrating rod with their vibrissae while we varied both the frequency and the amplitude of the vibrations. By monitoring the animals' response (we gave them food when they responded to certain vibrations) we found they were most sensitive to higher frequencies (about 2.5 kHz). This finding was somewhat unexpected because it is the opposite of what occurs in other animals, whose tactile systems are most sensitive to lower frequencies. We calculated that at a distance of 43 centimeters the wave created by the tail beat of a herring-size fish would attenuate to the seal's lower threshold, theoretically enabling the seal to home in on the fish and capture it.

These predictions were given partial support when we clipped the seals' vibrissae (they grow back in a few weeks). In a set of before-and-after experiments we compared the speed at which seals could capture live fish when the vibrissae were intact with the speed at which they caught fish when the vibrissae were removed. The removal of the vibrissae had no significant effect on the length of time the seals needed to find and capture a fish in clear water. We repeated the experiment in murky water, and although they still showed no difference in prey-location time, with or without vibrissae, some dewhiskered seals took longer to actually capture fish in their mouth.

What other purpose might the vibrissae serve? In 1967 William Montagna of the Oregon Health Sciences University in Portland suggested they might function as a speedometer: their bending would correspond to the animal's swimming speed. My colleague Linda Gaborko and I set out to test his theory. We began by training a seal to swim at a speed of six kilometers per hour through hoops placed around a 17-meter course. The seal was rewarded with herring for each circuit in which it maintained a constant speed of six kilometers per hour. When it swam either too slowly or too fast, a buzzer sounded and no reward was given. Once we were convinced that the seal could maintain a steady speed (even after a 17-day break in the training), we clipped its vibrissae and repeated the trial. Its swimming speed was not affected; it thus appears that the vibrissae are important sensory receptors, but their precise function has yet to be determined.

Because the seals seem to function in many instances without conventional sensory channels, it occurred to me that they might be sensitive to magnetic fields. It is widely believed that birds can detect the earth's magnetic field and use it for compass orientation; why could the same not be true for seals? To test that hypothesis, we trained captive seals to swim through a hollow, double-walled fiber-glass culvert in which we placed a Helmholtz coil: two hoops wrapped in copper wire through which a current is passed. By sending a current of 2.1 amperes at 70 volts through the coil, we were able to shift the earth's magnetic field within the culvert eastward by 60 degrees. Measurements with a magnetometer revealed that the field was uniformly shifted inside the coil and was not affected 30 centimeters outside it. Two buoys were anchored to a float three meters from the coil, one to the left of the coil, the other to the right. We hoped to train the seals to swim through the culvert on command and to touch the left buoy on emerging from the culvert when the magnetic field inside it was deflected eastward and to touch the right buoy when the coil was not powered.

The inside walls of the culvert were filled with fish oil (which prevented seawater from leaking in and also provided ballast). When the setup was complete, we conducted a number of sessions involving three seals. The first seal (a male) swam through the culvert 2,005 times in more than 30 separate sessions, the second seal (also a male) swam through it 927 times in 17 sessions and the third seal (a female) swam through it 1,227 times in 25 sessions. To our dismay the seals did not differentiate between the two electromagnetic fields.

We cannot conclude, however, that in their natural habitats seals are not sensitive to the earth's magnetic field. It must be emphasized that negative results are always difficult to interpret because they are often brought on by methodological omissions and errors. It could be that the seals failed to respond to the magnetic cues we provided because their environment (a small training tank) was artificial.

Although some aspects of seal biology are now known, many are still puzzling. It is clear that the animal, which traverses dark, murky waters in search of prey and spends considerable time on land, has either unusual sensory receptors or extraordinary cognitive abilities. Whether or not *Phoca vitulina* relies on echolocation in order to detect prey and other objects in the ocean and on the earth's magnetic field in order to navigate are questions awaiting resolution.

FURTHER READING

BIOLOGY OF THE HARBOUR SEAL (*PHOCA VITULINA*) IN EASTERN CANADA. J. Boulva and I. A. McLaren. Bulletin 200, Fisheries Research Board of Canada, 1979.

HARBOUR SEAL. Michael A. Bigg in *Handbook of Marine Mammals,* edited by Sam H. Ridgway and Richard J. Harrison. Academic Press, 1981.

BEHAVIORAL CAPABILITIES OF SEALS AND SEA LIONS: A REVIEW OF THEIR HEARING, VISUAL, LEARNING AND DIVING SKILLS. Ronald J. Schusterman in *The Psychological Record,* Vol. 31, No. 2, pages 125–143; Spring, 1981.

ATTACHMENT BETWEEN HARBOUR SEAL (*PHOCA VITULINA*) MOTHERS AND PUPS. Deane Renouf, John Lawson and Linda Gaborko in *Journal of Zoology,* Vol. 199, pages 179–187; 1983.

SENSORY FUNCTION IN SEALS. D. Renouf in *Mechanisms and Evolution of Behaviour in Seals,* edited by D. Renouf. Methuen & Co., in press.

The Trireme Sails Again

It took years of detective work to reconstruct the fabled Greek warship. Sea trials indicate that the ship might indeed have performed the feats of speed attested to in ancient records

by John F. Coates

In June of 1987, near the ancient Athenian port of Piraeus, a working reconstruction of the Greek warship called a trireme was launched for the first time since this remarkable type of ship disappeared nearly 2,000 years ago. Only a month later, crewed by 170 rowers of diverse nationality, background and rowing skill, *Olympias* was able to reach a top sprint speed of seven knots (13 kilometers per hour). The turning radius for the ship, starting from full speed, was 1.25 boat lengths, or about 46 meters. These capabilities accord with ancient accounts of the feats of triremes. Perhaps the most celebrated is Thucydides' account of a nonstop 340-kilometer voyage in 427 B.C. from Athens to Mytilene that took just over 24 hours, implying a cruising speed of 7.5 knots.

In antiquity the highly maneuverable ship was equipped with a bronze ram, by which the hulls of enemy ships were holed in battle. In 480 B.C. triremes won the Battle of Salamis for the Hellenes of the Greek mainland against a larger Persian fleet. This was one of the decisive battles of Western history, for had the mainland Greeks then fallen under Persian rule, none of the later cultural achievements of Greece, and of Athens in particular, would have come to pass. The trireme continued for another century and a half to play a significant role in providing military defense and protecting trading ships in the pirate-infested Mediterranean waters, thereby safeguarding the conditions that enabled Athens to develop the art, literature and philosophy constituting the Greek legacy to the Western world. For all that, scant material evidence about this singularly important ship has come down to present-day scholars. No trireme wrecks have ever been found, and surviving references in literature and art are fragmentary.

Scholars have debated the ship's actual form and capability since the Renaissance. Many of the most important questions have been answered over the past 50 years by John S. Morrison of the University of Cambridge. By the early 1980's Morrison had gathered and analyzed enough ancient literary, epigraphic and iconographic evidence to arrive at a promising conceptual design for the trireme. At about the same time I had been applying my own experience with warship design to the question. Moreover, underwater excavations near Marsala, Sicily, in the early 1970's by Honor Frost of the Society of Nautical Research in London had turned up crucial evidence about the shape and construction of long, oared ships in the ancient Mediterranean. The combined evidence severely constrained the design. And so it was that in 1981, when Frank Welsh, a Suffolk banker and writer with a long-standing interest in triremes, proposed the Trireme Project with the aim of building a full-scale reconstruction, Morrison and I were convinced that the time had indeed come for such an effort. With funds provided mainly by the Greek national authorities, *Olympias* was completed by a shipbuilder in Piraeus in 1987.

The ship that has emerged from some five years of design, research and development is highly integrated and sophisticated. Clearly the trireme capped an impressively determined and sustained development effort in ancient Greece and possibly other parts of the eastern Mediteranean beginning sometime in the seventh century B.C. It can now be seen that the ancient shipbuilders had reached an optimum design, within the limits of the materials and techniques available at the time, without the benefit of modern-day knowledge of hydrostatics, the stability of ships, structural mechanics, physiology and physics. Indeed, the trireme was almost certainly the fastest oared ship ever built. Its development provides further evi-

JOHN F. COATES has been mainly occupied for the past decade with the design and development of the replica of the ancient Greek trireme. After earning an M.A. in engineering science from the University of Oxford, he worked on ship design, structural research and ship repair for the British Royal Navy and was appointed Chief Naval Architect in 1976. Since his retirement in 1979 he has pursued an interest in applying naval architecture to historical ship research. Coates lives in Bath, England.

***OLYMPIAS*, the reconstructed trireme, cruises in the Saronic Gulf, where 2,500**

dence that the technology of ancient Greece reached a level generally unsurpassed until the latter part of the 18th century A.D.

It cost about $700,000 and took two years to build *Olympias*. In about 482 B.C. Athens, with a population of some 250,000 citizens, foreign residents and slaves, built approximately 200 triremes. This was certainly the main effort in a major arms program undertaken partly to meet the urgent needs of a naval war with the rival city-state of Aegina but mainly in anticipation of a second Persian invasion. It demonstrates a capacity to organize ship production on a level comparable to that achieved by the U.S. shipbuilder Henry J. Kaiser in World War II. It is regrettable, but not surprising in view of the general lack of industrial documentation in ancient Greece, that no details survive of the methods by which this great production rate was achieved.

Ancient literature indicates that in mainland Greece and Sicily only the richer city-states, and in Asia Minor only those subsidized by Persia, could afford triremes. The rest made do with 50-oared penteconters. For the cities that did build the costlier ships, triremes presumably offered a valuable military capability that could not be attained any other way. Theoretical calculations of the main performance characteristics of triremes and their much smaller predecessors indicate that maneuverability—so important in combat for ships armed principally with the ram—was little different from that of the much cheaper penteconter.

The chief advantage of the trireme, therefore, was its greater speed and larger crew. Triremes would have been some 30 percent faster under oar than the penteconter. They could have overtaken any other existing type of ship, which made it possible to adopt an offensive tactic in battle. The larger crews might have been valuable militarily on shore; there is no evidence that oarsmen were ever enlisted to board enemy ships at sea.

Surprisingly, it appears that the trireme developed directly from the two-level penteconter, which was powered by files of 12 and 13 oarsmen on each side. The literature does not indicate that any warship of intermediate size was ever adopted, although there is possibly a hint in the *Iliad* of a ship with 60 oarsmen on each side on two levels. Why, then, was such a large jump in size and cost undertaken, from the penteconter to the trireme? Historians may at some time be able to shed some light on that question.

Excavation more than a century ago of the foundations of trireme sheds, hundreds of which were built at Zea, part of Piraeus, showed that the trireme could not have been more than 5.6 meters wide and about 37 meters long. It was known from inscriptions that the ship carried 170 oarsmen, along with a captain, about a dozen deckhands and officers and 14 soldiers and archers. From references in the literature of the time it can be deduced that the oarsmen were arranged in three files

years ago an Athenian trireme fleet routed a Persian navy. The ship is 36.8 meters long, 5.4 meters wide and 3.6 meters high from keel to canopy deck. Fully loaded, it displaces 45 tons. Eyes painted above the ram serve as a talisman to ward off evil.

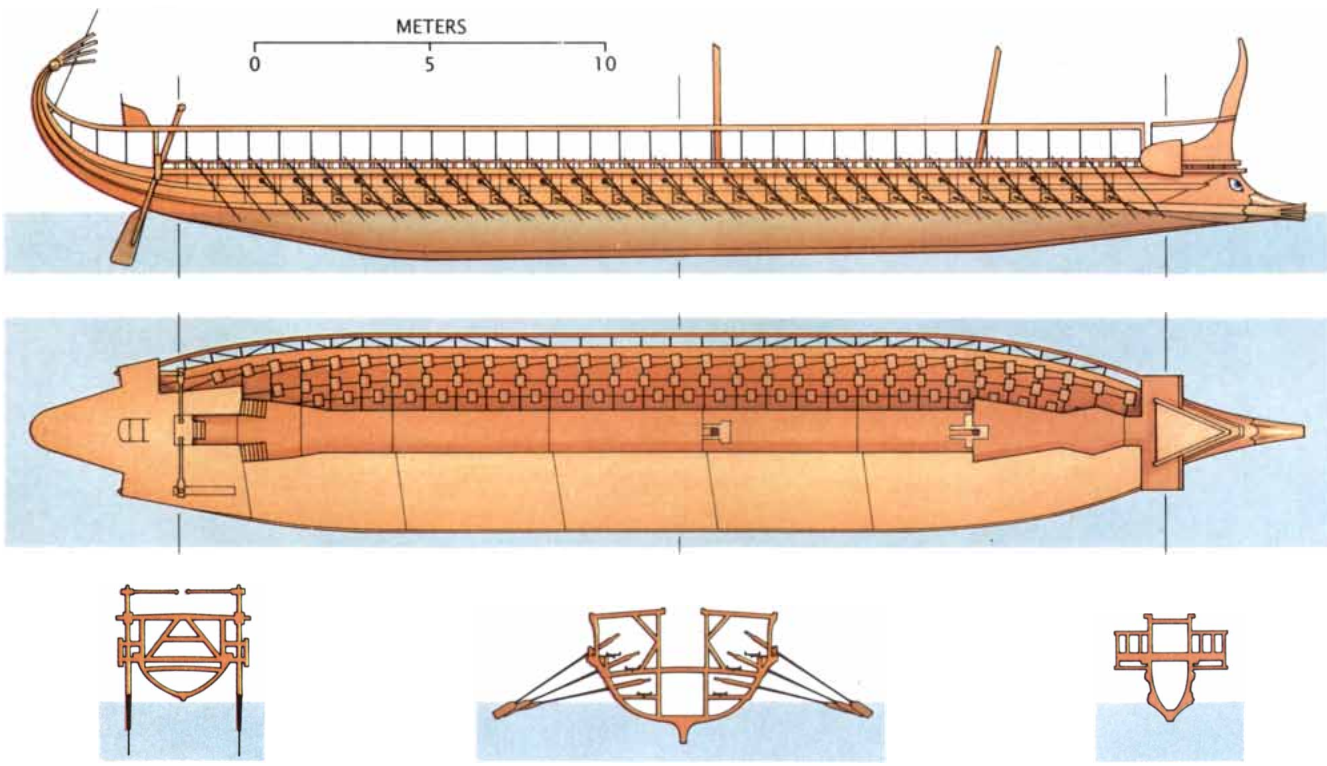
of 31, 27 and 27 oarsmen on each side and that their seats were fixed rather than sliding. What was not clear was how so many rowers could have been fitted into such a confined space. It was also hotly debated by trireme scholars in general whether triremes under oars alone could have attained the sprint speeds, 9.5 knots or more, that are inferred from the sustained speeds attested to in ancient literature, or whether the oarsmen were assisted by sails.

Olympias was in the early stages of design when Vernard Foley and Werner Soedel published the last major article on the subject in this magazine [see "Ancient Oared Warships," by Vernard Foley and Werner Soedel; SCIENTIFIC AMERICAN, April, 1981]. That article established some important points about the trireme. The authors rightly emphasized that the trireme was an extremely light vessel for its size. They estimated its weight to be "less than 40 tons" on a length of 121 feet. Subsequently they reduced that estimate even further. All previously proposed reconstructions on paper had been much more massive, indicating a general failure to recognize the extraordinary refinement of this type of ship. Foley and Soedel also drew attention to the trireme's attested speed and agility, for which lightness would have been essential. They did not, however, proceed numerically into matters of the ship's speed and power.

It was clear that more than the available evidence from literature, inscriptions and art was needed to define the trireme as a complete and workable ship. It was also necessary to determine the form of its hull and structure. These essential features were recently brought to light by underwater archaeologists working in the Mediterranean, who have now established the hull form and structure both of merchant sailing ships and of long, oared ships. All these requirements, together with the demands of physical law, made it possible to infer the form and structure of triremes within close limits—on the order of inches in cross section and a few feet in length.

Because the design was tightly confined, it was also reasonable to hope that some of the questions about triremes might be resolved during the process of carrying out the reconstruction. For example, it had been commonly held that triremes required ballast in order to stay upright because of their height above water, necessitated by the three stacked files of oarsmen. Ancient accounts of sea battles, however, mentioned that the victors towed away the wrecks of the defeated fleet. Also, the Greek word usually translated as "sunk" can also mean "swamped." As Augustin F. B. Creuze observed in the Encyclopædia Britannica in 1841, "the vessels spoken of as sunk were evidently merely stove in and waterlogged." In that condition they must have carried either no ballast or not enough to drag the wood structure to the sea bottom. In addition, ballast would have added mass, undesirable in a fast ship. It was satisfying, therefore, to discover while designing the reconstruction that triremes needed no ballast for stability. One implication of this finding is that the chances of coming on a trireme wreck on the floor of the Mediterranean today are slim and that such a discovery, although not completely improbable, is not worth waiting for.

The weight of a ship equals the weight of the water displaced, according to Archimedes' principle. A critical problem for us, then, was to design the reconstruction so that when it was fully loaded it would float at just the right waterline to allow the oars to be worked properly. It was particularly important to design the hull carefully, because the hull's shape and weight primarily determine the



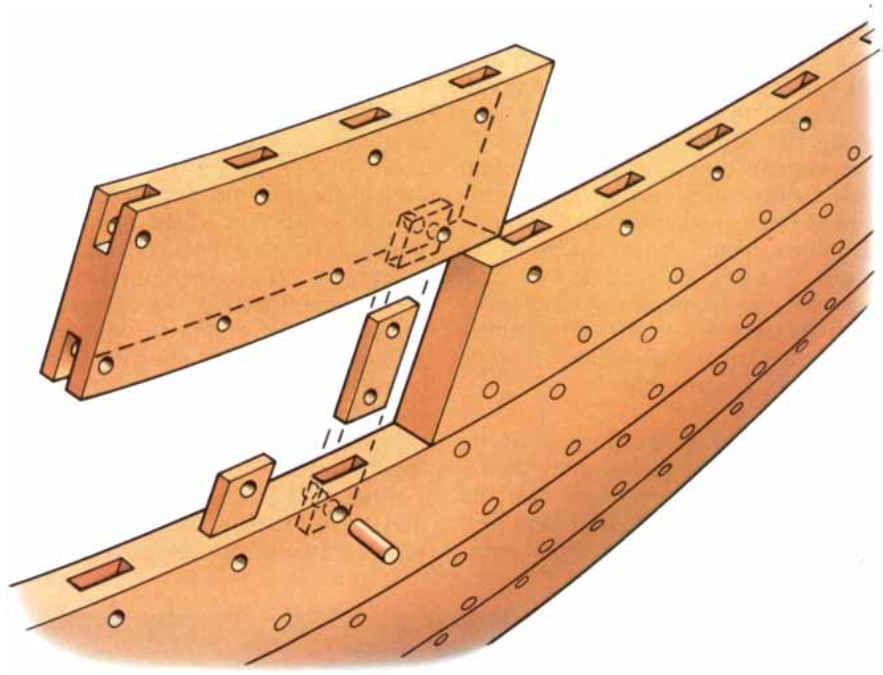
SIDE, TOP AND SECTION VIEWS of *Olympias* are shown in these scale drawings. The 170 rowers were arranged in three files on each side: 31 in the top bank and 27 in each of the two lower banks. The six files are arranged in a V shape, with the lowest bank farthest inboard and the top one farthest outboard. The pivots for the topmost oars were placed on outriggers. The seats of the rowers in the top and bottom files were angled a few degrees inboard, so that all the oar tips were evenly spaced.

ship's displacement and waterline. These features are in turn dictated by the actual methods of construction employed at the time. In this respect Frost's excavation of long, oared ships has been vital to the reconstruction.

All previous proposals for triremes had hulls that were wrong for several reasons: they were too heavy, because they were derived from medieval galleys or later timber ships, and they carried ballast. In what we now call traditional wooden-ship construction, the hull is built by first setting up the framework—keel, stem and sternpost—and then attaching numerous transverse members to it like ribs. The shape so generated is clothed with planks nailed or bolted to the outside of the frame; the inside of the frame is often also covered with planking. The hull is made watertight by hammering fibrous caulking material, usually cotton or hemp, tightly between the planks, which are prevented from sliding on each other only by such frictional forces as may be developed through the caulking. There are also massive longitudinal members inside the frame: a keelson above the keel and "clamps" or "shelves" under the ends of deck beams on each side of the ship. The ship is in essence built from the inside outward.

In the ancient Mediterranean, until the latter half of the first millennium A.D., ships and boats were built from the shell inward by completely different procedures, not all of which have been recovered. Shell-first construction demands that planks be fastened to each other edge to edge, progressively as a new plank is added to the one below it, to form a shell from the keel upward. The planks were fastened by numerous tenons of hardwood, set tightly in individual mortises, or sockets, cut in the edge of the plank. The tenon was locked in place by two pegs, one through each plank at each end of the tenon. The method is laborious but makes a very strong joint, and planks so joined form a true shell in the modern engineering sense: it could sustain sheer stresses in the plane of any of its elements. Various excavations of ancient Mediterranean ships have confirmed the existence of edge-to-edge joinery in shipbuilding.

The planking could be made thinner in this type of construction than in frame-first construction because of its greater strength. Indeed, the framing itself could be lighter too, because the purpose of the ribs was reduced to that of stiffening the shell against local and overall deformation. The ribs could have been inserted, typically in three unconnected tiers, progressively as the shell was built up far enough to receive them. They are not essential for holding the planking together, as they are in the more modern construction (which, although undoubtedly cheaper, is cruder in concept). The trireme was able to hold 170 oarsmen because the method of construction did not require large longitudinal and transverse timbers for the interior.



SHIP HULLS were constructed in the ancient Mediterranean by joining planks edge to edge with tenons inserted into individual mortises cut into the edges of the planks. Each tenon was held in place by two pegs. The hull was built from the keel up: frame timbers were added as the hull was completed far enough to receive them.

That method of construction also turned out to be necessary to produce a hull that is both light in weight and small enough in displacement to balance it at the correct waterline when the trireme was fully loaded. The ancient hull cross section has a wineglass shape, which is necessary for the satisfactory pegging of the tenons joining keel to garboard plank (next to the keel), as well as to form with the keel a flexurally stiff girder along the middle of the ship's bottom. In addition, the area of the wineglass-shaped hull cross section below the waterline is small, reducing the displacement to roughly 40 tons—very small for a ship with a length (at the waterline) of more than 30 meters. In a more modern hull with a flat bottom and firm bilge but of the same draft and breadth, the volume of the immersed hull would be considerably larger, raising the displacement to the vicinity of 70 tons or more.

As we designed the reconstruction it became apparent that the ancient hull section and method of construction fulfilled several other requirements. The system of joints and pegs was needed to carry the shearing loads generated by the variable distribution of weight and buoyancy along the length of such a slender hull. The heavy keel and adjacent planks that formed the "cusp" of the wineglass cross section made the hull deep enough to provide sufficient longitudinal bending strength. The hull section also had enough breadth at the waterline, while preserving low displacement volume, to provide adequate stability. The extension of the keel below the main part of the bottom ensured lateral resistance for reasonable performance under sail. In addition, the hull section had just enough flare to house the oarsmen so that each file was outboard of and higher than the one below it. Finally, with this type of hull section it would have been easy to clean, repair and recoat the outer bottom of the ship on slipways and beaches.

We found that the long, slender hull of the trireme had to sustain tensile bending stresses that approached the limit of what a wood structure could withstand. It had no deck to act as the upper flange of the hull and to help prevent the keel from buckling in the middle and possibly snapping in two.

In triremes long flaxen ropes (*hypozomata*) were stretched tightly fore and aft and then twisted to an even greater tension to reduce tensile stresses along the top edges of the hull. In experiments on a small scale we found that flaxen ropes relaxed under sustained stress and broke in an unpredictable way, and so we decided to use ropes of artificial fiber. Lack of time prevented us from solving some problems that arose with the ropes on the full-scale ship and we had to resort to steel cable temporarily.

Speed was the main virtue of the trireme, and in designing the reconstruction we found that the length and form of the hull were well suited to overcome the various sources of resistance to its progression through the water. Except in high-powered modern craft, the main component of resistance arises from the viscosity of water, or frictional resistance. It is proportional to the wetted area of the hull and its roughness and is also a function of the speed (proportional to the speed raised to the power of 1.825). As explained by Foley and Soedel, the creation of waves by the passage of a ship through water generates a second source of resistance. At low speed the effects of such waves are negligible, but at high speed they account for the major part of the total resistance.

Wave-making resistance varies with hull form, displacement, length and speed. Waves are formed continuously at the bow and stern of a moving ship and create a pattern around the ship as it advances. The waves travel with the ship, those at the bow starting with a crest and those at the stern starting with a trough. As the ship's speed increases, the bow and stern waves alternate between being in phase and being out of phase. Wave-making resistance increases with ship speed, most rapidly when the bow waves are in phase with the stern waves (reinforcing them); under such conditions it takes a very large infusion of power to make the ship move any faster. Conversely, when the bow waves are out of phase with the stern waves, the latter are partially cancelled, the resistance increases less rapidly and it is easier to make the ship move faster.

The ship's length, together with its speed, then, determines the interaction of the bow and stern waves and therefore the wave-making resistance. It is remarkable that the trireme's length is such that its wave-making resistance is increasing only slowly at just about the maximum speed (from 9.5 to 11 knots) at which the ship can be expected to be propelled under oar for a short sprint. To achieve the next slow increase in wave-making resistance, I calculated that it would be necessary to build a ship about 50 percent longer, with a crew of some 250 oarsmen; what is more, the gain in speed would be slight. In any case the endeavor would come to naught for lack of the longitudinal bending strength necessary to keep such a long ship from breaking apart. The trireme was just the right length to achieve the greatest speed in a practicable ship.

HULL VIEW of *Olympias* shows the smooth construction and shallow draft of the ship. Leather sleeves cover the lowest oar ports, which are only .4 meter above the waterline. A projecting outrigger provides the pivots for the topmost file of rowers. In the foreground the author (*middle*) discusses the shape of the ram with colleagues.

The performance of the trireme is, of course, determined not only by its size, form and structure but also by the power available from an oar crew of 170 men, the likely relation of required effective horsepower and ship speed, and the mechanics of oars worked from fixed seats in a fast oared ship. The sprint speed of triremes would have demanded that the oars be of exactly the right mass, balance, blade area and "gearing" (the ratio of the length of the oar outboard of its pivot to the length inboard). In high-performance oars these are critical matters, as every racing rower knows, and the oars of the trireme would not have been an exception. Most hypothetical reconstructions, by neglecting oars, have unwittingly been hard on the oarsman.

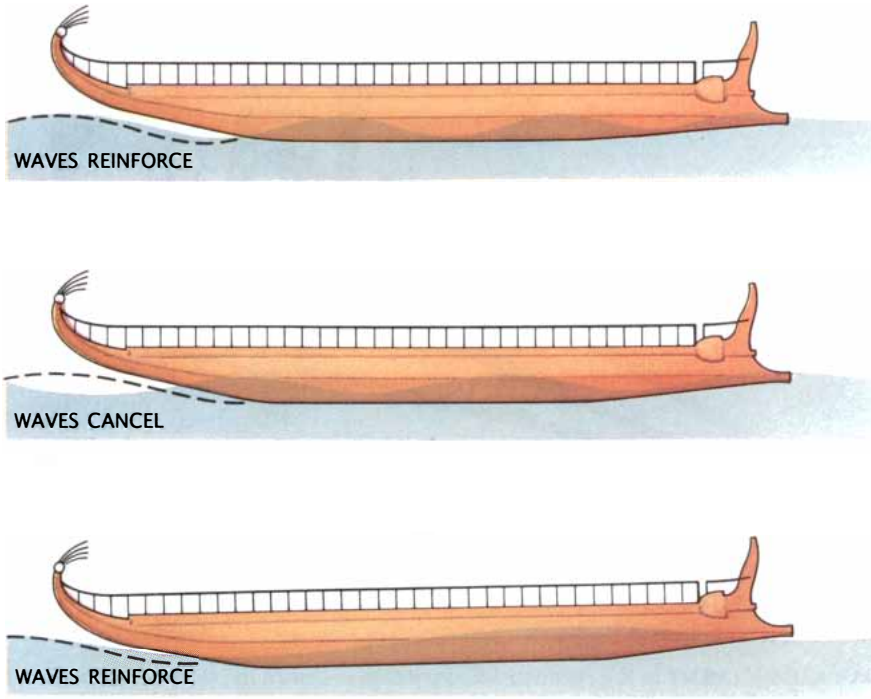
It is known from inventories of the naval dockyards of ancient Athens,

some of which have been found engraved on stone tablets in Piraeus, that the oars of triremes of the third quarter of the fourth century B.C. were nine and 9.5 cubits long (4.0 and 4.22 meters). There was one man to an oar, and the shorter oars were deployed at the ends of the trireme. The Roman engineer and architect Vitruvius recorded in the first century B.C. that the oars were set two cubits (.888 meter) apart. In addition, two ancient depictions are generally accepted as being reliable representations of the oar arrangement. The Lenormant relief, a fifth-century B.C. limestone relief from the Acropolis, shows a partial profile of a trireme. The Ruvo vase, a red-figure Attic krater, depicts the Argonauts on board a ship that Morrison identified as a trireme; it provided information about oar ports for the bottom two tiers of oars and the outrigger through which the top files of oars were worked.

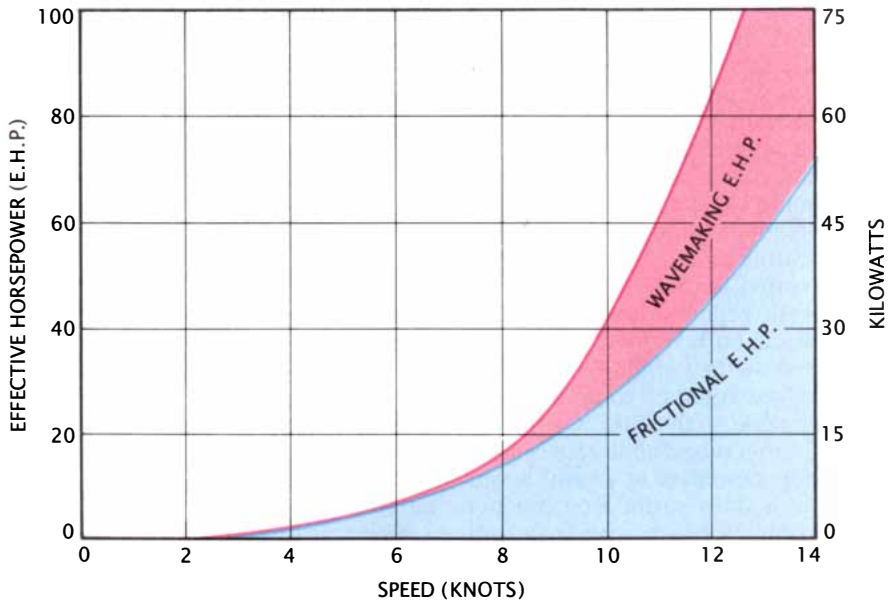
Since the maximum possible length of swing of the oar handles was .85 meter and the stroke rate was at most 50 per minute (the fastest rate at which oars of the stated length can be worked effectively), we deduced that in order to achieve 9.5 knots the gearing of the oars had to be high: about three to one. In addition the width of the hull at the waterline had to be about 3.7 meters in order to provide adequate stability with the fully loaded ship's center of gravity about .8 meter above the waterline. The lowest file of oars had to be at least .4 meter above the water to be workable under ordinary surface-wave conditions. (Even so, the lowest ports were protected by leather sleeves.) We had found out by experiment in a full-size rowing mock-up that the oars in the top file had to be inclined at an angle of less than 35 degrees from the horizontal; at steeper angles the oars could not be worked as effectively.

These oar characteristics have settled the old debate about whether the trireme had to be augmented by sails to achieve its top speed. The answer is that it did not. Except under unusual wind conditions, the sails would have caused the trireme to heel over far enough so that the rowers on the down side would have had the oar handles pushed into their laps and would not have been able to lift the blades over the top of the waves.

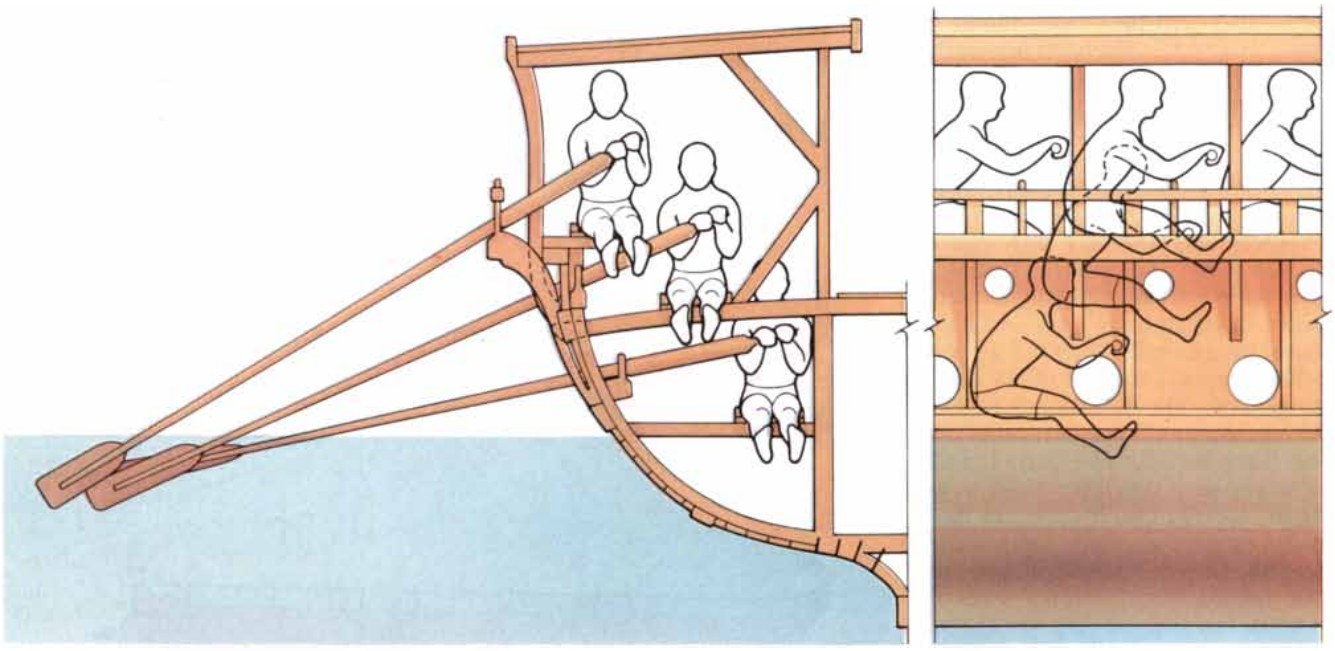
How much effective power could a trireme's crew have delivered to their ship? Data on the performance of oarsmen in fixed seats are scarce. One source is provided by



WAVE-MAKING RESISTANCE arises because standing waves formed at the bow interact with waves formed at the stern (*broken line*). At low speed the effect is slight, but as the ship moves faster the wavelength of the standing wave increases, so that the bow and stern waves are alternately in phase (*top*) or out of phase (*middle*). When the waves are in phase, the resistance increases rapidly; when they are out of phase, it increases less rapidly. Eventually the ship reaches a speed at which half of the length of the standing bow wave equals the ship's length (*bottom*). This critical speed can be increased by lengthening the hull. The wave height is exaggerated here for clarity.



EFFECTIVE HORSEPOWER needed to propel *Olympias* at various speeds was calculated on the basis of ship-model tests made at the National Technical University of Athens. The blue area shows the proportion of total power that is absorbed by frictional drag. The red area indicates the amount of power needed to overcome wave-making resistance. The graph shows that at the speed of seven knots attained by the ship in 1987 the maximum effective power was 10.5 kilowatts, or .062 kilowatt per person. If the crew could match the performance of U.S. Navy light racing cutters (.128 kilowatt per person, or 21.8 kilowatts for the entire crew), it should be possible to drive the trireme at 9.3 knots if the rudders were raised.



OAR ARRANGEMENT in the reconstructed trireme is shown in a frontal cross section (*left*) and in a side view (*right*). The hull needed enough flare to house the oarsmen. The outrigger at the top of the hull placed the top oar pivot .6 meter outboard.

the U.S. Navy's fixed-seat light racing cutters, which many years ago were standardized throughout the service and engaged in competitions over courses of three nautical miles. A model of the standardized cutter was tested for resistance in a water tank. The tests showed that when racing at seven knots over the three-mile course, the 12-man oar crew developed 2.05 effective horsepower, or 1.53 kilowatts; that is, an average of .128 kilowatt per man.

If the rowers in a trireme were all as effective as those in the U.S. Navy cutters, then the 170 rowers in the trireme would be able to generate 21.76 kilowatts for 25 minutes (the duration of the cutter races). That power may be calculated to drive the ship at 9.3 knots if the rudders were raised or at 8.6 knots if both rudders were fully immersed, as they would have to be if the ship was required to be instantly maneuverable. In the 1987 sea trials a crew of rather mixed ability succeeded in driving *Olympias* at nearly seven knots for a short sprint after practicing for only 25 hours, with both rudders immersed.

The maximum effective power per rower was .062 kilowatt, only half that of the average U.S. Navy cutter oarsman. The trireme crew, however, was hampered by being mixed not only in experience and ability but also in leg length and in capacity to adapt to fixed-seat rowing, which requires a large swing of the body to achieve adequate stroke length. (Sliding seats allow a longer drive with the legs and so do not call for such a large body swing.) Moreover, the oars were quite unbalanced, so that each rower needed to expend almost 20 watts just lifting the blades out of the water. During the second series of trials in 1988, after more training, the ship achieved about nine knots. The speed targets for the reconstruction seem therefore to be attainable with additional experience.

After a few trials it became clear that if oarsmen are to deliver power effectively, the oars must have exactly the right design. In a restoration they must also, of course, be made by ancient technology, which excludes the use of waterproof glues and therefore of composite sections. Authors of the time refer to oars made of spruce, and indeed this is a suitable timber, since it is light and yet relatively stiff. At 50 strokes to the minute the kinetic energy of the swinging oar becomes significant, and so the moment of mass of the oar about its pivot needs to be minimized.

With a gearing of three to one it is impossible to achieve an acceptable oar balance without either using timber of the quality of spruce or counterweighting the handles. Counterweighting adds to the mass of the oar and would probably prove unacceptable when rowing fast; also there is no evidence for it in ancient sources, although one would expect to find mention if the practice had been common. Finally, with a three-to-one gearing and an acceptable balance on the handle to lift the blade out of the water with a downward force of no more than 3.63 kilograms (based on normal sporting oars), the buoyancy of the oar blade has to be kept down to little more than 1.2 kilograms so that the blade floats mainly immersed when the hands are off the handle. That calls for blades of small volume, which is difficult to achieve without making them either too small in area or too thin for strength.

Reaching speeds beyond 9.5 knots will depend primarily on the quality of the crew and on arriving at the best design for the oars. The hull must also be kept smooth. Crews will certainly need periods of practice and adjustment before *Olympias* can emulate its ancient originals. The search for the right oar design is still in progress. That such a search is necessary demonstrates the degree to which the trireme was refined to achieve its remarkable performance.

By reconstructing a trireme and demonstrating that the working ship conforms with archaeological evidence, ancient literature and pictorial representations, we have passed the first test of authentication. It remains to be seen whether it is possible to achieve the levels of performance under oar attested to in ancient literature as well as those that can be inferred from modern physics and by experimentation. One target is to sustain 7.5 knots continuously for a

day, thereby matching the legendary 24-hour voyage from Athens to Mytilene. The other speed target is 9.5 knots for three nautical miles, a speed inferred from the data on U.S. Navy racing cutters.

Olympias has amply and gratifyingly demonstrated its maneuverability in the Aegean over the past two years. A test that has not been carried out very effectively yet is the important maneuver of backing water, necessary not only following the ramming of an enemy ship but in many other circumstances in the melee of battle. One objective is to measure and translate trireme maneuverability into terms that will provide inputs for a computer simulation; this would put the study of trireme tactics on a quantifiable basis. The performance of *Olympias* under sail also has yet to be evaluated thoroughly, but it is already clear that it sails much better than the simple square rig would suggest to modern eyes. A study of sailing performance would tell something about how fleets were deployed in the ancient Mediterranean.

In view of the archetypal position of the trireme in influencing the development of larger oared warships, the knowledge gained from the trireme reconstruction should provide a basis for designing and possibly reconstructing the most important of the later ships, the quinquereme (a three-tiered ship with two men on each oar in the two upper files). Ultimately we hope that by building and carrying out experiments with *Olympias*—and possibly, in due course, with similar reconstructed ships—we can reach a clearer and more practical understanding of the exercise of sea power in the Mediterranean, which so vitally determined the course of history during the classical period of Greece.

FURTHER READING

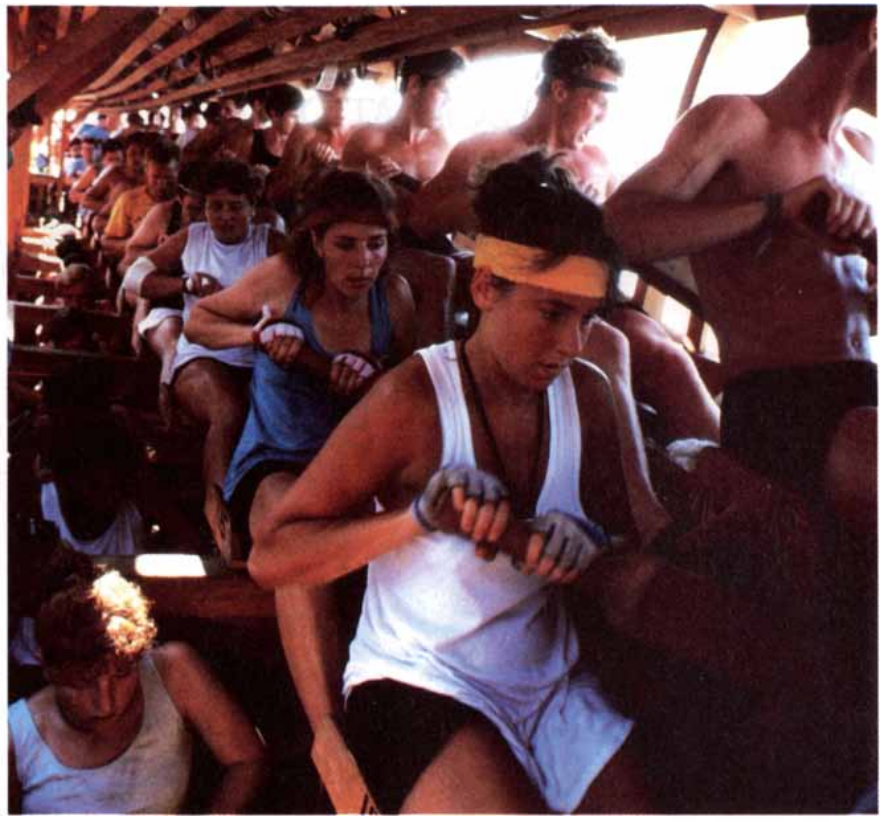
GREEK OARED SHIPS, 900-322 B.C. J. S. Morrison and R. T. Williams. Cambridge University Press, 1968.

SHIPS AND SEAMANSHIP IN THE ANCIENT WORLD. Lionel Casson. Princeton University Press, 1971.

FIRST SEASON OF EXCAVATION ON THE PUNIC WRECK IN SICILY. Honor Frost in *International Journal of Nautical Archaeology and Undersea Exploration,* Vol. 2, No. 1, pages 33-49: March and September, 1973.

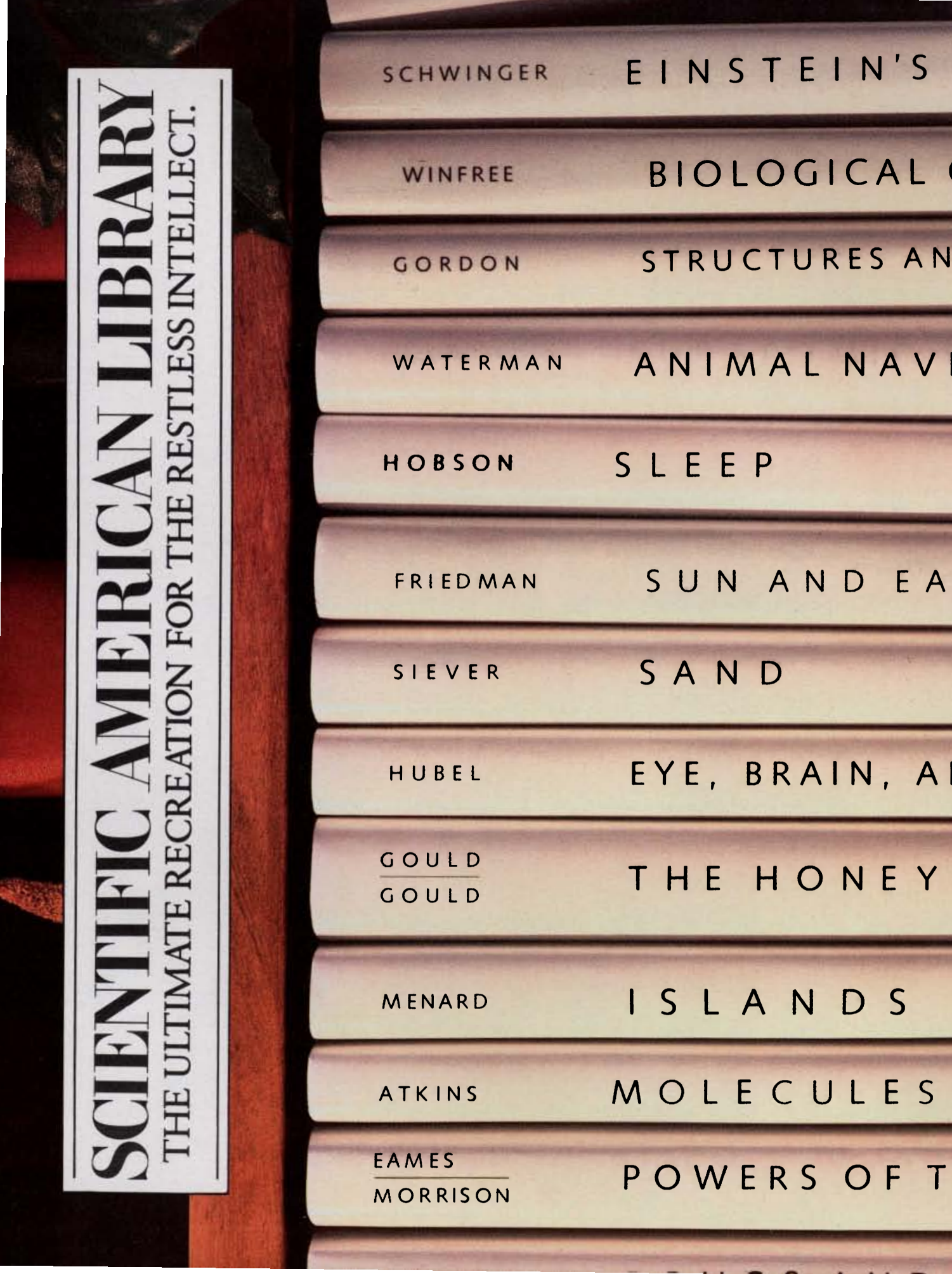
ARCHAEOLOGY BENEATH THE SEA. George F. Bass. Walker & Co., 1975.

THE ATHENIAN TRIREME: THE HISTORY AND RECONSTRUCTION OF AN ANCIENT GREEK WARSHIP. John S. Morrison and John F. Coates. Cambridge University Press, 1986.



MIXED CREW of volunteers is arranged in crowded oar banks on board *Olympias*. They sat on fixed, leather-covered seats and pulled on oars made of Oregon pine.

ALL THREE TIERS of rowers attack the water with their oars as they drive *Olympias* into a sprint. In sea trials the ship proved itself capable of more than nine knots under oar or sail. Coaches, recorders and assistants stand on the canopy deck.


SCIENTIFIC AMERICAN LIBRARY
THE ULTIMATE RECREATION FOR THE RESTLESS INTELLECT.
SCHWINGER
EINSTEIN'S
WINFREE
BIOLOGICAL
GORDON
STRUCTURES AN
WATERMAN
ANIMAL NAV
HOBSON
SLEEP
FRIEDMAN
SUN AND EA
SIEVER
SAND
HUBEL
EYE, BRAIN, A
GOULD
GOULD
THE HONEY
MENARD
ISLANDS
ATKINS
MOLECULES
EAMES
MORRISON
POWERS OF T

Join the most exciting book club in science—where the only commitment you make is to reward your own curiosity

This is the Scientific American Library, where the mysteries of the human and natural worlds unfold before your eyes in lavishly illustrated volumes you'll want to read and refer to for years to come.

Here's why 75,000 readers like you have already become members of the Scientific American Library

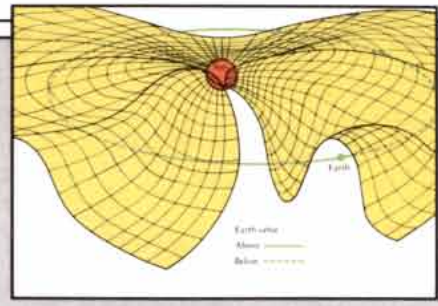
■ **15-day Free Trial**
You will receive a new volume to examine approximately every two months. (Your first volume is shipped immediately via UPS at no extra charge.) You have 15 days to look over each volume. If you decide to keep it, send in your payment. Otherwise, simply return the book in the same shipping carton and we'll immediately credit your account.

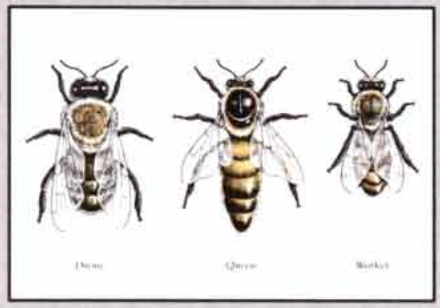
■ **No Prepayment and No Minimum Purchase**
Send no money in advance. You pay only for the volumes you keep. You may choose as many or as few volumes as you like: there is no yearly minimum.

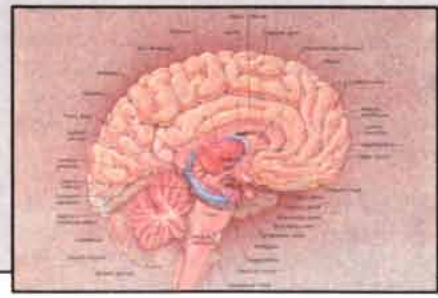
■ **Low Members' Price**
You may purchase any of these handsome volumes for only $24.95—25% below the bookstore price.

■ **No Strings, No Risk, No Commitment**
You may cancel your obligation-free membership at any time for any reason, without notice, without penalty.

CREDITS (top to bottom)—From Sun and Earth, *the "ballerina skirt" model of the solar-interplanetary current sheet. From* The Science of Structures and Materials, *bubble formation due to the surface tension of a liquid; courtesy Stephen Dalton, NHPA. From* The Honey Bee, *drone, queen, and worker; drawing by John Hatzakis. From* Einstein's Legacy, *Chaplin and Einstein in Hollywood, 1931; courtesy The Granger Collection. From* Drugs and The Brain, *major anatomical divisions of the brain; painting by Bill Andrews.*







A one-of-a-kind book club for those who take science seriously.

Each volume in the Scientific American Library was created exclusively for this series, written by preeminent scientists—including many Nobel laureates—who describe their work in an exciting, accessible way.

Choose your premier volume

■ **Extinction,** Steven M. Stanley. In this geologic detective story, a renowned paleobiologist explores the causes of the mass extinctions that have periodically decimated thousands of thriving species. "Smack up to date scientifically...the best overview of the changing relationship between the physical environment and life.... Enthralling...."
—John Gribbon, *New Scientist*

■ **Molecules,** P. W. Atkins. A distinguished chemist takes us below the surface of visible reality to show how molecules make soap "soapy," determine the taste of barbecue, and give rise to the brilliant colors of fall foliage. "Undoubtedly the most beautiful chemistry book ever written.... It is more than a book of facts; *Molecules* is a work of art."
—John Emsley, *New Scientist*

Join us. You have nothing to lose. And a world of scientific discovery awaits.

To become a member of the Scientific American Library, simply detach and mail the facing reply card.
Or write to: Scientific American Library, P.O. Box 646, Holmes, PA 19043

Children of the Garden Island

In 1955, 698 infants on the Hawaiian island of Kauai became participants in a 30-year study that has shown how some individuals triumph over physical disadvantages and deprived childhoods

by Emmy E. Werner

Kauai, the Garden Island, lies at the northwest end of the Hawaiian chain, 100 miles and a half-hour flight from Honolulu. Its 555 square miles encompass mountains, cliffs, canyons, rain forests and sandy beaches washed by pounding surf. The first Polynesians who crossed the Pacific to settle there in the eighth century were charmed by its beauty, as were the generations of sojourners who visited there after Captain James Cook "discovered" the island in 1778.

The 45,000 inhabitants of Kauai are for the most part descendants of immigrants from Southeast Asia and Europe who came to the island to work on the sugar plantations with the hope of finding a better life for their children. Thanks to the islanders' unique spirit of cooperation, my colleagues Jessie M. Bierman and Fern E. French of the University of California at Berkeley, Ruth S. Smith, a clinical psychologist on Kauai, and I have been able to carry out a longitudinal study on Kauai that has lasted for more than three decades. The study has had two principal goals: to assess the long-term consequences of prenatal and perinatal stress and to document the effects of adverse early rearing conditions on children's physical, cognitive and psychosocial development.

The Kauai Longitudinal Study began at a time when the systematic examination of the development of children exposed to biological and psychosocial risk factors was still a bit of a rarity. Investigators attempted to reconstruct the events that led to physical or psychological problems by studying the history of individuals in whom such problems had already surfaced. This retrospective approach can create the impression that the outcome is inevitable, since it takes into account only the "casualties," not the "survivors." We hoped to avoid that impression by monitoring the development of all the children born in a given period in an entire community.

We began our study in 1954 with an assessment of the reproductive histories of all the women in the community. Altogether 2,203 pregnancies were reported by the women of Kauai in 1954, 1955 and 1956; there were 240 fetal deaths and 1,963 live births. We chose to study the cohort of 698 infants born on Kauai in 1955, and we followed the development of these individuals at one, two, 10, 18 and 31 or 32 years of age. The majority of the individuals in the birth cohort—422 in all—were born without complications, following uneventful pregnancies, and grew up in supportive environments.

But as our study progressed we began to take a special interest in certain "high risk" children who, in spite of exposure to reproductive stress, discordant and impoverished home lives and uneducated, alcoholic or mentally disturbed parents, went on to develop healthy personalities, stable careers and strong interpersonal relations. We decided to try to identify the protective factors that contributed to the resilience of these children.

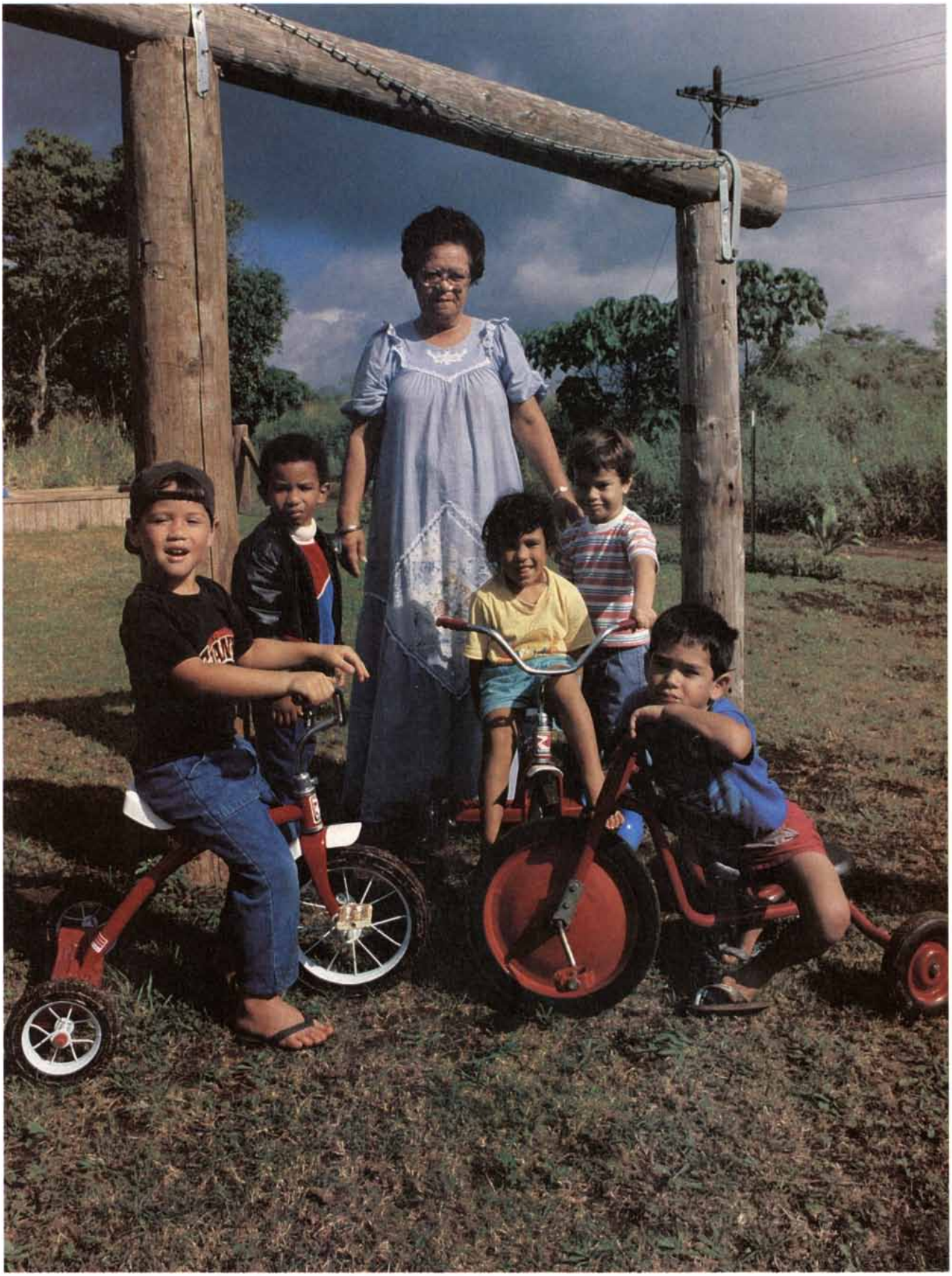
Finding a community that is willing or able to cooperate in such an effort is not an easy task. We chose Kauai for a number of reasons, not the least of which was the receptivity of the island population to our endeavors. Coverage by medical, public-health, educational and social services on the island was comparable to what one would find in communities of similar size on the U.S. mainland at that time. Furthermore, our study would take into account a variety of cultural influences on childbearing and child rearing, since the population of Kauai includes individuals of Japanese, Pilipino, Portuguese, Chinese, Korean and northern European as well as of Hawaiian descent.

We also thought the population's low mobility would make it easier to keep track of the study's participants and their families. The promise of a stable sample proved to be justified. At the time of the two-year follow-up, 96 percent of the living children were still on Kauai and available for study. We were able to find 90 percent of the children who were still alive for the 10-year follow-up, and for the 18-year follow-up we found 88 percent of the cohort.

In order to elicit the cooperation of the island's residents, we needed to get to know them and to introduce our study as well. In doing so we relied on the skills of a number of dedicated professionals from the University of California's Berkeley and Davis campuses, from the University of Hawaii and from the island of Kauai itself. At the beginning of the study five nurses and one social worker, all residents of Kauai, took a census of all households on the island, listing the occupants of each dwelling and recording demographic information, including a reproductive history of all women 12 years old or older. The interviewers asked the women if they were pregnant; if a woman was not, a card with a postage-free envelope was left with the request that she mail it to the Kauai Department of Health as soon as she thought she was pregnant.

Local physicians were asked to submit a monthly list of the women who were coming to them for prenatal care. Community organizers spoke to wom-

EMMY E. WERNER is professor of human development and research child psychologist at the University of California, Davis. She received her Ph.D. from the University of Nebraska in 1955 and then joined the Institute of Child Development at the University of Minnesota; she served as visiting scientist at the perinatal research branch of the National Institutes of Health from 1959 to 1962 and as an associate research child psychologist in the School of Public Health at the University of California, Berkeley, between 1965 and 1969. Werner has taught at Davis since 1962.



SCHOOLCHILDREN ON KAUAI represent an ethnic mixture typical of the island's inhabitants. In addition to the Hawaiians of Polynesian descent, there are people of Japanese, Pilipino, Portuguese, Chinese, Korean and northern European descent.

en's groups, church gatherings, the county medical society and community leaders. The visits by the census takers were backed up with letters, and milk cartons were delivered with a printed message urging mothers to cooperate. We advertised in newspapers, organized radio talks, gave slide shows and distributed posters.

Public-health nurses interviewed the pregnant women who joined our study in each trimester of pregnancy, noting any exposure to physical or emotional trauma. Physicians monitored any complications during the prenatal period, labor, delivery and the neonatal period. Nurses and social workers interviewed the mothers in the postpartum period and when the children were one and 10 years old; the interactions between parents and offspring in the home were also observed. Pediatricians and psychologists independently examined the children at two and 10 years of age, assessing their physical, intellectual and social development and noting any handicaps or behavior problems. Teachers evaluated the children's academic progress and their behavior in the classroom.

From the outset of the study we recorded information about the material, intellectual and emotional aspects of the family environment, including stressful life events that resulted in discord or disruption of the family unit. With the parents' permission we also were given access to the records of public-health, educational and social-service agencies and to the files of the local police and the family court. My collaborators and I also administered a wide range of aptitude, achievement and personality tests in the elementary grades and in high school. Last but not least, we gained the perspectives of the young people themselves by interviewing them at the age of 18 and then again when they were in their early 30's.

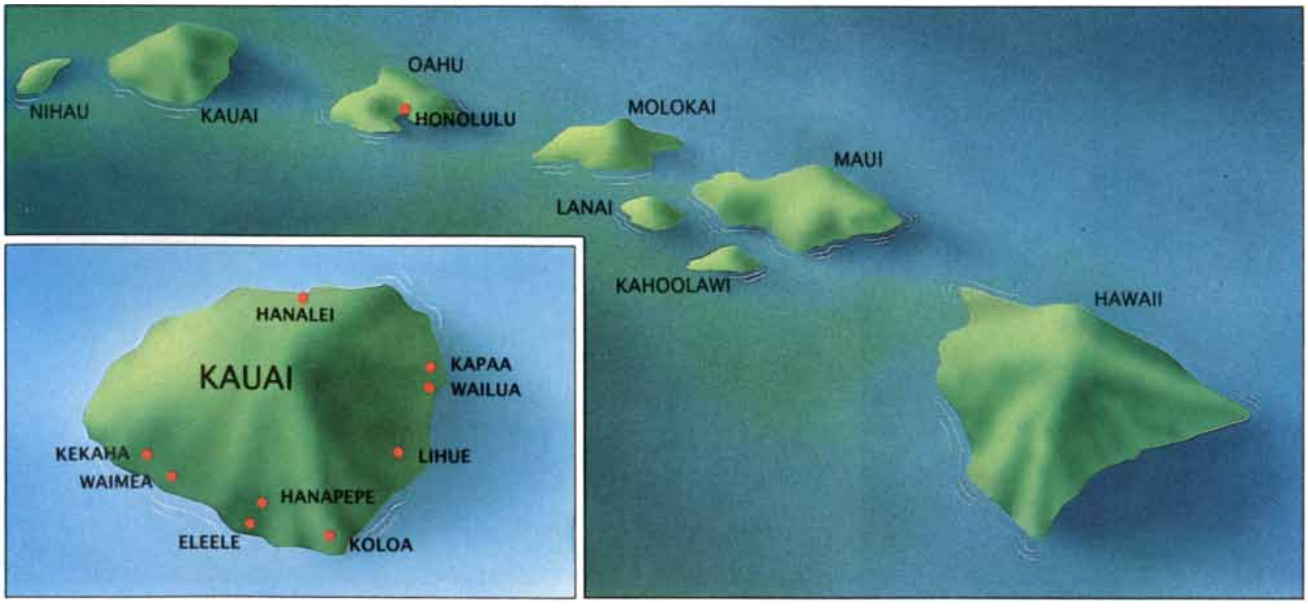
Of the 698 children in the 1955 cohort, 69 were exposed to moderate prenatal or perinatal stress, that is, complications during pregnancy, labor or delivery. About 3 percent of the cohort—23 individuals in all—suffered severe prenatal or perinatal stress; only 14 infants in this group lived to the age of two. Indeed, nine of the 12 children in our study who died before reaching two years of age had suffered severe perinatal complications.

Some of the surviving children became "casualties" of a kind in the next two decades of life. One out of every six children (116 children in all) had physical or intellectual handicaps of perinatal or neonatal origin that were diagnosed between birth and the age of two and that required long-term specialized medical, educational or custodial care. About one out of every five children (142 in all) developed serious learning or behavior problems in the first decade of life that required more than six months of remedial work. By the time the children were 10 years old, twice as many children needed some form of mental-health service or remedial education (usually for problems associated with reading) as were in need of medical care.

By the age of 18, 15 percent of the young people had delinquency records and 10 percent had mental-health problems requiring either in- or outpatient care. There was some overlap among these groups. By the time they were 10, all 25 of the children with long-term mental-health problems had learning problems as well. Of the 70 children who had mental-health problems at 18, 15 also had a record of repeated delinquencies.

As we followed these children from birth to the age of 18 we noted two trends: the impact of reproductive stress diminished with time, and the developmental outcome of virtually every biological risk condition was dependent on the quality of the rearing environment. We did find some correlation between moderate to severe degrees of perinatal trauma and major physical handicaps of the central nervous system and of the musculoskeletal and sensory systems; perinatal trauma was also correlated with mental retardation, serious learning disabilities and chronic mental-health problems such as schizophrenia that arose in late adolescence and young adulthood.

But overall rearing conditions were more powerful determinants of outcome than perinatal trauma. The better the quality of the home environment was, the more competence the children displayed. This could already be seen when the children were just two years old: toddlers who had experienced severe perinatal stress but lived in middle-class homes or in sta-



KAUAI, the Garden Island, lies at the northwest end of the Hawaiian archipelago. The towns that participated in the Kauai Longitudinal Study are shown in the inset. Lihue is the county seat; it is about 100 miles from Honolulu, the capital of Hawaii.

This isn't just a baggage tag. It's a contract.

We don't think a system of handling baggage should leave anything to chance. Which is why we take so many precautions. Like writing baggage tags by computer and not by hand (handwriting, unfortunately, is often subject to interpretation).

Our system works so well that we guarantee you *and* your baggage a smooth trip. And we back that guarantee with cash. If you or your baggage miss a connecting Lufthansa flight when you fly Lufthansa First or Business Class across the Atlantic, we will pay you $200.

Call it a contract. We do. Everywhere we fly. On 6 continents, in 82 countries, in 161 cities.

People expect the world of us.

Lufthansa
German Airlines

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

SOME OF THE WORST THINGS IN LIFE ARE FREE, TOO.

Parkinson's disease has no known cause and no known cure. It affects 1.5 million Americans. And like all diseases, it doesn't cost a thing — until it strikes.

It struck Diane Ruby, R.N., at age 38. Fortunately, through care and treatment by the National Parkinson Foundation, her condition has been kept in check and she's resumed her nursing career.

Research is always under way at the NPF. Because even though victims of neurological disorders can learn to cope, they must still endure the unpredictable course of Parkinson's.

That's why the struggle for a cure for Parkinson and allied disorders goes on. And why every bit of progress leads to a victory. You can help with your tax-deductible gift. For information, contact the NPF today.

NATIONAL PARKINSON FOUNDATION

1501 N.W. 9th Ave. (Bob Hope Rd.)
Miami, Florida 33136-9990
1-800-327-4545; in Fla. 1-800-433-7022

Prepared as a public service by Newman, Blitz & McConnell, Inc., Miami, Florida

Expert to Expert

Save 20% on new titles from
Computer Science Press
The Expert-to-Expert Publisher

A LOGICAL LANGUAGE FOR DATA AND KNOWLEDGE BASES

Shamim A. Naqvi and Shalom Tsur, *MCC Systems Technology Laboratory, Austin, Texas*

Providing material available in no other book, this unique new work on Logical Design Language examines language features and theoretical background required to work in the area of logic and deductive databases. Based on LDL software developed at MCC Systems Technology Laboratory, the book covers the fundamentals necessary to write simple programs as well as the most advanced and detailed material available on the subject. Its language-centered presentation describes successive containing subsets of the language, each more powerful than its predecessor, to build a complete, in-depth framework for the reader.
1989, 304 pages, $39.95
SPECIAL PRICE $31.99

PRINCIPLES OF DATABASE AND KNOWLEDGE-BASE SYSTEMS Volumes I and II

Jeffrey D. Ullman, *Stanford University*

This two-volume set integrates classical database concepts and the new trends in knowledge systems and object-oriented systems. Volume I covers data models and their commercial languages, as well as design theory, physical storage, and security. Volume II covers old and new optimization techniques for efficient implementation of query languages, including the more expressive logic languages of knowledge-base systems.
Volume I: 1988, 631 pages, $41.95
SPECIAL PRICE $33.59
Volume II: 1989, 528 pages, $41.95
SPECIAL PRICE $33.59

DATA COMPRESSION: Methods and Theory

James A. Storer, *Brandeis University*

Data compression is rapidly becoming an essential component of communications and data storage software and hardware. This book provides advanced coverage of data compression techniques in digital text by combining the traditional tools of information theory with those of modern computer science. The focus is on lossless data compression, where the decompressed data must be identical to the original. However, applications to lossy data compression (e.g., where a decompressed digitalized image need only "look" as good as the original) are also discussed.
1988, 413 pages, $44.95 **SPECIAL PRICE $35.99**

Computer Science Press
An imprint of W. H. Freeman and Company, the book publishing arm of Scientific American

THE TURING OMNIBUS 61 Excursions in Computer Science

A.K. Dewdney, *University of Western Ontario*

From artificial intelligence and simulation to computer vision and cryptography, A.K. Dewdney takes you on a fascinating tour through the landmarks of computer science, offering dozens of intriguing glimpses into the current state of the computer scientist's art. *The Turing Omnibus* is a portable reference you'll use and refer to for years to come.
Available for shipment April 21, 1989.
1989, 400 pages, $24.95 **SPECIAL PRICE $19.99**

THE ELEMENTS OF ARTIFICIAL INTELLIGENCE: An Introduction Using LISP

Steven L. Tanimoto, *University of Washington*

"Clearly the best AI book available, by far. It gives lucid, cogent, coherent developments of important topics. LISP and LISP programs are beautifully incorporated into the text, so that LISP and AI can be learned together, each helping and illuminating the other."
—Leonard Uhr, University of Wisconsin, Madison
1987, 530 pages, $37.95 **SPECIAL PRICE $30.59**

Your satisfaction matters. Return any book in good condition within 15 days of receipt and we'll send you a prompt refund. **Offer expires June 30, 1989.**

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

All orders must be prepaid in U.S. dollars only.

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

Acct no._____

Signature_____
(Credit card orders must be signed.)

Name_____
(Please Print)

Address_____

City/State/Zip_____

Qty.	Author	ISBN	Special Price	Total
	Naqvi/Tsur	8200-6	$31.99	
	Dewdney	8154-9	$19.99	
	Ullman, I	8158-1	$33.59	
	Ullman, II	8162-X	$33.59	
	Storer	8156-5	$35.99	
	Tanimoto	8028-3	$30.59	

Postage & handling* (Add $1.95 for first book, $1.25 for each additional book) $_____
NY, MD, CA and UT residents add sales tax $_____
TOTAL $_____

*For Canadian residents, postage & handling is $2.25 for first book, $1.50 each additional book. Please allow four weeks for delivery. Items ordered together may be shipped separately.

0318

ble family settings did nearly as well on developmental tests of sensorymotor and verbal skills as toddlers who had experienced no such stress.

Prenatal and perinatal complications were consistently related to impairment of physical and psychological development at the ages of 10 and 18 only when they were combined with chronic poverty, family discord, parental mental illness or other persistently poor rearing conditions. Children who were raised in middle-class homes, in a stable family environment and by a mother who had finished high school showed few if any lasting effects of reproductive stress later in their lives.

How many children could count on such a favorable environment? A sizable minority could not. We designated 201 individuals—30 percent of the surviving children in this study population—as being high-risk children because they had experienced moderate to severe perinatal stress, grew up in chronic poverty, were reared by parents with no more than eight grades of formal education or lived in a family environment troubled by discord, divorce, parental alcoholism or mental illness. We termed the children "vulnerable" if they encountered four or more such risk factors before their second birthday. And indeed, two-thirds of these children (129 in all) did develop serious learning or behavior problems by the age of 10 or had deliquency records, mental-health problems or pregnancies by the time they were 18.

Yet one out of three of these high-risk children—72 individuals altogether—grew into competent young adults who loved well, worked well and played well. None developed serious learning or behavior problems in childhood or adolescence. As far as we could tell from interviews and from their record in the community, they succeeded in school, managed home and social life well and set realistic educational and vocational goals and expectations for themselves when they finished high school. By the end of their second decade of life they had developed into competent, confident and caring people who expressed a strong desire to take advantage of whatever opportunity came their way to improve themselves.

They were children such as Michael, a boy for whom the odds on paper did not seem very promising. The son of teen-age parents, Michael was born prematurely, weighing four pounds five ounces. He spent his first three weeks of life in a hospital, separated from his mother. Immediately after his birth his father was sent with the U.S. Army to Southeast Asia, where he remained for two years. By the time Michael was eight years old he had three siblings and his parents were divorced. His mother had deserted the family and had no further contact with her children. His father raised Michael and his siblings with the help of their aging grandparents.

Then there was Mary, born after 20 hours of labor to an overweight mother who had experienced several miscarriages before that pregnancy. Her father was an unskilled farm laborer with four years of formal education. Between Mary's fifth and 10th birthdays her mother was hospitalized several times for repeated bouts of mental illness, after having inflicted both physical and emotional abuse on her daughter.

Surprisingly, by the age of 18 both Michael and Mary were individuals with high self-esteem and sound values who cared about others and were liked by their peers. They were successful in school and looked forward to the future. We looked back at the lives of these two youngsters and the 70 other resilient individuals who had triumphed over their circumstances and compared their behavioral characteristics and the features of their environment with those of the other high-risk youths who developed serious and persistent problems in childhood and adolescence.

We identified a number of protective factors in the families, outside the family circle and within the resilient children themselves that enabled them to resist stress. Some sources of resilience seem to be constitutional: resilient children such as Mary and Michael tend to have characteristics of temperament that elicit positive responses from family members and strangers alike. We noted these same qualities in adulthood. They include a fairly high activity level, a low degree of excitability and distress and a high degree of sociability. Even as infants the resilient individuals were described by their parents as "active," "affectionate," "cuddly," "easygoing" and "even-tempered." They had no eating or sleeping habits that were distressing to those who took care of them.

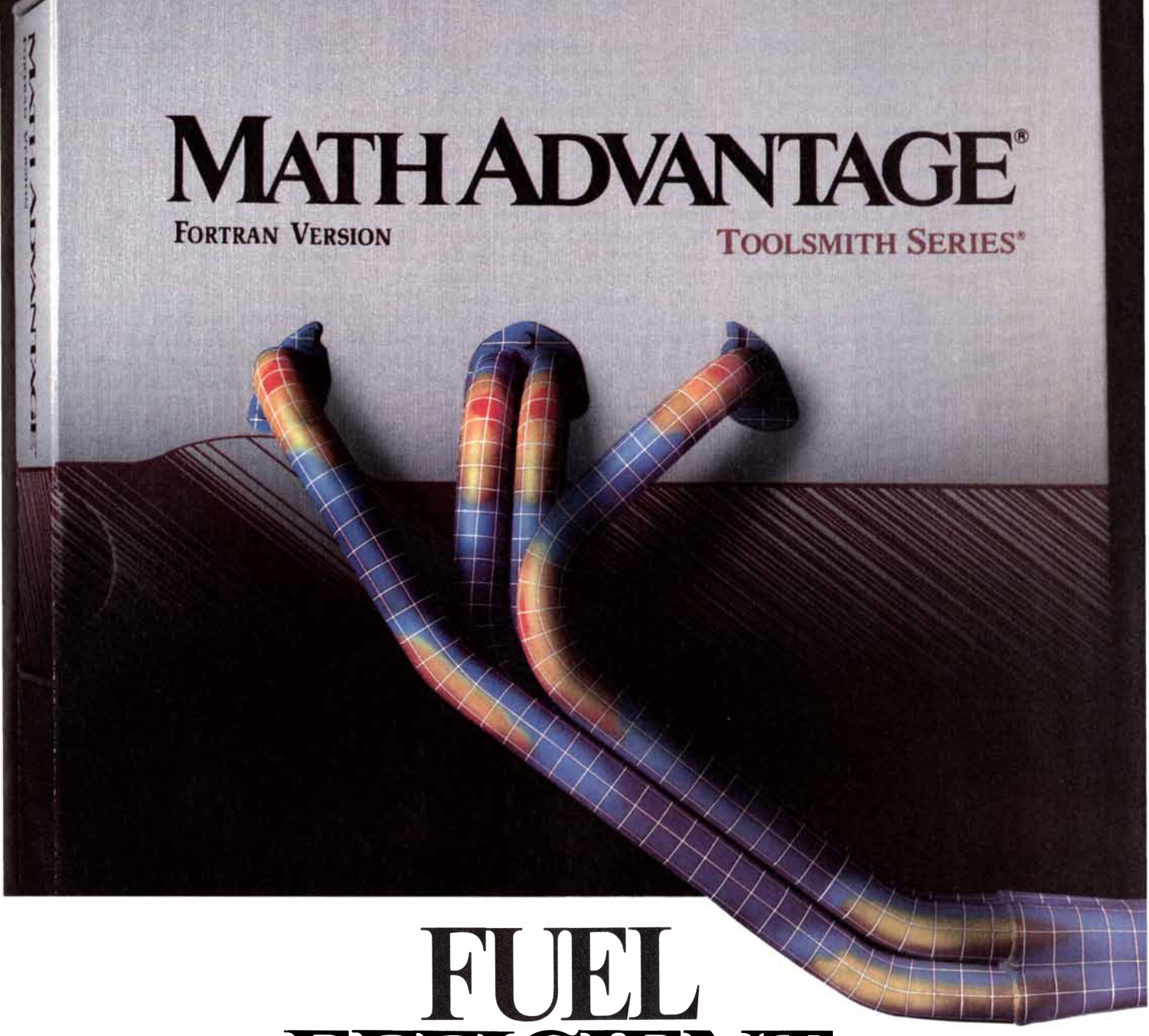
The pediatricians and psychologists who examined the resilient children at 20 months noted their alertness and responsiveness, their vigorous play and their tendency to seek out novel experiences and to ask for help when they needed it. When they entered elementary school, their classroom teachers observed their ability to concentrate on their assignments and noted their problem-solving and reading skills. Although they were not particularly gifted, these children used whatever talents they had effectively. Usually they had a special hobby they could share with a friend. These interests were not narrowly sex-typed; we found that girls and boys alike excelled at such activities as fishing, swimming, horseback riding and hula dancing.

We could also identify environmental factors that contributed to these children's ability to withstand stress. The resilient youngsters tended to come from families having four or fewer children, with a space of two years or more between themselves and the next sibling. In spite of poverty, family discord or parental mental illness, they had the opportunity to establish a close bond with at least one caretaker from whom they received positive attention during the first years of life.

The nurturing might come from substitute parents within the family (such as grandparents, older siblings, aunts or uncles) or from the ranks of regular baby-sitters. As the resilient children grew older they seemed to be particularly adept at recruiting such surrogate parents when a biological parent was unavailable (as in the case of an absent father) or incapacitated (as in the case of a mentally ill mother who was frequently hospitalized).

Maternal employment and the need to take care of younger siblings apparently contributed to the pronounced autonomy and sense of responsibility noted among the resilient girls, particularly in households where the father had died or was permanently absent because of desertion or divorce. Resilient boys, on the other hand, were often firstborn sons who did not have to share their parents' attention with many additional children in the household. They also had some male in the family who could serve as a role model (if not the father, then a grandfather or an uncle). Structure and rules in the household and assigned chores were part of the daily routine for these boys during childhood and adolescence.

Resilient children also seemed to find a great deal of emotional support outside their immediate family. They tended to be well liked by their classmates and had at least one close friend, and usually several. They relied



FUEL EFFICIENT.

Get maximum compute-engine performance for every application, with Math Advantage® from QTC.

Math Advantage is the de facto standard for math libraries in high-speed computing environments. It's developed and optimized by QTC's team of mathematicians, applications engineers, and supercomputer experts. It features over 260 frequently used but hard-to-program subroutines. And it runs on over 40 systems ranging from Cray X-MP to IBM PC, Cyber to VAX, Sun to Mac.

Best of all, Math Advantage gives you FORTRAN, C, and Ada versions of algorithms like FFTs, eigensolvers, matrix operations, and BLAS, all completely *ready-to-run*. So it helps you optimize the most valuable resource of all—your time.

Because it's from QTC, the experts in high-performance development tools, Math Advantage also comes with unmatched documentation, support, and technical excellence.

See for yourself—over 10,000 Math Advantage users already have. Contact QTC at 1-800-234-0168, or Telex 910 240 2827 for more information.

QTC
Quantitative Technology Corporation

on an informal network of neighbors, peers and elders for counsel and support in times of crisis and transition. They seem to have made school a home away from home, a refuge from a disordered household. When we interviewed them at 18, many resilient youths mentioned a favorite teacher who had become a role model, friend and confidant and was particularly supportive at times when their own family was beset by discord or threatened with dissolution.

For others, emotional support came from a church group, a youth leader in the YMCA or YWCA or a favorite minister. Participation in extracurricular activities—such as 4-H, the school band or a cheerleading team, which allowed them to be part of a cooperative enterprise—was also an important source of emotional support for those children who succeeded against the odds.

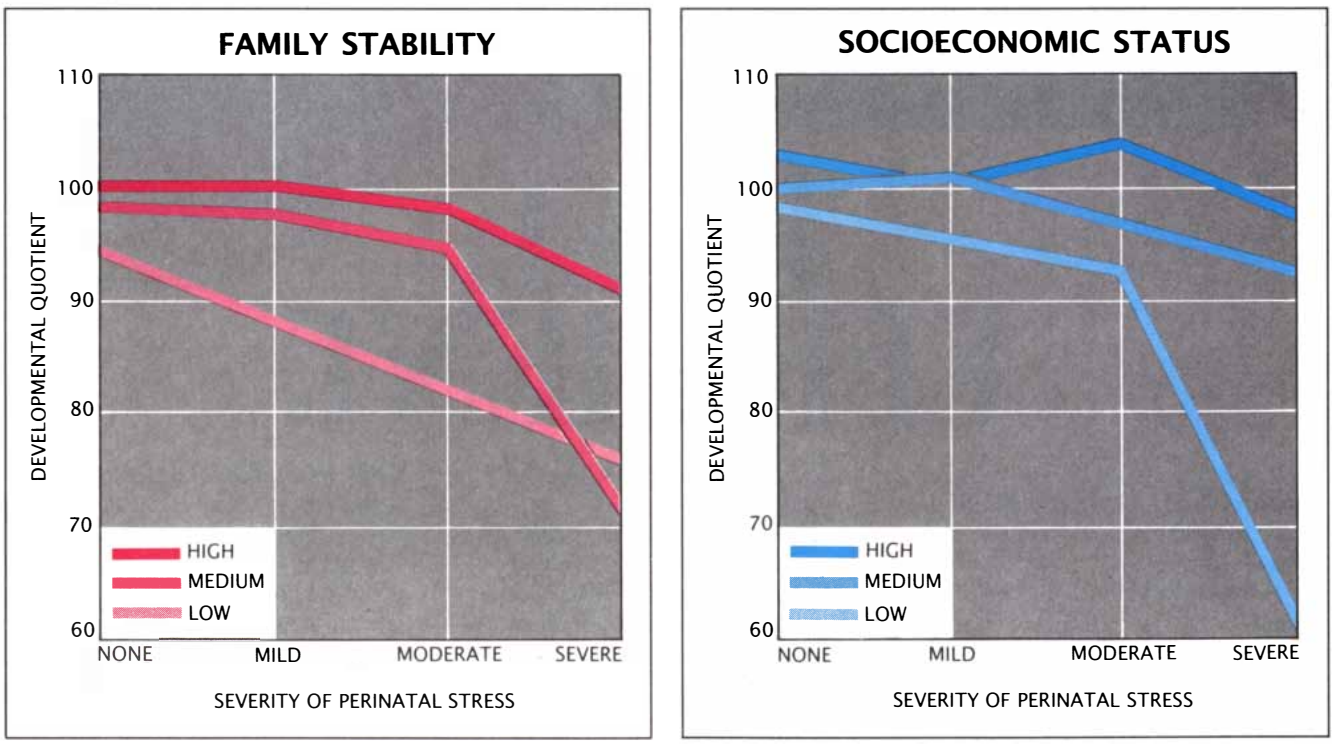
With the help of these support networks, the resilient children developed a sense of meaning in their lives and a belief that they could control their fate. Their experience in effectively coping with and mastering stressful life events built an attitude of hopefulness that contrasted starkly with the feelings of helplessness and futility that were expressed by their troubled peers.

In 1985, 12 years after the 1955 birth cohort had finished high school, we embarked on a search for the members of our study group. We managed to find 545 individuals—80 percent of the cohort—through parents or other relatives, friends, former classmates, local telephone books, city directories and circuit-court, voter-registration and motor-vehicle registration records and marriage certificates filed with the State Department of Health in Honolulu. Most of the young men and women still lived on Kauai, but 10 percent had moved to other islands and 10 percent lived on the mainland; 2 percent had gone abroad.

We found 62 of the 72 young people we had characterized as "resilient" at the age of 18. They had finished high school at the height of the energy crisis and joined the work force during the worst U.S. recession since the Great Depression. Yet these 30-year-old men and women seemed to be handling the demands of adulthood well. Three out of four (46 individuals) had received some college education and were satisfied with their performance in school. All but four worked full time, and three out of four said they were satisfied with their jobs.

Indeed, compared with their low-risk peers from the same cohort, a significantly higher proportion of high-risk resilient individuals described themselves as being happy with their current life circumstances (44 percent versus 10 percent). The resilient men and women did, however, report a significantly higher number of health problems than their peers in low-risk comparison groups (46 percent versus 15 percent). The men's problems seemed to be brought on by stress: back problems, dizziness and fainting spells, weight gain and ulcers. Women's health problems were largely related to pregnancy and childbirth. And although 82 percent of the women were married, only 48 percent of the men were. Those who were married had strong commitments to intimacy and sharing with their partners and children. Personal competence and determination, support from a spouse or mate and a strong religious faith were the shared qualities that we found characterized resilient children as adults.

We were also pleasantly surprised to find that many high-risk children who had problems in their teens were able to rebound in their twenties and early thirties. We were able to contact 26 (90 percent) of the teen-age mothers, 56



INFLUENCE OF ENVIRONMENTAL FACTORS such as family stability (*left*) or socioeconomic status (*right*) appears in infancy. The "developmental quotients" derived from tests given at 20 months show that the rearing environment can buffer or worsen the stress of perinatal complications. Children who had suffered severe perinatal stress but lived in stable, middle-class families scored as well as or better than children in poor, unstable households who had not experienced such stress.

(80 percent) of the individuals with mental-health problems and 74 (75 percent) of the former delinquents who were still alive at the age of 30.

Almost all the teen-age mothers we interviewed were better off in their early thirties than they had been at 18. About 60 percent (16 individuals) had gone on to additional schooling and about 90 percent (24 individuals) were employed. Of the delinquent youths, three-fourths (56 individuals) managed to avoid arrest on reaching adulthood. Only a minority (12 individuals) of the troubled youths were still in need of mental-health services in their early thirties. Among the critical turning points in the lives of these individuals were entry into military service, marriage, parenthood and active participation in a church group. In adulthood, as in their youth, most of these individuals relied on informal rather than formal sources of support: kith and kin rather than mental-health professionals and social-service agencies.

Our findings appear to provide a more hopeful perspective than can be had from reading the extensive literature on "problem" children that come to the attention of therapists, special educators and social-service agencies. Risk factors and stressful environments do not inevitably lead to poor adaptation. It seems clear that, at each stage in an individual's development from birth to maturity, there is a shifting balance between stressful events that heighten vulnerability and protective factors that enhance resilience.

As long as the balance between stressful life events and protective factors is favorable, successful adaptation is possible. When stressful events outweigh the protective factors, however, even the most resilient child can have problems. It may be possible to shift the balance from vulnerability to resilience through intervention, either by decreasing exposure to risk factors or stressful events or by increasing the number of protective factors and sources of support that are available.

It seems clear from our identification of risk and protective factors that some of the most critical determinants of outcome are present when a child is very young. And it is obvious that there are large individual differences among high-risk children in their responses to both negative and positive circumstances in their caregiving environment. The very fact of individual variation among children who live in adverse conditions suggests the need for greater assistance to some than to others.

If early intervention cannot be extended to every child at risk, priorities must be established for choosing who should receive help. Early-intervention programs need to focus on infants and young children who appear most vulnerable because they lack—permanently or temporarily—some of the essential social bonds that appear to buffer stress. Such children may be survivors of neonatal intensive care, hospitalized children who are separated from their families for extended periods of time, the young offspring of addicted or mentally ill parents, infants and toddlers whose mothers work full time and do not have access to stable child care, the babies of single or teen-age parents who have no other adult in the household and migrant and refugee children without permanent roots in a community.

Assessment and diagnosis, the initial steps in any early intervention, need to focus not only on the risk factors in the lives of the children but also on the protective factors. These include competencies and informal sources of support that already exist and that can be utilized to enlarge a young child's communication and problem-solving skills and to enhance his or her self-esteem. Our research on resilient children has shown that other people in a child's life—grandparents, older siblings, day-care providers or teachers—can play a supportive role if a parent is incapacitated or unavailable. In many situations it might make better sense and be less costly as well to strengthen such available informal ties to kin and community than it would to introduce additional layers of bureaucracy into delivery of services.

Finally, in order for any intervention program to be effective, a young child needs enough consistent nurturing to trust in its availability. The resilient children in our study had at least one person in their lives who accepted them unconditionally, regardless of temperamental idiosyncracies or physical or mental handicaps. All children can be helped to become more resilient if adults in their lives encourage their independence, teach them appropriate communication and self-help skills and model as well as reward acts of helpfulness and caring.

Thanks to the efforts of many people, several community-action and educational programs for high-risk children have been established on Kauai since our study began. Partly as a result of our findings, the legislature of the State of Hawaii has funded special mental-health teams to provide services for troubled children and youths. In addition the State Health Department established the Kauai Children's Services, a coordinated effort to provide services related to child development, disabilities, mental retardation and rehabilitation in a single facility.

GRANDPARENTS or other adults can provide a supportive, nurturing role for a child if a parent is absent or becomes ill.

The evaluation of such intervention programs can in turn illuminate the process by which a chain of protective factors is forged that affords vulnerable children an escape from adversity. The life stories of the resilient individuals on the Garden Island have taught us that competence, confidence and caring can flourish even under adverse circumstances if young children encounter people in their lives who provide them with a secure basis for the development of trust, autonomy and initiative.

FURTHER READING

KAUAI'S CHILDREN COME OF AGE. Emmy E. Werner and Ruth S. Smith. The University of Hawaii Press, 1977.

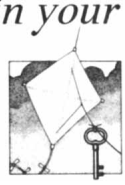
VULNERABLE BUT INVINCIBLE: A LONGITUDINAL STUDY OF RESILIENT CHILDREN AND YOUTH. Emmy E. Werner and Ruth S. Smith. McGraw-Hill Book Company, 1982.

LONGITUDINAL STUDIES IN CHILD PSYCHOLOGY AND PSYCHIATRY: PRACTICAL LESSONS FROM RESEARCH EXPERIENCE. Edited by A. R. Nichol. John Wiley & Sons, Inc., 1985.

HIGH RISK CHILDREN IN YOUNG ADULTHOOD: A LONGITUDINAL STUDY FROM BIRTH TO 32 YEARS. Emmy E. Werner in *American Journal of Orthopsychiatry,* Vol. 59, No. 1, pages 72–81; January, 1989.

THE AMATEUR SCIENTIST

How to build a Planck-mass accelerator in your solar system



by Antoni Akahito

Since the cave dwellers first collided flint against flint to produce fire, natural philosphers have had to resort to ever higher energies in their quest to unlock nature's minutest secrets. The Superconducting Supercollider is the latest instrument to be brought to bear in the physicist's eternal search for truth, but with the site allocation already taken care of, it is long past time to look toward the next step in mankind's greatest journey. Since one can expect congressional fluctuations to obstruct the progress of science for some years to come, this month I wish to encourage proponents of small-scale science—amateurs in particular—to grasp the torch and construct a Planck-mass accelerator.

The Planck mass, or Planck energy (the equivalence of mass and energy by $E=mc^2$ makes the terms interchangeable), is the largest energy that physics as now constituted can deal with in any sensible fashion. It is the energy an average particle had 10^{-43} second after the big bang when the forces of quantum mechanics and gravity are thought to have been unified. A Theory of Everything, of which current superstring theories may be dim precursors, would explain the unification and in fact could be experimentally tested by a Planck-mass accelerator. In principle there is little difference between such a machine and its ancestors: protons or electrons are accelerated up to Planck energies and collided head on. During the collisions the projectiles convert their energy into Planck-mass particles, particles that existed at the earliest instants of creation. Boiled down to its essentials, a Planck-mass accelerator simulates the big bang.

Beside such a machine, existing accelerators pale into insignificance. Consider the proton. Its mass is about 10^{-24} gram, orders of magnitude below the sensitivity of the best laboratory balances. Through $E=mc^2$ it harbors an equivalent energy of about one billion electron volts (or one giga electron volt, abbreviated GeV). The world's largest accelerator, Fermilab's Tevatron, can accelerate protons to energies of 2,000 GeV, now usually abbreviated as 2 TeV for two tera electron volts. The Tevatron, then, can impart to protons about 2,000 times their rest mass in energy. If two such protons are collided together in the Tevatron, the energy can be used to create new particles with masses of approximately 10^{-21} gram, still far below the sensitivity of any laboratory balance. The Superconducting Supercollider (SSC) is designed for 20-TeV operation, only 10 times higher than Tevatron energies.

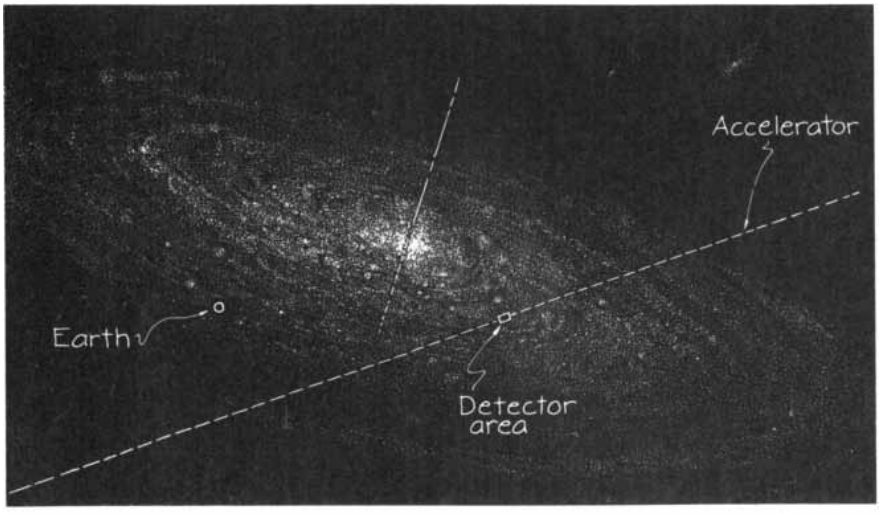
Cosmic rays do somewhat better: the highest-energy cosmic rays, believed to be produced by astronomical objects such as Cygnus X-3, are measured at roughly 10^6 TeV. Particles created in cosmic-ray collisions would weigh in at about 10^{-15} gram.

Grand unified theories (GUT's), however, which profess to combine the strong, weak and electromagnetic forces into one strong-electroweak force, are thought to begin operating at about 10^{12} TeV. This is a million times higher than the most energetic cosmic rays and 100 billion times higher than the expected attainments of the SSC.

Still, we are going for the Max. The Planck energy, where the Theory of Everything is presumed to come into play, corresponds to approximately 10^{16} TeV. This is 10 billion times more energetic than the most energetic cosmic ray. It is 1,000 trillion times more energetic than the particles that will be produced by the SSC. It corresponds to a mass of about 10^{-5} gram, which can be measured on today's laboratory balances.

The first and most difficult step in building the Planck-mass accelerator is finding a name for it. The Superconducting Supercollider has already overburdened the growing list of endeavors anointed with the adjective "super" (now elevated to the rank of nonhyphenated prefix): superconductors, supersymmetry, superparticles, superstrings, supercomputing centers and Supertuesdays. The Superconducting Supercollider has even managed to usurp two supers in as many words and has acquired a three-initial abbreviation in the bargain.

It is certainly easy to be sympathetic: the SSC will be 87 kilometers in circumference and will cost $5 billion (barring overruns). Still, all this is vaguely unsatisfactory. "Super" in the context of accelerators has very much the same ring as the term "postmodern" in literature. What does "post-



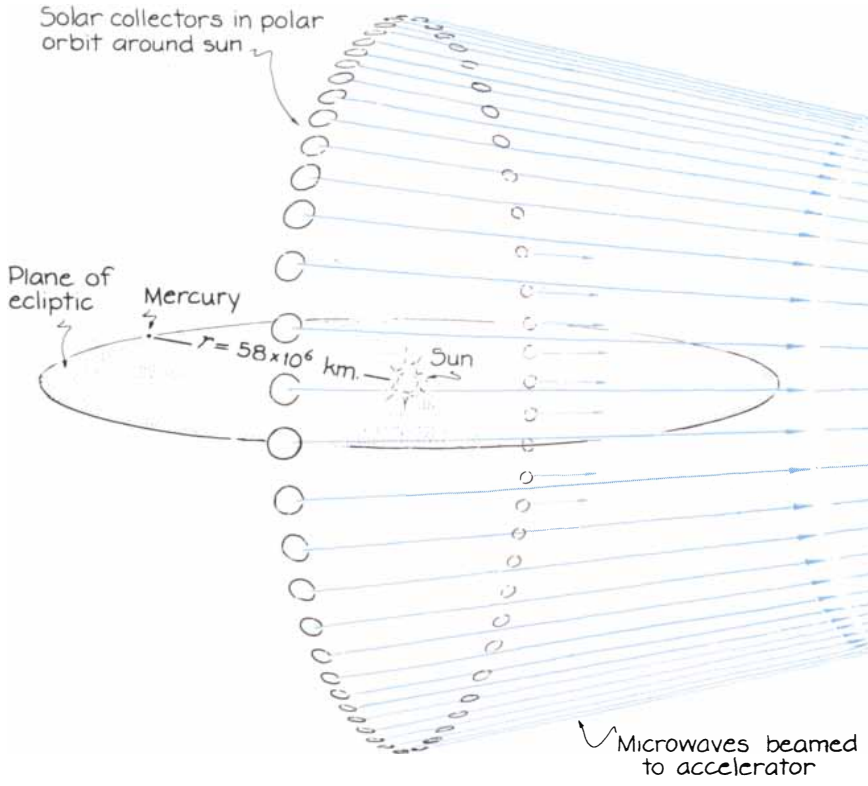
Stanford Linear Accelerator scaled up to Planck energies

modern" become after 10 years? If an accelerator designed for 20 TeV is to be anointed with the adjective "super," then in the next generation we shall be forced to go to "Hypercollider," and then no doubt to "Superhypercollider" and "Hypercolossalcollider," at which point accelerator naming begins to sound like *Sneak Previews*. Clearly what is super today is superfluous tomorrow. Therefore for the Planck-mass accelerator I suggest "Ultimate Collider," or UC. Modest though two initials may be to describe a machine of 10^{16} TeV, it will have to do; as I have said, according to our present conceptions of space and time, it does not make any sense to talk about anything larger.

The second step in designing the UC is to consider what kind of power source will be needed to accelerate protons or electrons up to the Planck energy and create Planck-mass particles. A simple arithmetic calculation reveals the first obstacle: the entire energy of a one-megaton atomic bomb converted into planckons (as I shall call them) will produce about three million. Three million planckons may seem like a lot, but it is negligible compared with the beam intensities achieved by today's accelerators. Machines such as Fermilab's Tevatron are typically capable of delivering 10^{12} to 10^{13} particles per second to the target. Consequently, to achieve today's beam intensities, the amateur will need the energy equivalent of roughly a million one-megaton bombs exploding per second.

This computation assumes, obviously, that 100 percent of the energy of the atom bomb goes into making planckons, which is overly optimistic. The actual efficiency of present-day accelerators is difficult to judge. A beam intensity of 10^{13} particles per second at an energy of 20 TeV represents a power of about 30 megawatts. If, as planned, a 300-megawatt power plant is to be built for the SSC, a 10^{13} particle-per-second beam intensity implies an efficiency of 10 percent; the rest is lost to refrigeration of the magnets, transmission lines and so on. If the beam intensity is only 10^{12} particles per second, the SSC will be about 1 percent efficient. Of course, with room-temperature superconductors (which the do-it-yourselfer can fashion empirically in the kitchen) refrigeration losses can be considerably reduced, if not eliminated entirely.

Nevertheless, I want to be on the safe side, and so I shall assume that the prototype UC will have an efficiency of only 1 percent. With a 1 percent efficiency, the power source for the UC will have to provide the equivalent of 100 million one-megaton bombs per second during operations. This is far above the megatonnage available in today's arsenals.



Solar collectors convert .1 percent of solar luminosity into microwaves

It does correspond, however, to approximately 4×10^{30} ergs per second, or only about a thousandth the luminosity of the sun, which is well within the range of a science-oriented society. The amateur, then, should begin by placing a system of solar collectors in orbit around the sun. If they are placed at the radius of Mercury's orbit, the combined collection area should be at least 4×10^{13} square kilometers, about 660 times the surface area of Jupiter. The solar energy should then be transformed into microwaves, for example, and beamed to the accelerator proper. A large capacitor bank is recommended, for it will significantly reduce the required collection area. (Canal Street in Manhattan has traditionally been a good hunting ground for junk parts.)

Having solved the problem of energy supply, the next task is to look into the design of the accelerator itself. Today's machines are predominantly of two types: linear accelerators, or linacs, and synchrotrons. As its name implies, a linear accelerator accelerates particles along a straight line. The world's largest linac currently is the Stanford Linear Accelerator—universally known as SLAC—with a length of three kilometers. The way a linac accelerates electrons, say, is fairly straightforward. A high-frequency alternating electric field, at approximately 1,000 megahertz, is passed down a microwave guide. The phase of the field is arranged so that it will push the electrons down the cavity. In other words, the electron is accelerated by getting it to ride the crest of a wave. A linac has the disadvantage that it can accelerate a particle only once—from beginning to end. The final energy of the particle is limited by the amount of energy the accelerator can impart to it on a single pass.

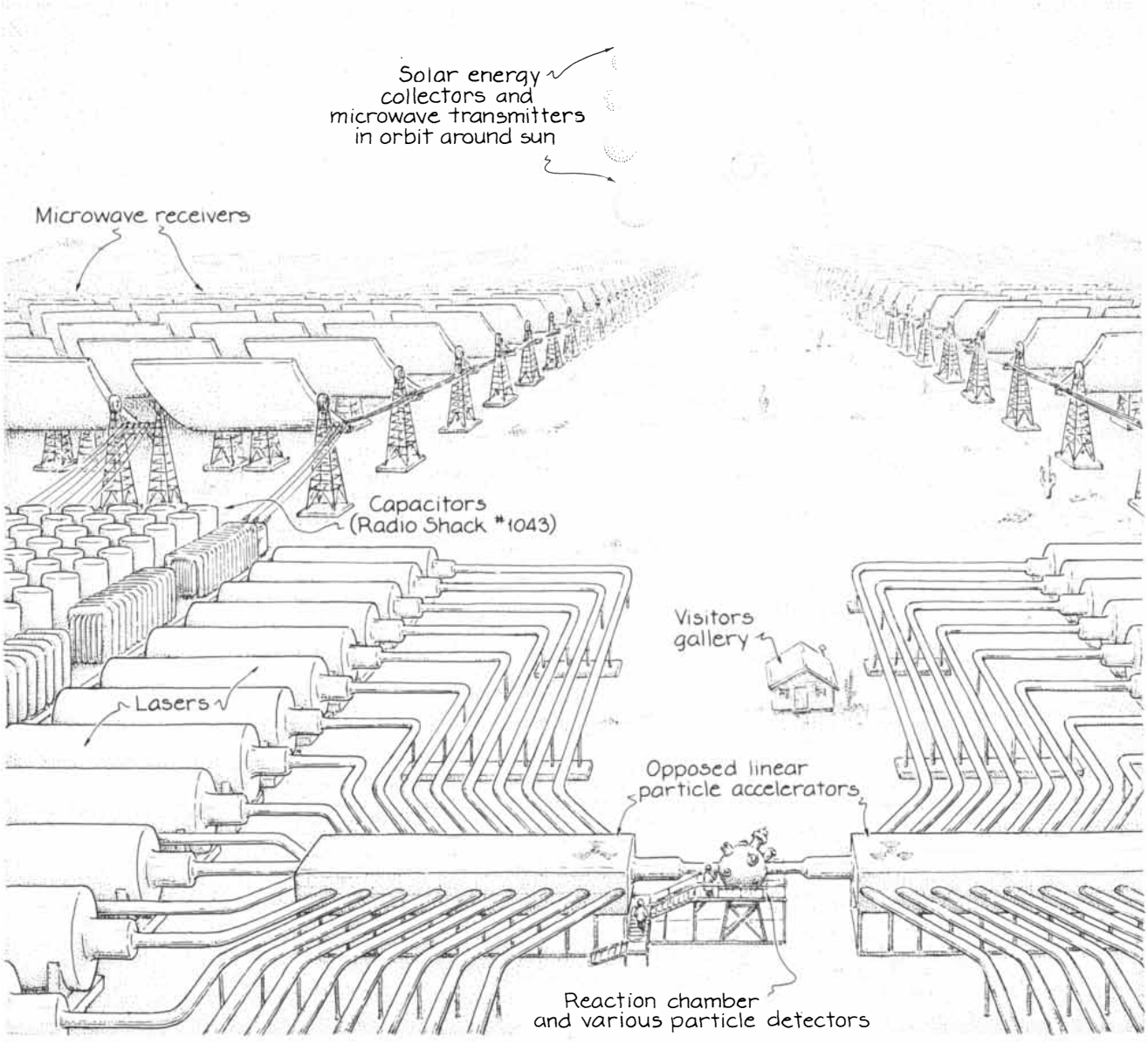
In contrast, a synchrotron accelerates particles repeatedly around a single circular track. Synchrotrons are thus capable of much higher energies for a given length than the linear accelerator is, and largely for this reason the SSC has been designed as a synchrotron. It will also utilize an increasingly popular technique known as colliding beams, which is why "collider" follows the second "super" in SSC. According to relativity, the energy available to create new particles is much greater when two protons or two electrons collide head on than it is when they hit a stationary target in the laboratory. A proton collider therefore cir-

culates two beams of protons in opposite directions until they attain the required energy and then forces them into a head-on collision. In the SSC the full 40 TeV of the two protons is then available to create new particles, each with an energy of 20 TeV. For a noncolliding-beam synchrotron to yield a pair of 20-TeV particles from a single proton smashing into a laboratory target, it would have to accelerate the proton to an energy of approximately 800,000 TeV.

For that reason colliding-beam synchrotrons are now considered the wave of the future. Unfortunately simple considerations show that synchrotrons—be they stationary-target or colliding-beam—cannot be the basis of the Ultimate Collider (without significant difficulties); the amateur is urged to avoid them.

According to a century-old result of Maxwell's theory of electromagnetism, any accelerating charged particle radiates energy. One of the basic problems any accelerator designer faces is knowing how much energy the electrons will lose as they hurtle down SLAC's vacuum chamber, or how much energy protons will give off as they circulate in the SSC's storage rings. Left to themselves, these circulating protons would sooner or later radiate away all their energy and stop. And so some fraction of the energy input of an accelerator simply goes into replacing the energy the particles lose as they are accelerated.

The amount of energy lost in an accelerator depends very crucially on the design. Synchrotrons are prey to an illness appropriately termed synchrotron radiation: the radiation emitted by any charged particle in a circular orbit. In Cornell University's 10-GeV synchrotron, a 10.5-MeV boost is given to an electron on each turn, but the losses from synchrotron radiation on each turn are about 8.85 MeV. And so, you see, at high energies most of the energy goes not into accelerating particles but into making up radiation losses. Unfortunately synchrotron radiation goes up as $(E/m)^4$, where E is the particle energy and m is its rest mass—in other words, very rapidly. By the time you reached only 10^4 TeV—5,000 SSC energies—an electron circulating in a synchrotron of 100-kilome-



A Planck-energy proton collider

ter radius would be radiating away a Planck mass of energy on every turn.

Radiation losses are, however, inversely proportional to the radius of the accelerator; an obvious strategy, then, is to make the radius of the accelerator larger. This is not very feasible. The radius necessary to keep a Planck-energy electron radiating at less than one Planck energy per turn is roughly 10^{27} times the size of the observable universe.

Because synchrotron radiation losses go as $(E/m)^4$, such losses are less severe for protons, which are much heavier than electrons. Specifically, the proton is almost 2,000 times more massive than the electron, and so at a fixed energy synchrotron radiation losses are about 10^{13} times less. But the factor of E^4 means that once a proton is accelerated to an energy 2,000 times higher than that of an electron, radiation losses will be the same: in an accelerator with a radius of 100 kilometers, at about 10^7 TeV radiation losses exceed one Planck mass per turn. To keep the radiation losses from Planck-energy protons within acceptable bounds, one would need to construct a synchrotron with a radius 10^{14} times the size of the observable universe.

Luckily for the amateur, there is a solution to the problem. Radiation losses in a linear accelerator turn out to be vastly less severe than those associated with synchrotrons. In a linac the power lost to radiation can always be kept below the input power simply by keeping the energy given to the electrons below the order of 10^6 TeV per centimeter. SLAC provides an "energy gradient" of roughly 10^{-7} TeV per centimeter, which is 13 orders of magnitude below the upper bound. Protons, because they are heavier, are again subjected to a less stringent limit, in this case about 10^{13} TeV per centimeter.

And so there we are. To be conservative, the Ultimate Collider prototype should be constructed as a linac. As long as we keep the energy gradient below the limit of 10^6 TeV per centimeter, we can attain arbitrarily high energies. What is more, we want to make it a colliding linac—two linear accelerators run in opposite directions to capitalize on the full energy available in head-on collisions.

The first obvious feasibility test is to scale up SLAC to Planck energies. At 10^{-7} TeV per centimeter, however, this calls for an accelerator 100,000 light-years long, somewhat greater than the size of the galaxy. A collider would be twice as long, with the laboratory area presumably at the center. Such unwieldy proportion makes data collection inconvenient; an experimenter, after throwing the "on" switch, would have to wait 200,000 years for the results.

Again we are saved by the fact that radiation losses in linacs are so small. If we choose to construct a machine with an energy gradient of 100 TeV per centimeter—still far below the limit of 10^6 TeV per centimeter—the length of the UC is reduced enough so that it would fit within the orbit of Pluto. At first glance an energy gradient of 100 TeV per centimeter—which is one billion times as large as the gradient at SLAC—strikes one as large, if not impossible. SLAC's accelerating field is produced by a bank of more than 200 high-frequency oscillators known as klystrons, and they produce about the maximum gradient attainable by conventional methods. The thought of increasing 200 klystrons to 200 billion does not seem very fruitful.

But the klystron is not the only device capable of generating large amounts of power. The highest power available today actually comes from lasers. Even a commercially available CO_2 laser can result in gradients 10 times as high as those at SLAC, and the HELIOS laser at the Los Alamos National Laboratory is already up by a factor of 1,000. Some investigators are now talking about future laser-driven machines with gradients of 10^{-4} TeV per centimeter, only a million times less than our goal.

At this stage the amateur must overcome a serious obstacle. Electrons are bound to atoms with energies of about 10 electron volts. Atomic dimensions are of order 10^{-8} centimeter. One therefore expects that field gradients larger than, say, 10 eV per 10^{-8} centimeter will tear electrons from their nuclei. In terms of our units this upper limit is about 10^{-3} TeV per centimeter, about 100 times larger than the SLAC gradient. Even smaller values—values too small for the UC by a factor of 100,000 or one million—would without doubt cause serious damage to the accelerator's support structure.

The amateur must therefore look to new media to construct the Ultimate Collider. One promising device is a laser-plasma accelerator [see "Plasma Particle Accelerators," by John M. Dawson; SCIENTIFIC AMERICAN, March]. In plasmas, or highly ionized gases, electrons are already detached from their nuclei and so they cannot be further disrupted. Accelerating fields in experimental "beat wave" accelerators have already reached 10^{-5} TeV per centimeter; 10^{-2} TeV per centimeter is theoretically possible with currently attainable plasma densities, which are obtained from hydrogen. This is only 10,000 times below what is required for the prototype UC. Going to denser materials, such as platinum, would increase the acceleration gradient to about a TeV per centimeter, or about 100 times below the goal. White-dwarf or neutron-star material, which is on the order of 10,000 times denser than platinum, could yield plasma densities and gradients that exceed the required magnitude.

The accelerating chambers of present plasma accelerators are only several millimeters long at best. Consequently the amateur will probably have to run a number of machines in "tandem" in order to produce the required energy.

Based on this picture, for the UC prototype one can imagine the rotating neutron star in Cygnus X-3—which spews out 10^6 TeV cosmic rays—harnessed to serve as a booster for the main accelerator (although I am not sure this would be very cost effective). A string of giant lasers stretching across the solar system would then take the particles up to the Planck energy. Admittedly this is an unwieldy prototype, but the creative amateur will no doubt find acceleration mechanisms that are even more efficient than the plasma accelerator and is encouraged to pursue them. Eventually, of course, one will want to go for the theoretical limit of 10^6 TeV per centimeter, at which a fully operational electron UC would be only 10^{10} centimeters long—about a fourth of the distance to the moon. A proton collider could be shorter by a factor of several million—for an accelerator length of order 10 meters.

The final question remains: funding. At the projected SSC price of $250 million per TeV, the UC would cost only $\$2.5 \times 10^{24}$, something more than the U.S. budget deficit (actually about 10^{11} times the gross world product) but already a bargain considering potential spin-offs. Surely, though, one can expect that with increased technological sophistication this cost will decrease to give us the ultimate in big bangs for the buck. In any case, one cannot put a price tag on the philosophical benefit and change in world outlook that such a project will give to our grandchildren and our grandchildren's grandchildren, for with the help of the Ultimate Collider they will have the best chance of understanding the Moment of Creation.

COMPUTER RECREATIONS

A matter fabricator provides matter for thought

by A. K. Dewdney

"Nil posse creari de nilo."
—LUCRETIUS, *De rerum natura*

I was not surprised to receive, nearly a year ago, a long missive from someone who claimed to have invented a matter fabricator. After all, among those who write to me suggesting interesting ideas there are a few whose assertions do stretch credulity. But since the essence of science is an open mind (if not a fully ventilated one), I try not to dismiss such letters until I have read them to the end.

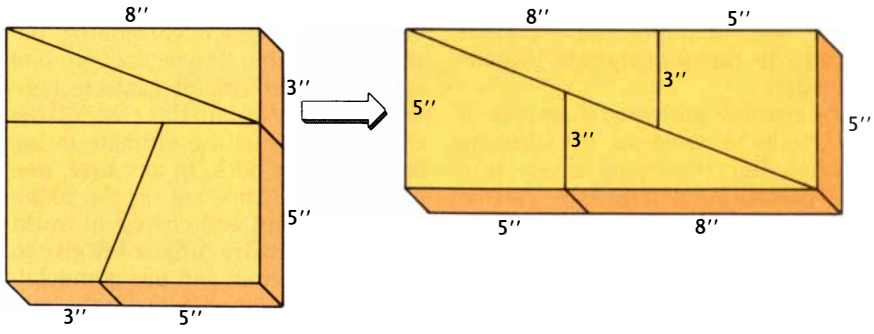
I am glad I did just that with this particular letter, because the inventor based his assertion on a legitimate mathematical result known as the Banach-Tarski paradox. Named for the Polish mathematicians who discovered it in the 1920's, the paradox reveals how under certain conditions an ideal solid can be cut into pieces and then reassembled into a new solid twice as large as the original one.

Indeed, the inventor turned out to be a professional mathematician who has many published papers to his credit. For reasons that will soon become clear, he shuns any kind of publicity and has asked that I call him Arlo Lipof. It was Lipof's familiarity with the Banach-Tarski paradox that first led him to investigate the possibility of applying the paradox to real matter instead of ideal matter. His investigation has paid off handsomely: he has written a computer program that gives precise directions for cutting a solid into many odd-shaped pieces and then reassembling them into a solid twice the size, leaving absolutely no space between the pieces!

Needless to say, the implications of Lipof's program are profound. To explain the paradox and how the program exploits it, I can hardly do better than to quote from Lipof's letter:

"The paradox is similar to the well-known puzzle involving tangrams, little pieces of paper cut into simple geometric shapes. Four such shapes can be assembled into a square that has an area of 64 square inches. Yet the very same pieces can also be assembled into a rectangle that has a greater area—65 square inches, to be precise. If you do not believe such a thing is possible, try cutting up the pieces as shown in my drawing" [*see illustration below*].

"If the little pieces of paper were instead pieces of gold, there would seem to be an automatic increase in wealth in going from the square to the rectangular configuration. Start with a square of gold, say, eight inches on a side and an inch thick. Then cut it according to the figure at the left. If the pieces are reassembled according to the figure at the right, an extra cubic inch of gold will have appeared. The extra cube would weigh approximately 4.3 ounces and at current prices would be worth about $1,800."



How to get a cubic inch of gold for nothing

Lipof goes on to concede that the appearance of getting something for nothing in the above example is purely illusory. But he asserts that although the Banach-Tarski paradox has "the same effect on one's mind," there is no flaw in the theory on which it is based. The Banach-Tarski paradox is real—at least in a mathematical sense.

The paradox arises from a proved theorem that, when stated in technical language, is almost comprehensible: if A and B are any two bounded subsets of $\mathbf{R}^3$, each having a nonempty interior, then A and B are equidecomposable. The theorem can be stated in less technical language if one initially considers a pair of bodies, of virtually any shape and size, that meet two criteria. Each body must be "bounded," or capable of being enclosed in a hollow sphere of some definite size. And each must have a "nonempty interior": it must be possible to envision a sphere somewhere inside the body that is entirely filled with the material of which the body is made.

The two criteria are actually rather modest ones. Indeed, almost any object we might imagine that violates them is hardly the kind of object we would normally call a body. A straight, infinitely long line, for example, violates both criteria: it is not bounded and its interior is empty in the sense that it has no interior to speak of. Also disallowed would be an imaginary cloud of points stretching to infinity in all directions—hardly a body in the usual sense of the word.

According to the theorem, then, any two such bounded bodies having nonempty interiors are "equidecomposable." This means that one can dissect both bodies into a finite number of pieces that are congruent in a geometric sense: a piece of one body can be made identical with a piece of the other body merely by rotating it. Hence one can in theory dissect a body into pieces and label them A_1, A_2, A_3... and dissect a different body into pieces and label them B_1, B_2, B_3..., so that the pieces A_1 and B_1, A_2 and B_2 and so on are identical. That is the essence of the Banach-Tarski paradox.

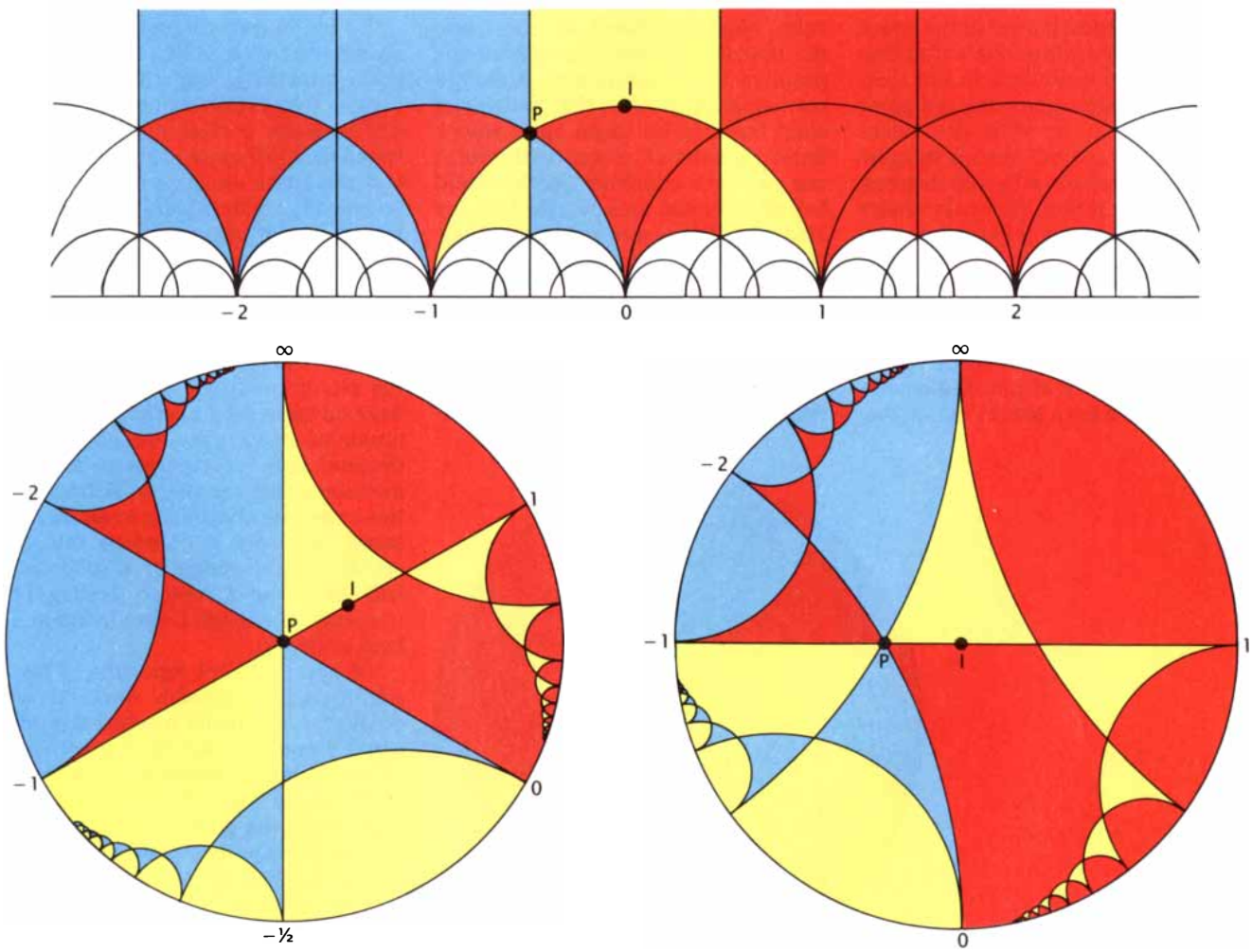
"It is thus possible," Lipof writes, "to take two solid spheres, one twice as large as the other, and cut them into pieces that are pairwise congruent. Forget the larger sphere and consider only the smaller one. Imagine that it is made of gold. In principle it can be cut into a finite number of pieces that can then be reassembled into a sphere twice as large."

There is no trickery here, although one must realize that there are certain topological qualifications implied in the innocent word "pieces." For one thing, they are not necessarily simple in form, or even composed of connected parts. Some parts of the same piece may come arbitrarily close to one another without actually touching. Moreover, the pieces cannot be measured in any precise way. For example, one cannot even imagine a way of gauging their exact volume. What would be the actual appearance of such pieces? Lipof says they are "like nothing you have ever seen before. They make fractals look like tangrams."

The Banach-Tarski paradox holds in its most general form only in spaces of three or more dimensions. There are, however, closely related theorems that illustrate the nature of the paradox in spaces of lower dimensions. A crude example of this phenomenon is given by the one-dimensional "space" consisting of all integers, since the subset of even integers represents simultaneously both half of and the whole of the set. The subset is half of the set of all integers in the sense that only every other integer is in the subset. Yet a simple transformation—dividing each element in the subset by 2—turns the even integers into the entire set of all integers; the set and its subset are the same size.

Most people do not find this fact very remarkable, because the size of both set and subset happens to be infinite. After all, infinity divided by 2 is still infinity. It would be more exciting to find a finite space that can be decomposed into paradoxical pieces, but that is not possible according to theory, at least if one limits oneself to spaces of a single dimension. The same applies for Euclidean spaces of two dimensions, or "flat" planes. It can happen, however, in certain non-Euclidean two-dimensional spaces.

A full explanation of the phenomenon is beyond the scope of this column, but I can give at least a glimpse of its paradoxical nature by projecting the exotic world of two-dimensional hyperbolic space onto an ordinary Euclidean disk, as is shown in the illustration below. The hyperbolic space occupies a half plane, as is shown in the upper part of the illustration. Its geometry is not Euclidean, wherein the shortest distance between two points is a straight line. Instead shortest distances are found along semicircles. In this illustration the hyperbolic space has been dissected into "triangular" regions that get smaller toward the bottom edge of the space. The triangles form the basis of a paradoxical decomposition of the space into three pieces that are colored red, blue



Two projections of hyperbolic space **(above)** ***on disks*** **(below).** ***The red piece is both one-half and one-third of the hyperbolic space***

and yellow. To mathematicians, in this context a "piece" may not be all of a piece, so to speak. It may be composed of an infinite number of fragments, triangular and otherwise.

The strange nature of the three pieces becomes most apparent when one views the hyperbolic space through a special mathematical porthole, that is, by projecting the space in two different ways onto a disk. The point labeled *P* in the hyperbolic space lies at the center of the left-hand disk and the point labeled *I* lies at the center of the right-hand disk.

In each disk a simple rotational symmetry involving the three pieces becomes evident. Consider the red piece in the disk at the left. If one imagines rotating the disk about its center by 120 degrees, or a third of a revolution, one sees that it would end up—fragment for fragment—on top of the yellow piece. Another 120-degree rotation would similarly match the red piece with the blue one. In other words, all three pieces are congruent, and together the three pieces make up the entire hyperbolic space.

The paradoxical nature of the space becomes evident when one's attention turns to the second disk. In this view of the very same space the red piece is congruent to the other two pieces combined! To see this, merely imagine rotating the red piece by 180 degrees, or half a revolution, about the disk's center. The red piece will overlie exactly both the blue and the yellow piece. There is therefore a piece of two-dimensional hyperbolic space (the red piece) that amounts simultaneously to a half and a third of the entire space.

The significance of this demonstration may have been lost in the mathematical shuffle. "Why," asks the reader, "should I be impressed?" The reason is that the three sets represented by the colored pieces are absolutely true congruences in hyperbolic space. The fact that the red piece does not appear congruent to a half and a third of hyperbolic space at the same time on a single disk is a consequence of the distortions of hyperbolic space associated with such projections.

There is no need for hyperbole or hyperbolic spaces in proving the most general version of the Banach-Tarski paradox, however. The fact is that in Euclidean three-dimensional space (which approximates the world we live in) any two bodies that satisfy the most modest conditions imaginable are equidecomposable. Unfortunately the proof is nonconstructive: it gives almost no clue to precisely how one would go about demonstrating the equidecomposability of two unequal solid balls.

At this point I quote again from Lipof's letter:

"I spent many years studying the Banach-Tarski paradox and related results. What fascinated me most was the nonconstructive character of the proof in three dimensions. Although mathematicians know that in theory a solid ball can be taken apart into a finite number of pieces with which one can then construct another solid ball of twice the size, no one had any idea what the pieces might look like, because the cutting of the pieces is based on what set theorists call the axiom of choice.

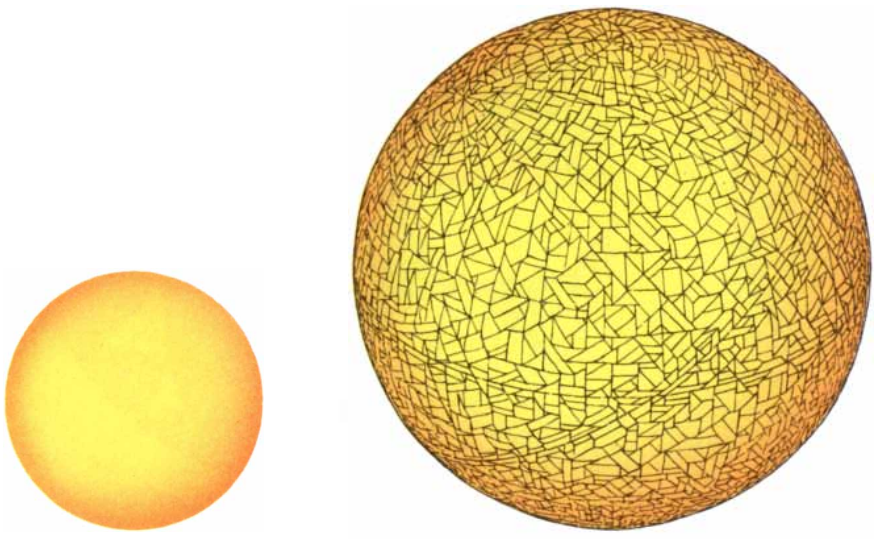
"The axiom gets its name not because mathematicians prefer it to other axioms but because it postulates that for any collection of sets, no matter how big, there is a way of choosing an element from each set in the collection. Indeed, many mathematicians prefer not to invoke the axiom, because it does not stipulate just how the matching is done.

"No one, therefore, had any idea of what the pieces of a paradoxical decomposition might look like until I began to investigate the question. I mechanized the proof of the Banach-Tarski paradox. The proof specifies that the second (larger) ball can be assembled by rotating the pieces of the first ball in two ways about its center, somewhat like the situation in hyperbolic space.... These rotations carry each piece of the first ball to its corresponding place in the second ball. Knowing the points that make up each piece and the necessary rotations, it was easy to construct a backtracking routine to cut the different pieces from a solid ball. Whenever the proof invoked the axiom of choice, I merely relied on a random-number generator in my personal computer to choose which points of the ball were to be elements in which sets.

"To be honest, throughout this research I had no idea that I was headed in the direction of a matter fabricator. I am no fool: I know that normally one has to make a clear distinction between the ideal spaces of mathematics and the space we live in. But when I completed my first simulation of the Banach-Tarski theorem, I realized I had in hand something like a recipe for doubling the size of any solid.

"It occurred to me to try an experiment involving a real material, but I initially held back. The dimensions of the pieces produced by my program were all expressed in triple precision numbers, an accuracy that might well demand that I cut atoms in two in manufacturing the pieces! Besides, at this stage I was beginning to doubt my sanity: the idea of carrying out an actual decomposition of a solid ball had given me a distinct feeling of unreality, as though I were living in a kind of dream.

"Of course, I kept repeating to myself, it couldn't possibly work. But to no avail. I eventually reached a point where I couldn't put the experiment off any longer. I cashed in a good part of my life savings to buy 12 ounces of gold. I had the gold cast into a ball, bought a tiny jeweler's saw and began to cut the ball up according to my program's recipe. A second computer program was most useful in this process. It catalogued the size and shape of each piece. In particular, the second program told me where each



Original gold ball **(left)** ***and Arlo Lipof's reconstruction of it*** **(right)**

piece was to go in the second sphere.

"The entire experiment took seven months from start to finish. I worked nights and weekends. When at last I had finished cutting the pieces, I began assembling them into a second sphere of twice the diameter. It was delicate and demanding work. I nearly lost my eyesight and began to get headaches, but I persevered. Slowly the second ball took shape—but not as a smooth ball. The pieces did not fit as well as I had hoped. There were tiny spaces between the fragments I had so painfully and carefully put in place with tweezers.

"After a few more weeks I finally completed the ball. I have sent you a drawing of the major joints on its surface [*see illustration on opposite page*]. Having your readers see the map gives nothing away: the surface of the ball is child's play compared with the intricate arrangement of its interior pieces. In any event, the actual ball was not as smooth and round as my picture implies. It was lumpy and irregular—downright ugly. But how I clutched the cloth bag that contained it on my way to the jeweler! The ultimate test of all my work, of course, would be to melt the ball down and find out whether I was indeed the owner of up to eight times as much pure elemental gold as I had started out with.

"The next day the jeweler handed me a bar of pure gold weighing 49.58 ounces. It was less than I had expected; those interstitial spaces had taken their toll. Yet the thing was no longer in doubt. The world's first practical application of the Banach-Tarski paradox had been made. For days I staggered around like a drunk, reeling from my discovery. At this point I am not sure what to do next."

After that first letter I received no further correspondence from Lipof for several months. Then one day last November the mailman brought me a short missive from him, postmarked in a South American country:

"You will no doubt be happy to learn that I have to some extent automated the procedure of producing large balls of gold from small ones. With the remainder of my life savings I have set up shop in the little town of _____. Here a few loyal employees assemble gold balls. There is a workroom lined with computers and with tables at which my people assemble the balls. The pieces are now not cut out but rather are cast directly and worked by my employees. There is always excess gold at the end of the process with which to begin anew. We produce approximately five pounds of gold a week from nothing. Is this not the philosopher's stone?

"The time will soon come to move on. I do not think I will write again; to communicate with you is dangerous. Excuse me, my friend, but one becomes paranoid in the presence of such potential. There is much that I need to do!"

I have not heard anything more from Lipof. But last December, out of curiosity, I began to track the price of gold from day to day. For nearly three months it has been in a slow but steady decline. Perhaps that is the ultimate proof for those who thought the Banach-Tarski paradox was merely a plaything of mathematicians.

I have, of course, been in touch with other mathematicians on the subject of the paradox. I owe a particular debt of gratitude to Bruno W. Augenstein of the Rand Corporation in Santa Monica, Calif. It was Augenstein who suggested that I use hyperbolic space as an example of the paradoxical properties of space.

Although he does not subscribe to Lipof's claims, Augenstein does concede that there may well be a relation between the Banach-Tarski paradox and the real world. One of Augenstein's papers, "Hadron Physics and Transfinite Set Theory," points out a relation between particle physics and paradoxical decompositions of objects in three-dimensional space. The paper suggests analogies that "give directly a large number of known physical results and suggest additional ones testable in principle. The quark color label and the phenomenon of quark confinement . . . have immediate explanations via analogies with the decomposition theorems." This much might interest the physicists, if not the alchemists, among the readers of this column.

There are a few self-professed "bit flippers" in the world, readers who keep close watch for the appearance of a project that promises number-crunching complexity equal to their talents. The two-part series on cryptology last fall (in the October and November issues) raised a cheer in this quarter; in particular, the description of the Data Encryption Standard (DES) in the second part provided ample grist for a bit flipper's mill. The DES is a scheme for encrypting computer messages that is used not only by commercial institutions but also quite possibly by various military installations around the world. It is long and complicated, but that is just the way Mike Rosing of Darien, Ill., likes it. Eschewing software that is too soft, Rosing writes his own programs in 68000 assembly code, a low-level computer language that lies at the hardware heart of the 68000 microprocessor chip.

There is nothing like writing and testing a program (at any level) for revealing bugs in the original specifications. The input for the P permutation table in the F module was mislabeled; instead of 48 bits it should be 32 bits. Decryption also produced problems for Rosing and the others. The 64 bits of the original key are not fed in reverse order, the 48-bit "subkeys" are. They are fed into the central block starting with key 16 and ending with key 1.

Charles Kluepfel of Bloomfield, N.J., wondered what parts of the DES were arbitrary. For example, must the E bit-selection table really have the form I gave in order for a successful data-encryption system to emerge? And what about the substitution tables? Wafting rumors hold that the designers of the DES deliberately included "side doors" in certain parts of the cryptosystem that make it somewhat easier to decrypt DES ciphertext without knowing the original key.

Daniel Wolf of Santa Maria, Calif., has written several programs of interest in assembly code for his 68000-based Amiga computer. Amiga owners may request copies of a cryptosystem based on the famous Enigma machine (described in the October issue) or on the RSA algorithm (in the November issue). A major virtue of software cryptosystems that are written in assembly language is their blinding speed. Wolf can be reached at Box 1785, Santa Maria, Calif. 93456.

Still another journal of cryptology, *Cryptosystems Journal,* was brought to my attention by Tony Patti of Burke, Va., who edits and publishes it. The first two issues pursue Patti's goal of describing and distributing state-of-the-art cryptosystems for IBM PC's and compatible computers. Interested readers can get in touch with Patti for further information at 9755 Oatley Lane, Burke, Va. 22015.

FURTHER READING

TANGRAMS—330 PUZZLES. Ronald C. Read. Dover Publications, Inc., 1965.

HADRON PHYSICS AND TRANSFINITE SET THEORY. B. W. Augenstein in *International Journal of Theoretical Physics,* Vol. 23, No. 12, pages 1197–1205; December, 1984.

THE BANACH-TARSKI PARADOX. Stan Wagon. Cambridge University Press, 1985.

BOOKS

The shape of life, where maps come from, the world's ships and the country's people

by Philip Morrison

AXIS AND CIRCUMFERENCE: THE CYLINDRICAL SHAPE OF PLANTS AND ANIMALS, by Stephen A. Wainwright. Harvard University Press ($22.95).

Thinking reed, featherless biped, immature ape? Reader, let the philosophers wrangle over their metaphors; you and I, along with most plants and animals, are a set of linked cylinders: twigs and digits, limbs and trunk. The sculptor well knows it, and now a reflective and witty morphologist gives us this gem of a brief, provocative essay in which he sketches out just how and why the bodies of sessile and mobile life alike are built of cylinders. Six apt chapters on shapes, on the mechanical properties of the key materials and on the conjectural evolution of the form of organisms are supported by visual evidence in a fine set of documented line drawings by a number of illustrators. They make a lively case for viewing evolutionary engineering from the top down in a search for overall adaptive form. "Structure permits the functions that exist . . . and . . . changing structure is function."

Posture and movement in life arise out of cylindrical form; that form in turn arises unavoidably out of the mechanical properties of the structural materials all life shares. Scale comes first. The unicells are small, and for them the microenvironment is similar in all directions, although it changes in time. They are spheres, or close to it, but they secrete mucoid coats of various polymers, and those molecules determine structure anew. Self-adhesive, the secretions promote sphericity and isolation; when the cells adhere to a substrate or to partners, selective adhesivity may generate the mat forms of embryonic multicellularity.

Now consider more complex polymers, with cross-links and even pads of ionic charge, that are able to control molecular size and shape by forming and breaking links as the chemical environment may determine. Here we envision a sequence of natural materials that tend toward increasing stiffness. The organisms change. At first there are sticky-backed worms. Then a cross-linked coating develops from the mucous substance to offer a horny carapace. Finally crystals of calcite become bound by the protein collagen and we have imagined hard-shelled life, with bone one step away.

Cross-linked polymers are often oriented into fibers, particularly when cells are motile. Simple movement is linear, and any marked anisotropy of interaction leads under linear motion to parallel orientation. With that a cylindrical plan is born—a plan fit to encounter a streaming world with its inescapable directionality, whether of flow, of gravity or of a solid substrate.

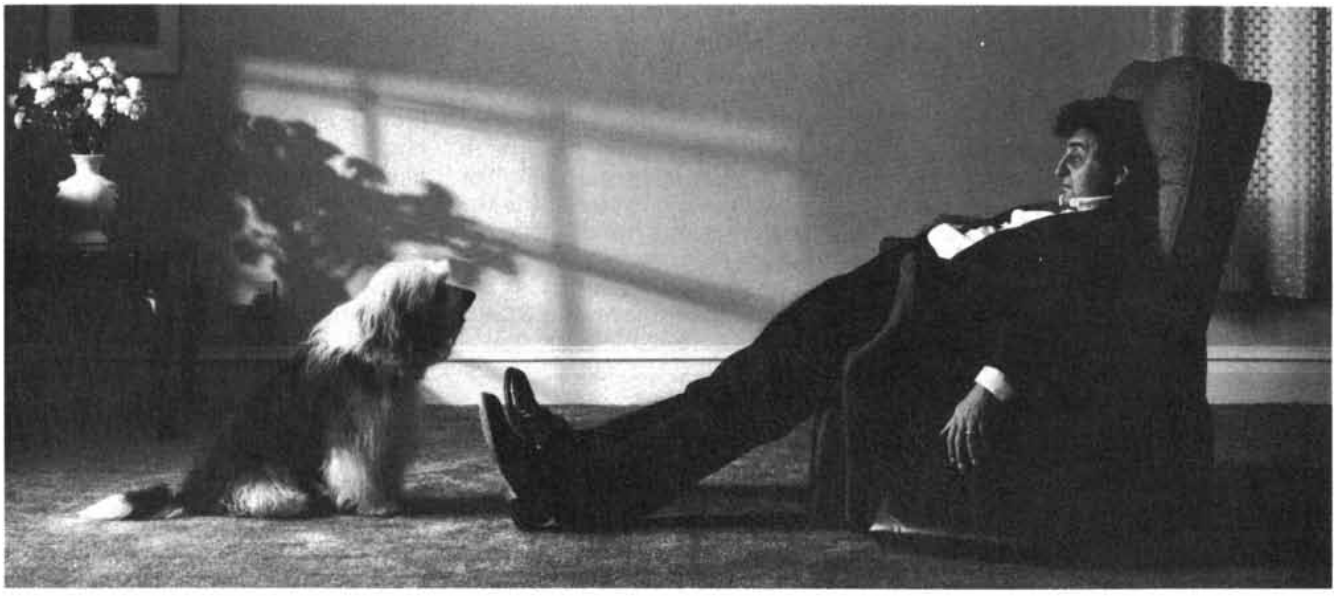
That single axis gave the higher kingdoms a choice. Plants (and many animals) simply settled down to become larger and more complex cylinders. "The basic, original multicellular plant was a filament." Plate-shaped, leaflike organisms and lumpy, bloblike ones abound, but they were broadly outcompeted by those that could instead branch repeatedly. Those many-branched cylinders support the increasingly specialized parts out where there is more space, and where light, food, water, gases to exchange and incoming gametes sometimes flow past. Compare a tree of wood with a chair of wood; the giant sequoia is "the envy of engineers" for its integrity of jointing. Manmade joints are often weakened by discontinuity or grow wasteful and ugly through overdesign. The purposeful sleekness of yachts and aircraft shows what the human designer begins to do.

Unlike plants, most animals chose to move about freely. We may reasonably suppose the first mobile animal was some tiny filamented worm of the sea floor, a "sock full of meat." That sock was woven of helical collagen fibers; the meat within was a set of muscle cells able to swell and shrink. Such a simple hydrostatic combination allows motion with an entire range of mechanical consequences. One of the book's more unexpected passages reminds us of the simple relation between changes in the diameter and the length of cylinders of constant volume. In a short, fat disk a small contraction of length makes a small but forceful increase in diameter; in a long, thin cylinder a small decrease in radius by circumferential contraction makes a sharper increase in length, although at low force. Hydrostatics almost alone can thus build quite subtle moving mechanisms; the body cavities of many phyla of animals are in fact hydrostatic devices made for maintaining posture under stress and for motion. Linked skeletons came later, like branching plants; more cylinders appear, but severer and more diverse material specifications as well.

A choice ends this small book just as a choice opens it. For every axis there is a circumference. Either may be stiffened for strength. The lobster and the bamboo are stiffened outside, on the circumference; there is the optimal placement of hard material. In contrast, we vertebrates are stiffened on the axis. An axial column of strong calcified lumps stiffens a creature, but not too much; softer interspersed disks allow graceful bending. The composite resists both compressive and tensile strain. (A few years ago Professor Wainwright published a paper called "To Bend a Fish.") In the end big muscle masses arrived to allow forceful effort or fast motion through strong, fibrous tensile links joined to long levers of a stiff composite: bone.

Three alternative designs determine the positioning and often the motion of good-size organisms, cylinders still. The stiffish, many-branched cylinder with continuous joints fills the forests and decks the reefs. A "fiber-wound, fluid-filled hydrostat" bends in the meadow, wriggles with the worm, squirts with the squid, twirls with every swift tongue. A flexible multicylindrical framework of posts and beams with artfully tied joints "blesses brittle stars, crinoids, arthropods, and vertebrates."

There are exceptions in plenty, in no way damaging to the theme: the platy lettuce and the box turtle, the lumpy sponge, the radiating urchin. Few books on living forms manage to be so pointed, brief and provocative of reflection. Still, no axial enthusiast



How to solve the energy crisis.

If you're like a lot of people the energy crisis has nothing to do with oil embargos and solar power.

You're suffering from a personal energy shortage. You know, no pep, no get-up-and-go, no spunk.

The reason for this phenomenon? Lack of exercise.

Exercise equals energy.

Research has proven that people who exercise on a regular basis have more stamina and feel better about themselves.

And health and fitness experts agree that no form of exercise is more efficient and effective than cross-country skiing. Not walking, running, rowing, or biking.

NordicTrack® gets you back on the right track.

NordicTrack simply duplicates the cross-country skiing motion, but you don't have to know how to cross-country ski to use it.

This no-impact, total body workout provides a better cardiovascular workout than exercise bikes, rowers, and treadmills.

You burn more calories. You have more energy.

With NordicTrack, you can burn more calories in a 20-minute workout than with any other type of exercise machine.

Plus, you'll feel stronger, more alert, and less stressed-out.

Why NordicTracks never show up at garage sales.

People love their NordicTracks. In fact, 7 out of 10 owners are still using their machines more than 3 times a week, 5 years after purchasing one.

So call NordicTrack.

But do it soon. Before you run out of gas for good.

Free Brochure & Video.
Call Toll Free 1-800-328-5888.
In Canada 1-800-433-9582.

☐ Please send me a free brochure.
☐ Also a free video tape ☐ VHS ☐ BETA

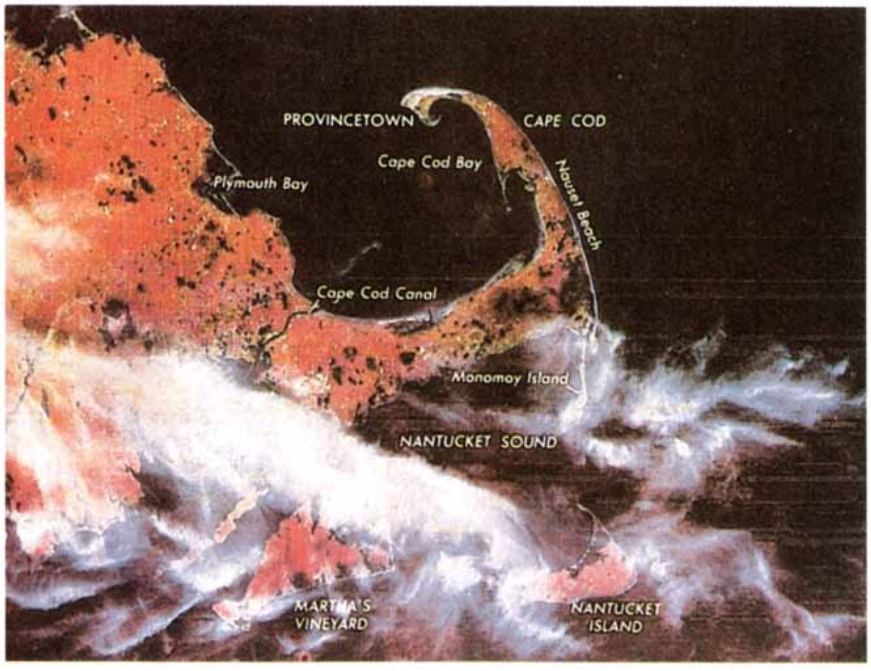
Name_____
Street_____
City_____ State_____ Zip_____
Phone ( )_____
NordicTrack 141 Jonathan Blvd. N. Chaska, MN 55318
320D9

NordicTrack
A CML COMPANY

should forget that it was within a definitely spheroidal head that the amusing and instructive relations of so many living cylinders were arrayed.

THE MAP CATALOG: EVERY KIND OF MAP AND CHART ON EARTH AND EVEN SOME ABOVE IT, Joel Makower, Editor, Laura Bergheim, Associate Editor. Vintage Books, Random House, Inc. (paperbound, $14.95).

A serious cyclist? You certainly enjoy well-mapped bike routes. A half-dozen sources issue maps, some on waterproof paper with ink that will not run, for 10-mile jaunts or coast-to-coast across a dozen sheets, even through the inviting viniferous provinces of France, from Bordeaux to Côtes du Rhône. Old-map buff? "For reasons no one can explain" it is the National Oceanic and Atmospheric Administration that now maintains and supplies reproductions of the fine old work of the cartographers of the Confederate Army.



A Defense Mapping Agency map and a* Landsat *image of Cape Cod

Your own wall can sport a map prepared by the CIA. Timely and well-informed overt operations by the Company's expert cartographers have brought forth a variety of clear new maps of many countries, plus close-ups such as a map of the Strait of Hormuz, the world geography of narcotics or a pocket street guide for Leningrad. The space-borne cameras of Landsat, and nowadays also the higher-resolution scans of its commercial French competitor SPOT as well, image the world for consumers. SPOT's pictures are spectacularly detailed. They are correspondingly pricey: a color print 10 inches square at SPOT's top-of-the-line quality runs close to $800.

Maps old and new of the lands of the earth are described and explained, arranged in some 40 categories. Sources are listed in detail, both governmental and commercial; what is available from each is briefly sketched. The richly useful mapping of the elements of air and water is not omitted. Those intricate aeronautical charts, with their worldwide entries of the network of landing grounds, navigational aids and local regulation of flight, are summarized by type. Sailors' charts of the watery world, riverways, deep blue sea and the coastal waters between, with their tides and currents, are also included. Astronomical maps are listed, maps of stars and of planetary surfaces, although not the massive surveys and catalogues, old and new, of professional astronomers.

Accounts of what globes and relief maps are to be found carry the reader some distance into the third dimension. The forest of software—maps that are not on paper but in pulse—is visited as well, although it is growing so luxuriantly that no book can be expected to provide more than initial help; the capacious new video-disk format CD-ROM is not here at all. Lists of sources for official maps—Federal agencies, all the states and some 100 overseas—and of U.S. and Canadian map libraries and map stores complete the references.

The authors have reproduced a black-and-white map excerpt or two for nearly every category. Their book offers a sample of a dozen colorful maps: the eastern U.S. with education and income correlation by census area, the World War II battle for Cherbourg, a weather-satellite map of air temperatures centered on the North Pole.... Only one category is missed: those fictional maps, from Treasure Island to the *New Yorker*'s U.S. The

catalogue contains serious pleasures both for browsers and for those who will act on its ample guidance.

COMBAT FLEETS OF THE WORLD 1988/89: THEIR SHIPS, AIRCRAFT, AND ARMAMENT, edited by Jean Labayle Couhat and Bernard Prézelin. Naval Institute Press, U.S. Naval Institute, Annapolis, Md. 21402 ($96.95).

During the 1980's combat at sea has more and more made the news of war. It is timely to pay attention to this hefty and comprehensive reference. Its topic is whatever forces are afloat out there, an annotated accounting of just about all the navies, coast guards and paranaval organizations in the world. Not at all a newcomer to a watery taxonomy long dominated by London, this well-established series from Paris has recently been adapted for readers of English, with new editions every couple of years at a reasonable price.

More than 150 countries are included—down to minimal fleets such as that of the Turks and Caicos Islands, which operates a single 40-foot fiberglass patrol craft for police work—but almost a third of the book is needed to treat the two superpowers. The capital ships that currently take grim pride of place are the fleet ballistic-missile submarines, doubly nuclear (both in armament and in propulsion) and now found in five navies.

Good close-ups show the world's largest, the Soviet Typhoons. Eight of them are in service or still building; those double pressure hulls and massive "sails" are "evidently intended to operate beneath the Arctic ice pack, breaking through to launch." In harbor they dwell sheltered in deep pens cut into the living rock on the Kola Peninsula, able to flee to hide and fire with only half a day's voyage from home. Their cousins, small Alfa-class nuclear attack subs, are the fastest and deepest-diving of all, rated at 45 knots submerged and able to dive below 700 meters. They have titanium pressure hulls powered by compact reactors cooled by a dense fluid mixture of lead and bismuth. These half-dozen high-performance Alfas have lower standards of safety and maintenance than would be acceptable in the West.

Enough of Leviathan. Consider the five French minehunters of the Circé class. They are silent and are built of nonmagnetic plywood; special rudders have small propellers of their own, so that the ships can turn on a half-franc piece even at low speed. They carry a team of divers to destroy deep mines, and they mount a wire-guided underwater sled able to bring its video eye close to a mine and lay alongside it 100 kilograms of high explosive to be set off by ultrasonic pulse once the sled has drawn away. German and British versions are to be seen as well; these ships nourish a brood of drones—small boats or rafts under radio control, able to seek out mines and mark cleared paths as they go. Seamen are less at risk.

Two small Swedish submarines appear to fly. It is only that these new diesel subs, steel cylinders of 1,000 tons, were launched by simply lifting them bodily into the waters at Malmö with the giant cranes at the Kockums shipyard; the photo opportunity was irresistible. The next round aims at a new closed-cycle underwater power scheme: fuel burned with liquid oxygen and waste gases condensed.

The U.S. Navy has this year launched the first of a set of "remarkably stable" small ships. They are meant to stay at sea for three months even under the impossible conditions of high-latitude winter gales. Out there amidst the high combers they will keep on trawling for information. They tow a couple of kilometers of plastic cable fitted with many a hydrophone, and in real time they relay (by satellite) all they hear to analytical computers ashore. These ships—the prototype, T-AGOS 19, is named *Victorious*—are catamarans, their hull arching high between twin supporting plates held on two well-submerged pontoons, each pontoon rather resembling a huge torpedo 200 feet long. There is nothing much near the waterline for a towering wave to encounter.

A Soviet candidate for advanced design is the new Pomornik class of amphibious landing craft. These are the largest air-cushion warships in the world. Able to carry a few tanks or a couple of hundred troops, they hover over a flat surface on a powerful air draft while they speed along at 55 knots, driven by three big ducted airscrews mounted on gas turbines atop tall deck pylons. The Everglades airboat is here written very large, too big to be carried on shipboard. They must be intended for "short-range ... assault operations" on their own.

Navies have the most varied needs. You can find tankers for gasoline, for missile fuel and for fluid radioactive waste, laundry ships, floating cranes, docks, workshops and barracks, degaussing ships and barges full of batteries, humble ships with spinning drums to pick up oil spills and grand ships able to plot the flight trajectory of an ICBM. Here you can see a photograph of a fast "cigarette boat" run by the U.S. Coast Guard in the drug wars of the Florida coast against many a smuggler of the same type. And here is one of the 45-knot aluminum "Boghammar boats," named for its builders in Stockholm. Manned by Iranian Revolutionary Guards armed with hand-held weapons, about 40 of these warlike speedboats have harassed unarmed merchant ships in the past few years; a few were lost in unequal combat with U.S. Navy ships and aircraft.

The special quality of this volume is its depth of technical detail, including specifications and striking close-ups of aircraft, weapons and fittings. There are more than 3,600 photographs (look right down the stacks of the missile cruiser *Slava*) and for many larger ships there are line drawings, carefully labeled to enable readers to interpret the bristly intricacies on modern masts, bridges and decks. Yet the clubby ambience is missing: here are no advertisements for new ships or radars, few silhouettes, no tables of comparisons and no long lists of senior officers. (The only index lists about 10,000 ships by name.) The Gallic version is rational, businesslike and fascinating. The American editorial team has done a first-rate job, free of awkwardness or material error.

It seems probable that the world's costly navies—and the volumes that describe them—have grown to somewhere near their maximum size during these years. Ahead lie lower budgets, smaller ships and gradually dwindling fleets.

THE AMERICAN CENSUS: A SOCIAL HISTORY, by Margo J. Anderson. Yale University Press ($30).

Faithful to the Enlightenment, the Founding Fathers trusted in reason informed by fact. They protected the truth by prudent checks on the excesses of class and of self-interest. One simple paragraph of the prudent Constitution laid out the numerical root of congressional power: "Representatives and direct Taxes shall be apportioned among the several States ... according to their respective Numbers ... of free Persons." The very next clause had to record the shameful compromise then found essential to nationhood, assigning to "all other Persons" a weight of three-fifths and "excluding Indians not taxed." "The actual Enumeration shall be made ... every subsequent term of ten Years, in such Manner" as Congress

SCIENTIFIC AMERICAN

CORRESPONDENCE

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

Photocopying rights are hereby granted by Scientific American, Inc., to libraries and others registered with the Copyright Clearance Center (CCC) to photocopy articles in this issue of SCIENTIFIC AMERICAN for the flat fee of $1.25 per copy of each article or any part thereof. Such clearance does not extend to the photocopying of articles for promotion or other commercial purposes. Correspondence and payment should be addressed to Copyright Clearance Center, Inc., 21 Congress Street, Salem, Mass. 01970. Specify CCC Reference Number ISSN 0036-8733/89. $1.25 + 0.00.

Editorial correspondence should be addressed to The Editors, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. Manuscripts are submitted at the authors' risk and will not be returned unless accompanied by postage.

Advertising correspondence should be addressed to Advertising Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017.

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

Name

New Address

Street

City

State and ZIP

Old Address

Street

City

State and ZIP

"shall by Law direct." No other nation had before instituted a regular census with so much bite.

Professor Anderson, historian at the Milwaukee campus of the University of Wisconsin, has presented her lucid and quietly revealing account in good time for next year's bicentennial census. She shows plainly how our census has played a pervasive triple role: as a tool for the apportionment of power, as a rational baseline measure for much of American social science and as a theater of policy debate. For generation after generation those fateful census categories have formed the framework of social issues.

The census was clearly national; in the all but sovereign states of 1790 it was a rare form of direct action by Federal officials. It was plain that the nation would grow, and the periodic census—syncopated with the rhythm of elections—was aimed at adjusting for the expected growth. The first census conflict had a mathematical origin, one arising out of a similar contradiction that lies within every calendar of days and years. Days are stubbornly integers, counted by sunrises, but the year, an orbital time, answers to no commensurate count of days; there is always something left over. Congressional representatives too are to be counted only by whole numbers, but among hundreds of legislators elected by a few dozen states a strict proportionality to population is impossible. Some rounding off is necessary, and it requires two interrelated choices: the whole-number size of the House and a scheme for dealing with the fractional representatives implied by state-by-state populations.

The scheme for apportionment that first carried the new Congress in 1792 did not become law; it was rejected by the first of all presidential vetoes and the votes were not on hand to override President Washington. A proposal of Jefferson's finally became law. First you set the size of the House, then you give each state its whole-number proportion, dropping all fractional remainders. It sounds simple, even under the documented circumstance that Congress in those days did not know much arithmetic. Actually, treating the remainders—the fractional representatives—remains subtle and full of paradox. No solution has been found to satisfy all plausible consequences of one head, one vote.

The present scheme, survivor of four or five methods adopted in the past, was devised in the 1920's and fixed in law only after 1940. First the size of the House is set and then the seats are allocated among states "so that no transfer of any one seat can reduce the percentage difference in representation between those states." The scheme slightly favors the smaller states, in the few cases where the fraction of a seat at issue is between .414 and .5. Dropping all the remainders would favor the larger states.

Decade by decade the head count grew steadily, doubling every 25 years across the 19th century; the census questions became more complex; new issues arose. In 1840 an 80-column form was designed to help the census takers. The arrangement of the columns was confusing, in the sense that the entry places were not clearly marked as "white" or "colored." The heading was not omitted, but it was curiously placed; it was easy to wander to the wrong row.

In the parlance of that time, every senile person was entered among the "insane and idiots." A small number of the forms were marked in error—a few whites in the colored box, and the reverse. Since the error rate was small, that made no significant difference wherever the population was substantial, just a little noise. But the smallest census category tallied was that of colored people in free Northern states. A rather few wrong entries increased the indicated fraction of "idiots and insane" among free blacks by a factor of 10 over the well-measured rate under Southern slavery. That count was in effect all noise. Secretary of State John C. Calhoun, who held control over the census, "found the data too useful to discredit." The report was printed as it was gathered; historical consensus recognizes now that the data were wrong.

Knowledge became power. The five volumes of the census of 1860 supplied essential intelligence; the Census Office made maps of the number of whites, free blacks, slaves, livestock and crops county by county and sent them to "Generals in the field and military governors." General Sherman himself wrote, "No military expedition was ever based on sounder or surer data." A map of the entire Confederacy, reproduced here from the 1860 report, codes the slave population county by county; it is one of the first such presentations of the rich census data for the public. Between 1870 and 1900 the country industrialized, the frontier vanished and the immigrants poured in. The census became a guide to a changing people. The Census Office was made a permanent statistical agency in 1902.

The 1920's saw the Census Bureau

immersed in a disgraceful tribalism amidst a wave of pseudoscientific claims. The upshot was the end of open immigration and the establishment in 1924 of quotas by national origin, aimed at the newcomers from eastern and southern Europe. (That "racialist law" lasted until 1965.) The 1940 census shared the New Deal aim of socioeconomic equity; the census would thereafter seek to gauge, place by American place, the real standard of living.

By 1970 the census faced a serious undercount; grants, not taxes, were mainly at stake. Inexpensive enumeration by mail, particularly in big cities, clearly led to differential undercount of the poor and members of minorities. Five million people were missed, the bureau estimated in 1973, with a minority percentage several times that of the total population. By 1980 there were no fewer than 54 suits filed in Federal court against the bureau; it would lose some important ones, only to appeal. The 1980 data were improved over 1970, although at high cost, and the bureau declined to adjust its final figures.

The census of the U.S. has had major impact on the theory and practice of social statistics. Herman Hollerith, the inventor of the punched-card tabulator, was first encouraged and then employed by the leaders of the Census Bureau. He patented, made and then leased his machines for the census of 1890. (*Scientific American* pictured the patented system on its front page in August of that year.) Electrical tabulation was a great success. During the 1940's the bureau and its consultants developed in all detail the calculation of sampling error without the unrealistic assumptions of a normally distributed and infinite population. By 1953 their two-volume summary of methods was published; it became a standard text. By now the digital computer and the elaboration of stratified sampling have together created a "demographics industry" based on widespread private sample research as well as on the modern census, remarkably powerful in both marketing and politics.

The bicentennial census has already called forth its own lawsuits. Will the coming count include all who live here—citizens, legal and illegal aliens, poor and rich, homeless and settled alike? Will the 101st Congress pay for a fair and complete count? (The 1980 census cost $5 a head.) Will a more sophisticated sampling technique someday replace, or at least uniformly test, the direct count?

ESSAY

Fraud in science: causes and remedies

by Arnold S. Relman

We keep hearing about fraud in biomedical science. A number of widely publicized cases have generated comment in the media, studies and reports by scientific and academic organizations and even congressional hearings.

How widespread is misconduct among research scientists? Has its incidence—relative to the number of working scientists—really increased recently? Much depends on how one defines "misconduct." If the term is applied to all types of poor practices, such as carelessness or bias in the conduct or recording of experiments, incomplete reporting of results and minor mishandling of data, misconduct must be common. But if only deliberate, outright fraud is considered, the incidence is probably low. Fewer than a score of such cases have come to light in the past 10 years. Whatever one's definition, it is important to recognize that honest error is not a form of misconduct. Error is inevitable in all human endeavors; it must not be regarded in the same light as deliberate malfeasance.

The recent increase in reported cases of misconduct could be indicative of a trend or might simply reflect increased attention to a long-standing problem. In either case, there is no evidence to support the notion propagated by some critics in the media that there has been an alarming and widespread deterioration of ethical standards in science. Yet there is a problem and it needs attention.

There are two theories of the origin of scientific misconduct. One of them blames human nature and the other blames the circumstances in which modern research scientists work. The first theory simply says that scientists are human, like everyone else; a few of them are bound to commit major sins and more will succumb to the temptation to cut corners. The other theory attributes misconduct primarily to inadequate supervision of young research workers, pressures to publish and the increasing competition for funds, appointments, promotions and peer recognition.

Many senior members of the biomedical research establishment until recently held the first view. They tended to think the problem had been greatly exaggerated by media attention and that little or nothing could be or needed to be done about it. That conviction has softened under the heat of recent publicity, and now most leaders in science give credence to the other point of view, which attributes misconduct more to the milieu of research than to human frailties. This view—unlike the first—suggests both a need and a rationale for preventive measures. Many institutions are now reexamining their policies with the object of improving the supervision of junior workers, instilling a better sense of scientific values and lessening competitive pressures to publish hastily and excessively.

In my opinion there is little doubt that the "publish or perish" attitude has been a major cause of the problem. Nowadays success in a scientific career depends so heavily on number of papers that some overly ambitious investigators come to perceive frequent publication as an end in itself, without sufficient regard for the quality and integrity of their research. This leads them to prepare reports for publication that may be premature, shoddy or even fraudulent. Their superiors or senior collaborators may be too busy or too distant from the work to be aware of its defects and yet may lend their names as coauthors.

Many people assume that the peer-review process at most reputable scientific journals should be able to weed out fraudulent manuscripts before publication, but this is not so. Expert referees usually recognize work of poor quality, but they are not likely to identify fraudulent work that is internally consistent. The reason is that science is based on a certain kind of trust in one's colleagues. When a scientist reports that he has done certain things in his laboratory and has made certain observations, his colleagues have no acceptable alternative but to believe him—unless, of course, his claims defy reason or violate well-established basic principles. His colleagues may challenge his methods and experimental design and his analysis and interpretation of results, but they cannot challenge what he says *he did* without calling everything else into question and making collegial discourse almost impossible.

Some critics have suggested audits of raw data or inspections of laboratories as ways of detecting fraud or resolving doubts about the integrity of research. It has been proposed that teams of investigators appointed by journal editors or by Government agencies might be available to conduct random inspections on a regular basis or to be called into action when the integrity of a study is questioned. That is a bad idea—and unworkable too. The red tape and the suspicion and controversy inevitably accompanying such a practice would poison the atmosphere and cast a chill over scientific discussions, and it would also add considerable expense at a time when research support is already being curtailed by budgetary constraints. Moreover, such inspections are unlikely to be effective. As a practical matter any audits would have to be confined largely to inspections of laboratory notebooks—and notebooks can easily be doctored.

Only direct observation by co-workers in the laboratory is likely to catch the miscreant who makes notes on experiments never done or falsifies the observations he records. Clearly it is the responsibility of local colleagues, and in particular of coauthors and supervisors, to recognize unethical behavior in the laboratory and to ensure that the data reported at scientific meetings and in manuscripts submitted for publication are obtained honestly.

That is why I believe the National Institutes of Health and other agencies funding research must look primarily to their grantee institutions for the solution to this problem. As a condition for receiving grants, institutions should be required not only to establish policies and procedures for the prompt reporting and rigorous investigation of cases of suspected misconduct (as is now the case) but also to show that they have programs designed to promote good research practices, to diminish the pressures predisposing to misconduct and to identify malfeasance when it occurs. Editors of scientific journals can also help by asking that all coauthors of submitted manuscripts accept full responsibility for the integrity of the studies they are reporting. Misconduct in science is a public concern, but its causes and cure reside in the scientific community.

ARNOLD S. RELMAN is the editor-in-chief of the *New England Journal of Medicine*.

HELPING COMPUTER PROGRAMMERS MAKE FAST AND STEADY PROGRESS.

Until now, manual skills were a big part of writing complicated computer programs. They required a major investment in time, effort and money—and days of boredom for the applications designer.

McDonnell Douglas information systems specialists bypass the dull, turtle-speed tedium formerly necessary for good programs. Their Pro-IV® software maps out the best program path and selects routine computer instructions for them.

The result is a working program that emerges with hare-like speed, ready to use, allowing application systems designers to fully concentrate on the special characteristics of the project.

Designers like Pro-IV software because it lets them try out ideas in prototype programs, quickly and easily, selecting only the best. Managements like it because it's faster, cheaper and yields better results.

For more information, write: Software, McDonnell Douglas, Box 14526, St. Louis, MO 63178

MCDONNELL DOUGLAS

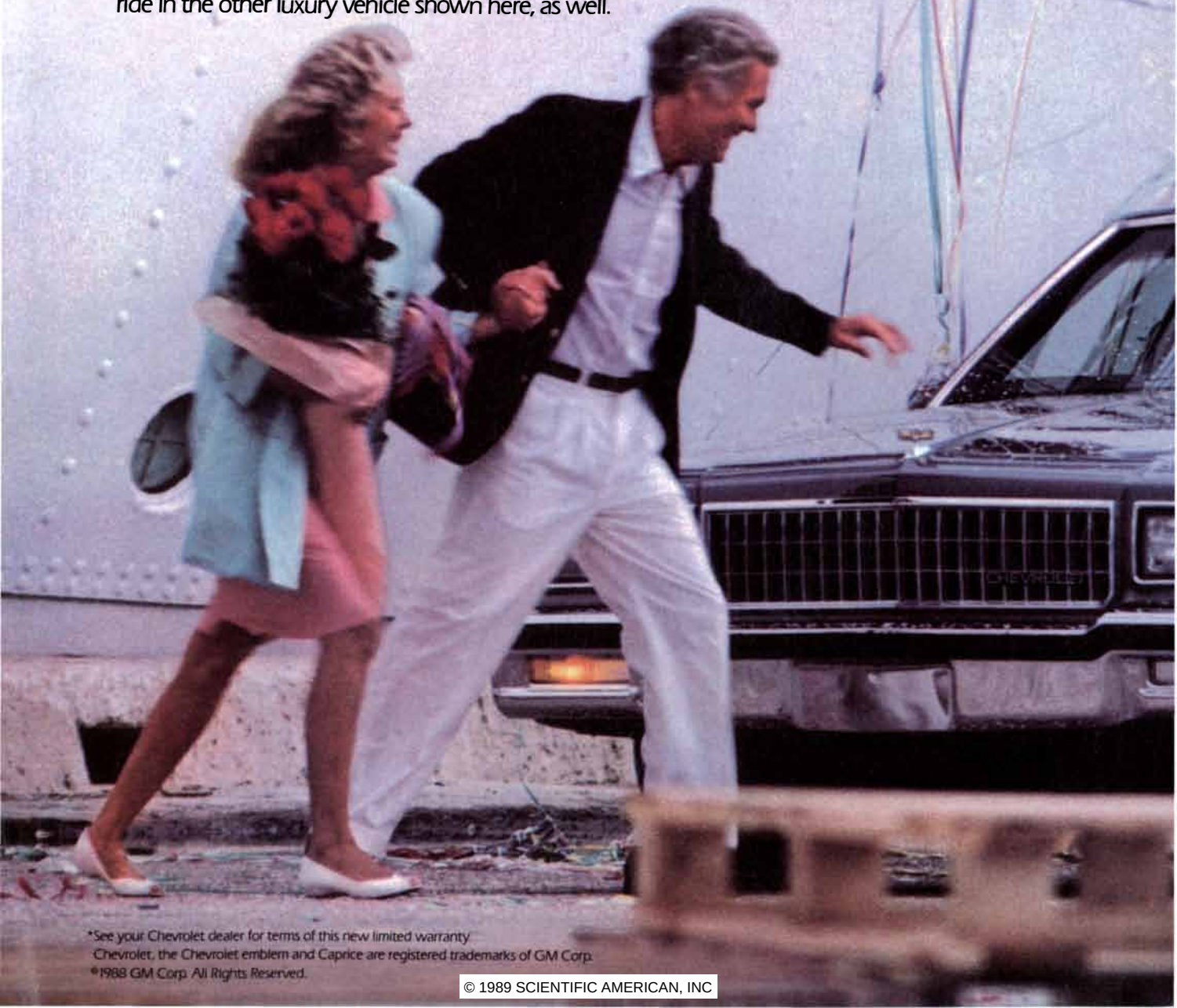
INFORMATION SYSTEMS *MILITARY & COMMERCIAL AIRCRAFT* *SPACECRAFT & MISSILES* *TRAVEL MANAGEMENT* *HELICOPTERS* *FINANCING*

Caprice Classic Brougham. Now you can comfortably afford to ride in the lap of luxury.

- Standard V8 with Electronic Fuel Injection.
- 4-speed automatic overdrive transmission.
- Standard AM/FM stereo sound system.
- Available genuine leather seating areas for six.
- New 3-year/50,000-mile Bumper to Bumper Plus Warranty. *

THE Heartbeat OF AMERICA TODAY'S CHEVROLET™

You see here two different ways to travel in the lap of luxury. The one in front is ours. Caprice Classic Brougham is so wondrously smooth, you sometimes forget you're moving. So comfortable, you'll never want it to stop. And so affordable, you'll probably be able to take a ride in the other luxury vehicle shown here, as well.



*See your Chevrolet dealer for terms of this new limited warranty.
Chevrolet, the Chevrolet emblem and Caprice are registered trademarks of GM Corp.
©1988 GM Corp. All Rights Reserved.

DEWAR'S PROFILE:

DAN RIZZIE

HOME: Dallas, Texas.

AGE: 37.

PROFESSION: Artist.

HOBBY: Cooking and eating. "I go to Italy every summer but not for art's sake; I go for the food. I've got my priorities straight."

LAST BOOK READ: *The De-Definition of Art*, Harold Rosenberg.

LATEST ACCOMPLISHMENT: Four one-man shows in the past year. In New Orleans, Dallas, LA and New York.

WHY I DO WHAT I DO: "More than anything else because I can't imagine not doing it."

QUOTE: "If I keep pressing, I'll get a really strange art form that's all mine. Which will either be very interesting. Or unbearable. You tell me."

PROFILE: Witty and open-minded. "Artistic types are supposed to live in New York, right? Wrong. Living in Dallas lets me do two things: Work. And leave the infighting to the experts."

HIS SCOTCH: Dewar's® "White Label®" and water. "Cocktail hour is my favorite hour of the day. Really. You can look it up."