SCIENTIFIC AMERICAN

NOVEMBER 1988
$2.95

How to make migrating tumor cells into easy targets.

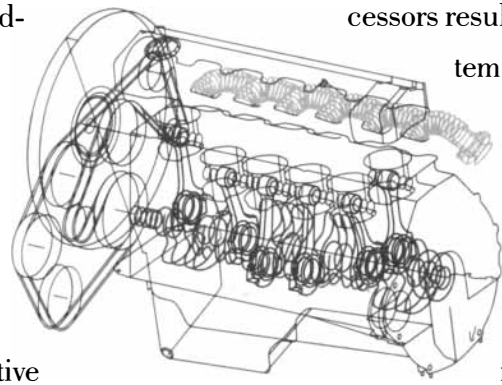
High-transparency materials for optical fibers.

Human walking: three million years of being upright.

Powerful puzzle: *a lightning flash discharges several hundred megavolts. Its source has defied physics for two centuries.*

THE MERCEDES-BENZ 300 CLASS: ITS TECHNOLOGY MAY BOGGLE THE MIND, BUT THE SOUL UNDERSTANDS IMMEDIATELY.

The engineers of Mercedes-Benz have long been noted for their uncanny ability to turn remote technical exercises into rewarding automotive experiences.



With the 300 Class, Mercedes-Benz engineers have done nothing less than turn the theoretical into the phenomenal. Distilling esoteric technology into what one automotive journal calls "a mechanical symphony." Creating an automobile of sublime driving pleasure.

An automobile whose electro-mechanically controlled six-cylinder engine produces robust power—177 horsepower in three-liter form. Power so smooth that even at test-track velocities, the engine's exertions can barely be felt.

Other advanced engineering concepts produce further refinements of the classic Mercedes-Benz driving character. A multilink independent rear suspension system preserves near-perfect geometry between tire and road, resulting in a ride that *Car and Driver* terms "nothing short of magical." While the same journal reports handling prowess so inspiring that "you can charge unfamiliar territory as though you were on your daily route to work."

Advanced aerodynamics have shaped a 300 Class sedan body that cuts drag to a slippery 0.31 Cd—and wind noise at cruising speed to near silence. The advanced application of microprocessors results in an Anti-lock Braking System that modulates braking pressure up to 15 times per second in sudden hard stops, preventing skidding and preserving steering control.

And advanced technology addresses the concern for occupant safety: microprocessors in the Supplemental Restraint System are primed to deploy a driver's-side air bag and front seat belt emergency tensioning retractors within *milliseconds* of a major frontal impact.

In their pursuit of a superbly capable automobile, the engineers of Mercedes-Benz have created one of the most sophisticated production cars ever built. But you need not be an automotive engineer to appreciate their achievement. You need only drive any one of the 300 Class automobiles: 260E and 300E Sedans, 300CE Coupe and 300TE Station Wagon. You—and your soul—will understand immediately.

Engineered like no other car in the world

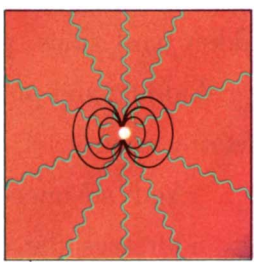
SCIENTIFIC AMERICAN

November 1988 Volume 259 Number 5

42 

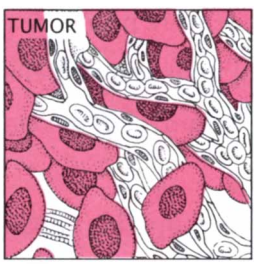
Educating Poor Minority Children
James P. Comer

The inadequate education of black children from urban ghettos condemns many of them to lives of chronic unemployment and desperation. Is there a way to bring such children into the mainstream of a technological society? The author built on his own childhood experiences to develop a strategy that is now working in 50 schools around the country.

50 

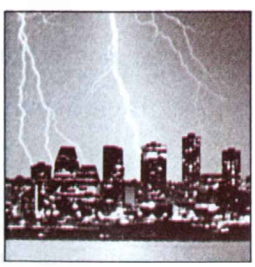
Quasi-periodic Oscillations in Celestial X-Ray Sources
Michiel van der Klis

Regular variations in the intensity of X-ray emanations can yield information about a celestial X-ray source: it may be orbited by an eclipsing star, for example, or may rotate like a lighthouse beam. What is one to make of powerful sources near the center of our galaxy whose X-ray intensity fluctuates not quite periodically but not quite at random?

60 

What Makes a Tumor Cell Metastatic?
Michael Feldman and Lea Eisenbach

The tumor cells that undergo metastasis, traveling through the bloodstream to establish new tumors, turn out to bear on their surface a distinctive pattern of molecules that enable them to evade the immune system. This feature and others that are unique to metastatic cells point the way toward what may one day be a new kind of cancer therapy.

88 

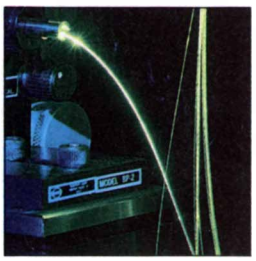
The Electrification of Thunderstorms
Earle R. Williams

An average thunderstorm generates several lightning flashes per minute and as much electric power as a small nuclear plant. To produce such large amounts of power a thundercloud must be charged to potentials of several hundred million volts. Models have been developed that explain aspects of the charging process, but the fundamental physics is still a mystery.

100 

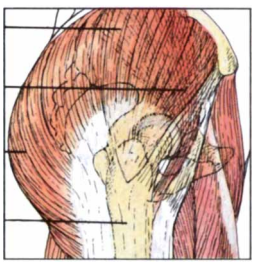
Superfluid Turbulence
Russell J. Donnelly

Liquid helium cooled to within about two degrees of absolute zero can flow without viscosity or friction, but it is not immune to a form of turbulence that is quantum-mechanical in nature. For example, the rates at which superfluid atoms circulate in a vortex are precisely determined by the same equation that fixes the radii of electron orbits in an atom.

110 

Infrared Optical Fibers
Martin G. Drexhage and Cornelius T. Moynihan

Optical fibers made of highly transparent silica glass can carry voice and data signals for as much as 50 kilometers without reamplification. New materials that transmit longer wavelengths of light can in theory be 20 times as transparent as silica glass. They may eventually be able to carry signals for hundreds or perhaps thousands of miles.

118 

Evolution of Human Walking
C. Owen Lovejoy

When did our ancestors begin to walk on two legs, and what anatomical changes made it possible? Analysis of the pelvis of Lucy, a three-million-year-old hominid, reveals that upright walking had been perfected by her time. The author thinks bipedality developed early because it enabled males to nurture a family: it freed their hands to carry food.

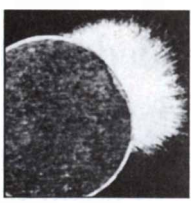
126 

Obstacles to Developing Vaccines for the Third World
Anthony Robbins and Phyllis Freeman

The developing nations urgently need new vaccines to control serious infectious diseases. Yet many of the vaccines—in particular those for diseases that are largely confined to the Third World—are unlikely to be produced if the decision is left to pharmaceutical companies. Neither the developing nations nor the United Nations can pay market prices.

DEPARTMENTS

8 Letters

12 

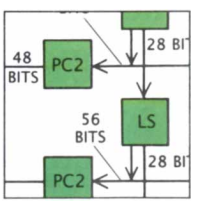
50 and 100 Years Ago

1888: What wou an eclipse of the moon lo k like to a person on the moon?

17 Science and the Citizen

134 Science and Business

138 The Amateur Scientist

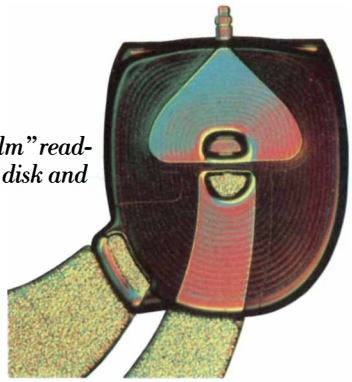
142 

Computer Recreations

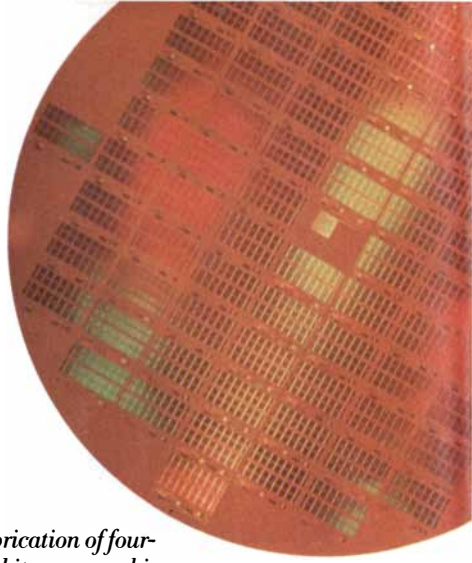
Computers encode messages by embedding a private key in a public problem.

146 Books

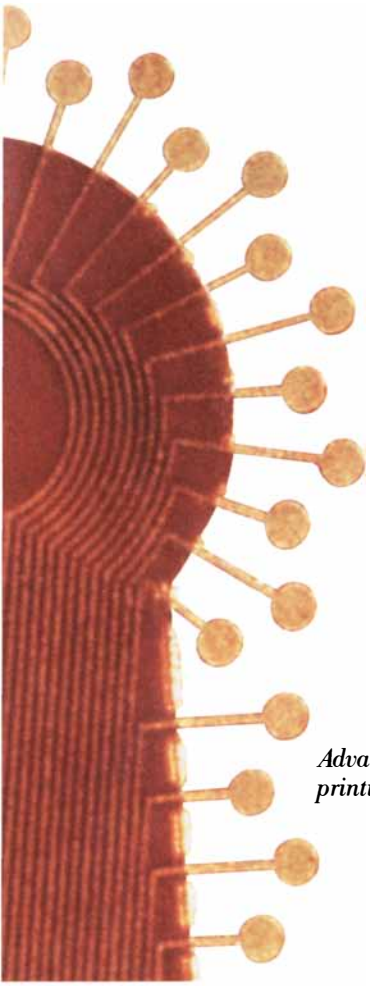
150 Essay: *Robert A. Weinberg*



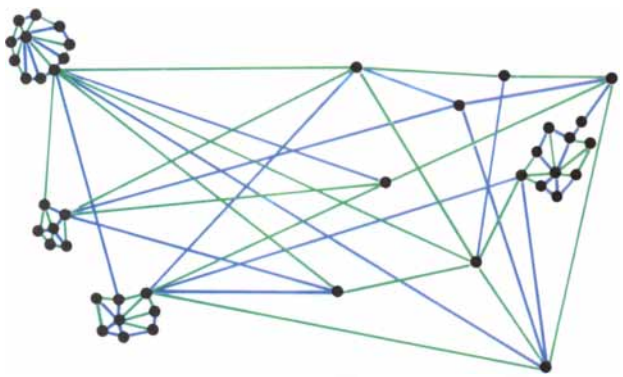
Invention of "thin-film" read-write technology for disk and tape storage.



First fabrication of four-million-bit memory chips.



Advanced "resistive ribbon" printing technology.



Most advanced, widely used architecture for computer networks.

"What does all this do for me?"

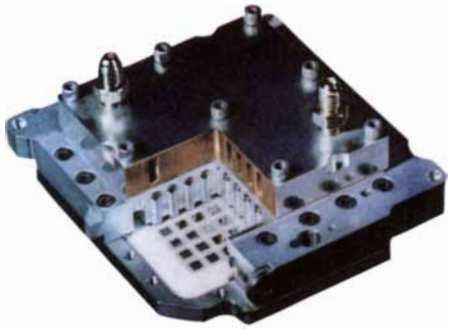
Our customers are practical people. They believe in results. And so do we.

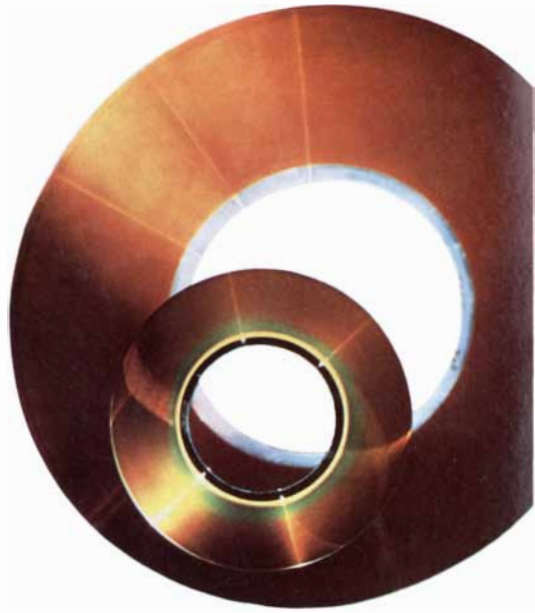
That's why IBM technology, like that pictured here, is intended to offer solutions to our customers.

For example, our customers can file billions of characters of information and retrieve three million characters per second, for a fraction of what it cost ten years ago—thanks to IBM's leadership in disk storage systems.

Customers using relational data bases, invented by IBM, can update and retrieve data simply by telling their computers what they want to do, not how to do it.

Device with the densest "packaging" of high-performance computer logic.





Invention of most major magnetic disk technologies.

Relational data base—the most convenient way to file, modify and retrieve computer information.

And people will benefit from our pioneering development of the four-million-bit memory chip.

It holds the equivalent of 400 pages of double-spaced typewritten text, and "reads" it all in one quarter of a second.

Technology like this is helping customers of all sizes—in industry, science, government and education—solve problems.

Because IBM is committed to leadership in state-of-the-art technology, our customers can depend on us to offer more state-of-the-art solutions. That's IBM technology at work.

IBM®

THE COVER photograph shows a cloud-to-ground lightning strike over a mountain near Tucson, Ariz. Lightning bolts typically represent potential differences of hundreds of millions of volts and currents of up to 10 kiloamperes (see "The Electrification of Thunderstorms," by Earle R. Williams, page 88). The large scale of this strike is made evident by the high-tension towers at the base of the mountain. The strike was so powerful that the mountainside appears to be covered by snow.

THE ILLUSTRATIONS

Cover photograph by Tom Ives

Page	Source
43	S. Varnedoe
44-45	Laurie Grace
46	S. Varnedoe
47	Laurie Grace
51	George Retseck
52	Andrew Christie
53	Michiel van der Klis
54	George Retseck
56	Andrew Christie
57	George Retseck
61	Daniel Zagury, University of Paris
62-85	Patricia J. Wynne
88-89	Greg Gilbert, *The Seattle Times*
90-94	Joe Lertola
96	Gordon Gerber (*top*); Richard E. Orville, State University of New York at Albany (*bottom*)
97	Los Alamos National Laboratory (*top*), Earle R. Williams (*bottom*)
98	Joe Lertola
99	National Aeronautics and Space Administration
101	Klaus W. Schwarz, Thomas J. Watson Research Center
102-104	George V. Kelvin
105	Russell J. Donnelly
106-108	George V. Kelvin
111	GTE Laboratories, Inc.
112-114	Hank Iken
116	James A. Harrington, Hughes Research Laboratories
119	Larry Rubens, Kent State University
120-123	Carol Donner
124	Carol Donner (*top*), C. Owen Lovejoy (*bottom*)
125	Carol Donner
127	Sean Sprague, United Nations Children's Fund
128	Johnny J. Johnson
130	Edward Bell
132	U.S. Centers for Disease Control
138-141	Michael Goodman
142-145	Andrew Christie
147	Graphische Sammlung Albertina, Vienna

ScientificAmerican (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). 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; Ari W. Epstein; Gregory R. Greenwell; John Horgan; June Kinoshita; Philip Morrison, *Book Editor;* Tony Rothman; Ricki L. Rusting; Russell Ruthen; 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 Moise; Dorothy R. Patterson

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; Nancy Mongelli; 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; Rob Gregory, *Advertising Manager;* Kate Dobson, *Advertising Manager;* Lisa Carden. CHICAGO: 333 N. Michigan Avenue, Chicago, IL 60601; Patrick Bachler, *Advertising Manager;* Litt Clark. DETROIT: 3000 Town Center, Suite 1435, Southfield, MI 48075; Bill Moore, *Advertising Manager.* ATLANTA: Reese & Associates. 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/INTERNATIONAL: Peter B. Kennedy

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

LETTERS

To the Editors:

The informative article "Capacitors," by Donald M. Trotter, Jr. [SCIENTIFIC AMERICAN, July], opening with a discussion of the 250-year-old Leyden jar, evoked nostalgia in this reader. Not that I am quite coeval with Leyden jars, but I well recall, as a boy in the first decade of this century, learning about them from an old book on "natural philosophy" found in an abandoned house and then making one from a widemouthed pickle bottle with interior and exterior electrodes of lead foil scrounged from a florist and, in the absence of a large rubber stopper, a plug carved from wood and boiled in paraffin wax. With a central brass rod tipped with a brass ball liberated from my mother's brass bed, I produced a quite serviceable Leyden jar that could store enough charge to deliver a tooth-rattling shock.

A little later an interest in wireless telegraphy, as radio was then called, descended on me. This required capacitors, then referred to as condensers, for both sending and receiving. We made condensers for receiving, able to handle low voltages, from kitchen waxed paper interleaved with foil. Those for sending had to withstand the 10,000 or 12,000 volts from an induction coil or a transformer, and those we constructed from eight-by-10-inch glass photographic plates discarded by photographers, from which we washed the gelatin emulsion with lye. There might be 10 or more glass plates, interleaved with foil electrodes, and the assembly was immersed in transformer oil supplied by a kindly employee of the Georgia Power Company. These capacitors worked, and worked well, giving us a range, on a good night, of several hundred miles. Here was simple but effective science, from which we learned. How much today's radio hams have missed by being able to buy their glossy equipment ready to plug in!

How far we have come since then in capacitor technology is shown by Dr. Trotter. I would have been interested, though, to read something about the voltage at which these miniaturized capacitors break down and about how the high-voltage capacitors used in power-distribution systems and nuclear research are made.

MAX HERZOG

Augusta, Ga.

To the Editors:

In "Computers in Music" [SCIENTIFIC AMERICAN, April] Pierre Boulez and Andrew Gerzso give an interesting description of how Boulez processes instrumental sounds by computer in his piece *Répons*. The description is framed, however, in a polemical insistence that musical applications of computers require large, general-purpose and highly flexible devices, such as the 4X and Matrix 32 developed at IRCAM. I find this approach to the use of computers in music anachronistic at best.

At the opening of the article the authors assert that computers provide "an entirely new means of musical expression" with musical potential "bounded only by imagination in creating an 'orchestra' of sounds." And they suggest that the combination of computer-generated sounds and the sounds of conventional instruments requires special musical ideas that must be "readily translatable into both mediums.... Otherwise the listener might wonder what role the computer was meant to have in relation to the other instruments and be puzzled ... by the lack of coherence."

The article closes with the statement that "no composer ... can be satisfied with a device that allows the study of only one method for analyzing, synthesizing or transforming sound." The authors end by lamenting the recent trend toward "specialized devices, each of which has its own method for processing digital signals," because trying to link such devices "inevitably results in problems of control and coordination."

Although the computer has certainly proved to be a remarkable and even revolutionary musical instrument, the claim that it is "an entirely new means of musical expression" recalls the claims made on behalf of the first computer synthesizers in the 1950's and the analog synthesizers of the 1960's. Certainly musical technology is advancing, but it seems to me that musical expression depends more on the people who make the music than on the technology they use. The contention that the use of computers with conventional musical instruments demands extraordinary musical concepts or extraordinarily fast, large and hence expensive general-purpose machines flies in the face of what is happening in the musical world today.

What is happening, briefly, is that the "specialized devices" the authors disparage have brought about a genuine musical revolution. These devices, which include synthesizers, samplers, effects generators and mixers, have existed in various forms for more than 20 years. In the past few years, however, an industrial standard has been adopted that allows all these devices to communicate with one another and with personal computers by means of an interface called MIDI (Musical Instrument Digital Interface).

Thus the musical use of computers has finally been passed from research centers such as IRCAM to a broad community of people who are primarily interested in making music. Teenagers playing in bands in their garage have found solutions to "problems of control-and coordination," and professional composers and performers are now able to put together personal-computer music systems in the form of ever changing collections of MIDI devices and personal computers.

Flutists and violinists have learned to overcome considerable problems of control and coordination in order to make music. Pianists play expressively even though their instrument offers only a single timbre. Similarly, now that computer music is widely available thousands of musicians, both professionals and amateurs, are busy making music without being terribly concerned about not having the biggest, fastest or most flexible devices. They are simply using the available instruments to make music. This is the real computer-music revolution. It is no different from the use of logs for drums and reeds for flutes: people take whatever tools for musicmaking their environment provides and make music.

I do not mean to dismiss totally the work of the authors or of others at IRCAM. The development of inexpensive computer-music instruments has depended on techniques pioneered at IRCAM, Stanford University, the Massachusetts Institute of Technology and similar centers of technological research in computer music. I trust that further advances in the fields of computer music and digital audio technology will continue to come out of these institutions.

NEIL B. ROLNICK

Department of Music
Rensselaer Polytechnic Institute
Troy, N.Y.

To the Editors:

There is no implication in our article that using the computer in music requires "special" or "extraordinary" musical concepts—only concepts that

Some computers are trying to make headlines with features Macintosh had four years ago.

Now for today's news.

The latest addition to the Macintosh family, Macintosh IIx. Its fundamental strength is its consistent, intuitive Macintosh way of working. But with its 68030 chip, up to 8 megabytes of RAM, NuBus™ architecture and math coprocessor, it's no slouch in purely technical terms, either. Macintosh II comes with a CPU (not shown) and a mouse. The keyboard and monitor are sold separately for the sake of custom configuration. In this piece, we have shown Macintosh IIx with the extended keyboard and 16" color monitor.

How things change.

Not too long ago, Macintosh® computers were something of a novelty in the business world.

Now MBAs stand in line to snap up tickets to the latest Macintosh productivity seminars.

Blue suits and gray suits outnumber sweatsuits at Macintosh trade shows and exhibitions.

Fortune 500 companies bring thousands of Macintosh computers in through the *front* door.

And the same computer-makers who just four years ago elbowed each other in the ribs at the sight of Macintosh are introducing computers with strangely familiar features.

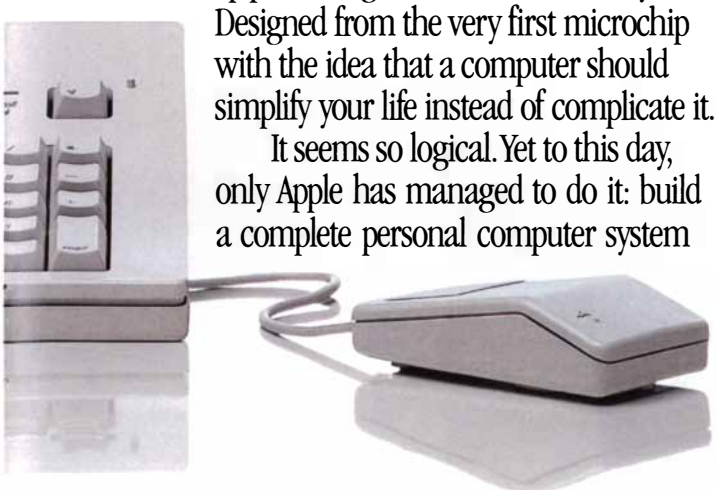
How things change, indeed. Just how much they've changed is indicated by a recent, independent survey of *Fortune 1000* MIS managers.

This notoriously uncompromising group, who buy computers by the truckload, have just rated the Macintosh system far higher than the other leading personal computer system. Specifically: 34% higher for ease of learning, 31% higher for ease of use and 27% higher for learning new programs.

Of course, other computer companies recognize a bandwagon when they see one.

But, in the computer business, playing "if you can't beat 'em, join 'em" isn't quite so easy.

Because Macintosh isn't just a mouse. Or icons. Or pocket-sized disks. Or adjustable windows. Or even desktop publishing. Macintosh is a total system.



Designed from the very first microchip with the idea that a computer should simplify your life instead of complicate it.

It seems so logical. Yet to this day, only Apple has managed to do it: build a complete personal computer system that accommodates the demanding needs of human nature and common sense. And American business is reaping the benefits.

Now the Macintosh system has grown. With computers and printers for almost every conceivable job. And thousands of programs that all work in the same graphic, intuitive way. Utilizing the innovative Macintosh features to the fullest extent.

HyperCard® software, the latest Macintosh exclusive, makes the personal computer more personal than ever.

No other computer system works like it. Not four years ago. Not today. Maybe not ever.

The obvious conclusion being, if you really want a computer that works like a Macintosh, it's a good idea to go out and get yourself a Macintosh.

Then you can take advantage of this year's real news: Macintosh as your own, personal key to computers of almost every size, shape and brand name. Allowing you, for the first time, to use a Macintosh and its point-and-click simplicity to work with virtually *all* computers, no matter how large or complex. Exploiting their vast resources with the ease and efficiency that Macintosh made famous.

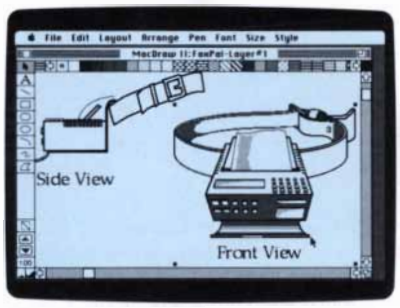
In other words, more power to be your best. No matter what it is that you do best.

Of course, all the original Macintosh features are still good news. It's just that they're old news. And if you turn the page, you'll understand why.

Old news.

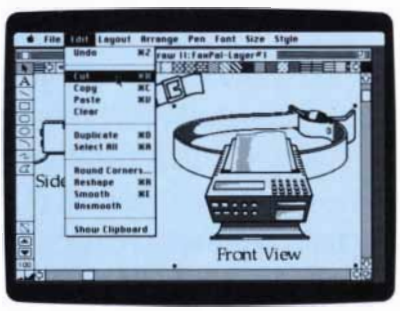
Point.

This basic human skill is also the basic Macintosh skill. So working with a Macintosh seems like second nature. And that's the whole point.



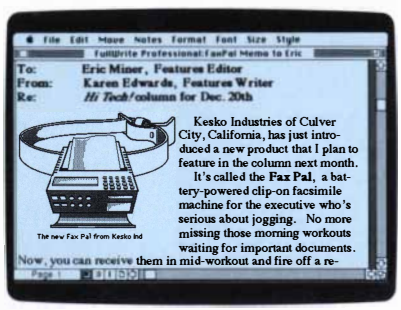
Click.

After you point at what you want to do, just push a button to do it. That's how you work with almost every Macintosh program, not just a select few.



Cut.

Try this with any other computer and you'll understand what makes Macintosh Macintosh. Pick any program, do some work, then cut...



Paste.

...and paste it into another program. It works with text, numbers, even pictures. And virtually every Macintosh program.

News.

Look.

HyperCard, a radically different kind of information organizer, now comes with every Macintosh. It lets you search through forests of knowledge with exhilarating speed.

Link.

After you've found what you're looking for, you can connect it to other pieces of information. And almost any form of information: text, graphics, sound, animation, even video.

Create.

HyperCard lets you customize information with splashy video effects and dynamic graphics you have to experience to believe. It's easy (obviously) and fun (really).

Control.

HyperCard includes the first programming language for the rest of us. To give you a personal say in how your Macintosh works. And something no other computer system offers.

Old news.

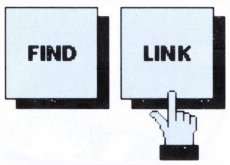
A mouse. *Used to be only Macintosh had these simple, elegant devices. But ours still stands out. It has only one button. So you can't push the wrong one.*

Hard disks. *Today, they're the most popular method of storing information. So Apple offers six different models with up to 80 MB of storage.*

Pocket-sized disks. *It doesn't take a Ph.D. from M.I.T. to realize that a disk in a protective case that slips in your pocket is better than a flimsy one that doesn't.*

What-you-see-is-what-you-get. *Macintosh users have always been able to print on paper what they see on the screen. And it's a built-in talent, not a costly option.*

News.



HyperCard buttons. *The biggest news in controlling your computer since the mouse. Now you can link information any way you want. And actually put it to work.*

AppleCD SC™ Drive. *CDs give you access to enormous archives of words, pictures and sounds. CD ROM is the future of information storage. And Apple has it now.*

Apple FDHD™ 1.4 Drive. *Reads MS-DOS and Apple® II as well as Macintosh formats, making it easy to transfer non-Macintosh files to work on a Macintosh IIx.*

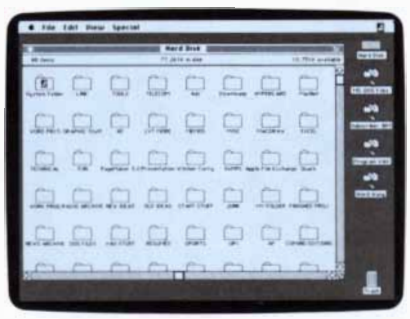
LaserWriter® family. *Choose the LaserWriter II that meets your needs. And print what you see on a Macintosh screen with stunning precision.*

Old news.

Basic applications. *Once, other computers had all the great software and Macintosh had all the promises. Now others are making promises Macintosh has already kept.*

Adjustable windows. *On a Macintosh, they've always been standard equipment. With most other computers, they're an expensive option.*

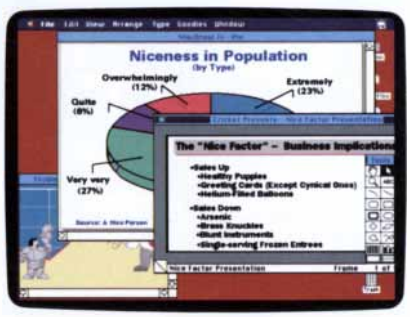
Icons. *Pointing at a familiar symbol is a lot easier than typing commands like "ERASE A:1NSREV .23." But today, we've taken the idea a lot further.*



The Macintosh desktop. *Perhaps the single most influential computer screen of all time. Simple, graphic and as easy to understand as the top of your desk.*

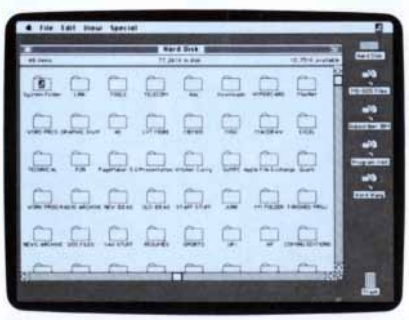
News.

Over 3,000 second-generation applications. *The most advanced business software is now available on Macintosh. And from word processing to CAD/CAM, it all works the same intuitive way.*



MultiFinder.™ *Just click to jump between programs. A big advantage on a computer that lets you cut work from one program and paste it into another.*

Macintosh connects. *Even the most dauntingly complex computer can now work as productively as a Macintosh. But only if you connect it to a Macintosh.*



The Macintosh desktop. *Thousands of improvements later, the desktop still looks familiar. So if you learn one Macintosh, you've learned them all. Even those we haven't invented yet.*

The good news abo
Brought to you

IBM Mainframe: Old news.
Connected to an IBM mainframe, a Macintosh works like any good computer terminal. But when you're not using it as a terminal, you get all the advantages of a Macintosh.

IBM Mainframe: News.
New programmers' tools like MacAPPC™ and MacWorkStation™ let you work with the same mainframe like you work with a Macintosh. A graphic improvement, as you can see.

IBM PC: Old news. *A Macintosh II or SE can run MS-DOS programs. Or, you can convert MS-DOS files to Macintosh files and enhance your information with the graphic capabilities Macintosh is so famous for.*

IBM PC: News. *An AppleShare® file server lets workgroups of different computers work together. Macintosh users see the group's work as Macintosh file folders. MS-DOS users see the work as subdirectories. Everybody's happy.*

ut other computers.
by Macintosh.

Digital VAX: Old news. *Macintosh can connect to a VAX minicomputer, too, and tap its enormous power like any workstation. Of course, if you don't know the VAX way of working, you'll have to learn.*

Digital VAX: News. *Now Macintosh can also bring you the same power in the familiar point-and-click Macintosh form. So if you don't know the VAX way of working, you won't have to learn it.*

Wang minicomputers: Old news. *Macintosh can work with the Wang word processors found in thousands of companies. Then give you the considerable advantage of Macintosh productivity when you work alone.*

Wang minicomputers: News. *Now Macintosh can work with the same Wang documents and bring them to life with a nearly infinite choice of fonts, type styles and sizes. Giving you Macintosh productivity all the time.*

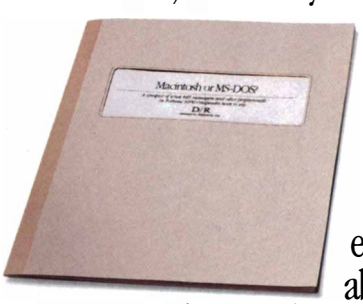
And now for tomorrow's news.

Apple invented the personal computer as you know it today. And we're working on how you'll know it tomorrow. One concept: The Knowledge Navigator,™ incorporating video images, a video camera, voice recognition and much of today's Macintosh.

What will the Macintosh computer of the future look like?

It probably won't fit in your wallet, hover in mid-air or come when you call it.

In fact, it will likely bear a striking similarity to the Macintosh of today. Except it will let you do things you've never done before.



Corporate America loves Macintosh. Read all about it: 800-446-3000, ext. 475.

Desktop video, for example. Imagine being able to create and edit multimedia presentations as elaborate as a Super Bowl halftime show without leaving your desk. We're working on it.

Macintosh will also help you do the things you've been doing all along. But with ruthless efficiency. Extending the Macintosh way of working to more and more information resources with better and better tools.

And no matter what new things we build into Macintosh, they'll be built on what you already know. Just like always. To protect not only what you've invested in your computers, but also what you've invested in yourself.

Speaking of which, it's time you found out just how good an investment Macintosh computers really are today. If you call 800-446-3000, ext. 475, we'll be happy to send you a copy of the survey of *Fortune 1000* MIS managers we mentioned way back in the beginning. It shows you, in hard numbers, what a powerful impression Macintosh has made on people who buy computers in quantities most of us reserve for groceries.

Of course, it's only the beginning of why Macintosh will continue to be the most exciting story in computers for years to come.

And that's the way it is.

The power to be your best.™

make sense in this new context. One does not write in the same way for a string quartet as one does for a symphony orchestra.

The problems of control and coordination referred to in our article apply only to machines, and indeed such problems do exist. For example, today it is impossible to synchronize MIDI devices at the level of samples (the numbers that describe a sound's waveform). Readers interested in this discussion can read Richard Moore's article "The Dysfunctions of MIDI" in the spring issue of *Computer Music Journal*, published by the MIT Press.

It is the role of IRCAM and other research centers specializing in computer music to try new musical ideas, which today may require machines that may seem large and complex. We all know that today's large and expensive machines become tomorrow's small and inexpensive ones.

PIERRE BOULEZ

ANDREW GERZSO

Institut de Recherche et Coordination Acoustique/Musique (IRCAM)
Paris

To the Editors:

The doped, electrically conducting rubber reported by Mrinal K. Thakur of the AT&T Bell Laboratories and described in "Son of Rubber" ["Science and the Citizen," SCIENTIFIC AMERICAN, August] is an interesting discovery, but it is not the first example of a nonconjugated electrically conducting polymer. In 1981 my colleagues and I showed that polysilane polymers become electrically conducting when they are doped with oxidizing agents. Like rubber but unlike other conducting polymers, the polysilanes are colorless, flexible and soluble. Since the polysilanes contain no multiple bonds at all, conjugation of the kind mentioned in your story is impossible.

The highest electrical conductivity observed for polysilanes was about .1 siemen, which is similar to the value reported for doped rubber by Thakur. Whether or not the mechanism of conduction is also similar in polysilanes and doped rubber remains to be established.

ROBERT WEST

Department of Chemistry
University of Wisconsin
Madison

DODGE DYNASTY LE. *With its combination of elegance and value, the 1989 Dynasty LE gives classic performance new meaning. Complete with comfortable room for six. A wide range of interior amenities. A standard V-6. Advanced features like an electronic 4-speed Ultradrive automatic and available anti-lock brakes. And our exclusive 7 year or 70,000 mile Protection Plan.* Dodge Dynasty. Classic performance defined.* 7/70

Dodge

THE NEW SPIRIT OF DODGE

THE PERFORMANCE DIVISION OF CHRYSLER MOTORS

*See this powertrain limited warranty & its restrictions at dealer.

BUCKLE UP FOR SAFETY.

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

NOVEMBER, 1938: "Beneath the rising walls of the Westinghouse building at the New York World's Fair 1939, the world's first Time Capsule has begun its long journey into the future—a journey that, it is hoped, will extend through 5,000 years of time and give future archeologists a considerable insight into our present civilization. The Capsule consists of an outer shell of Cupaloy (chosen because of its electrical qualities and resistance to corrosion) and an inner envelope of Pyrex glass, sealed, evacuated and filled with nitrogen. More than 40 articles of common use are included, along with a carefully prepared microfilm 'essay' on our times. (The complete September 1938 *Scientific American* is reproduced in this microfilm.)"

"Atomic particles of enormous penetrating power come into the earth's atmosphere from some sources outside the solar system, but exactly what these particles are, or where they come from, or how they were formed, or when, we still do not know. It is experimentally demonstrated, at least, that these cosmic rays are of extragalactic origin."

"A synthetic fiber having the appearance of wool can be manufactured from casein, a milk by-product. To make the fiber, casein is softened in water and dissolved in a solution of caustic alkali. It becomes a thick, sticky mass and is carefully worked into the proper consistency by aging, addition of modifying agents and dilution. The mass is then forced through multiple spinnerets, and the resulting fibers are separated and hardened in an acid bath."

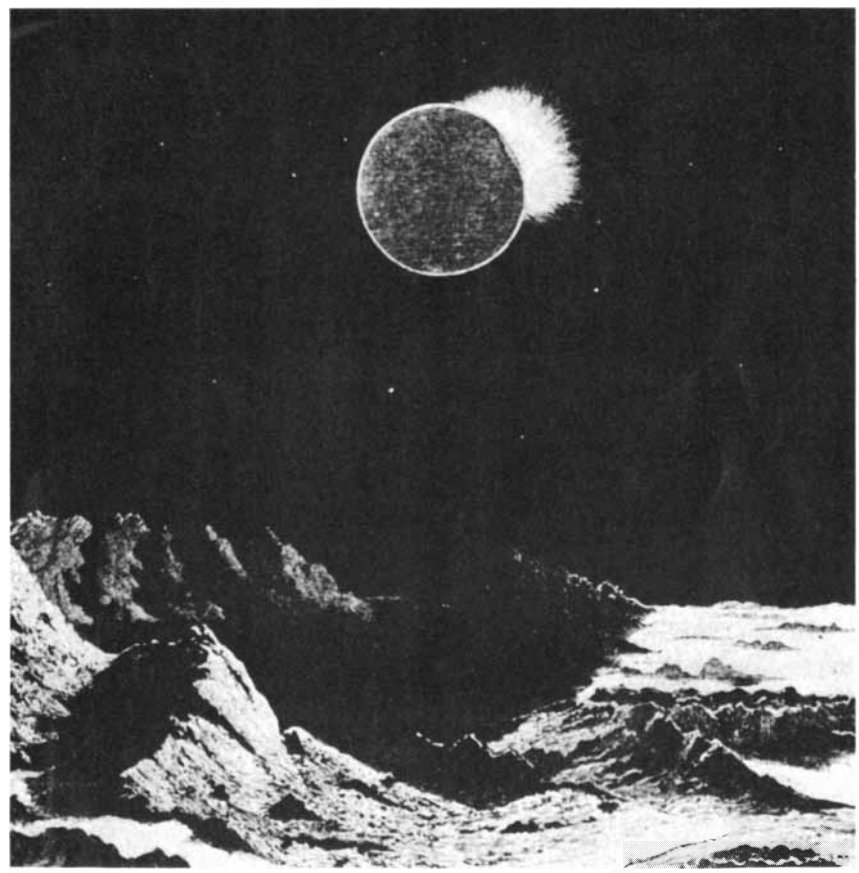
SCIENTIFIC AMERICAN

NOVEMBER, 1888: "The excavations commenced by Dr. Schliemann at Mycenæ are still being energetically carried on, and continue every day to bring to light fresh objects of great archæological and anthropological interest. The entire terrain around the town is full of tombs belonging to an epoch antecedent to Homer."

"At the recent *soirée* of the Royal Society, the principal feature of the evening was the soap bubble experiments of Mr. C. V. Boys. One of these afforded a beautiful illustration of the diffusion of gases through a medium such as a soap film, which remains intact the while. Thus a soap bubble was blown with pure oxygen gas, and immersed for a few seconds in a bell glass containing the invisible vapor of ether. When the bubble was withdrawn and approached to a flame, it exploded with a flame and a report, showing that during the short time of its exposure to the ether vapor, diffusion had occurred, and the original pure oxygen had given place to an explosive mixture of oxygen with the ether."

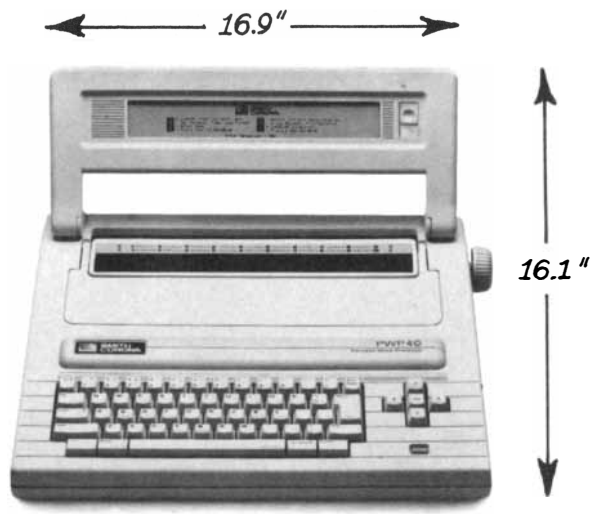
"If ships could be pulled through the water, instead of being driven by means of the ordinary propellers now in use, then a saving, on the average, of 40 per cent. in engine power would be gained. Both the screw and the paddle push the sustaining water from the ship, and thus augment its resistance, and it has been found by experiment that the thrust of the propeller surpasses the pull of a tow rope, on the average, by 40 per cent. to obtain the same speed in the same vessel. The solution is to propel the ship by means of revolving sails acting in the air. The air propeller is, in its outer shape, somewhat similar to the ordinary water screw, with sails or blades made of thin sheet steel."

"Since our sight, sharpened by the telescope, has been able to penetrate the distances of space and to recognize more clearly the special structure of other worlds, the question has arisen: How would the heavens appear from a different standpoint? This question suggests itself specially when watching a total eclipse of the moon, for the astronomer tells us that it is our earth which is throwing its shadow on its celestial companion, and that we really are the cause of this heavenly display. The astronomer, aided by the artist's eye trained to perspective, answers with the accompanying landscape, which places us on our neighboring world just at the time when the shadow of the mother planet is crossing it."



A lunar eclipse as seen from the moon

Announcing a breakthrough of minor proportions.



The Smith Corona PWP 40 Personal Word Processor.

Our biggest idea yet just happens to be our smallest idea yet.

Less than 17 inches across and weighing in under 17 pounds, the Smith Corona PWP 40 personal word processor redefines portability.

Yet for all its size (or, should we say, lack of it), there's never been a portable word processor better equipped to carry its own weight.

Using the PWP 40 is simplicity itself.

Your writing appears on the 8 line by 80 character backlit display in front of you. You can move blocks of copy, insert copy into existing paragraphs or delete copy in seconds. All before you ever touch a piece of paper.

As for PWP 40's memory, it's simply unforgettable. There's a 42,000 character internal memory as well as an external memory via DataDisk. Each DataDisk has a capacity of 100,000 characters!

Of course, PWP 40 wouldn't be complete without a correction system. And this one makes correcting mistakes as easy as making them.

The Spell-Right™ 50,000 word dictionary catches misspelled words and typing errors quicker than you can. WordEraser® erases your errors. Word-Right® AutoSpell® even displays the correct spelling and replaces the error for you.

Add deluxe features like automatic reformatting, automatic page numbering, bold type, dual pitch and much more, and you've got a lot of word processor in a surprisingly small package.

There is one other thing that will surprise you about PWP 40. Its price. You'll be happy to hear that PWP 40 is surprisingly affordable.

You see, we didn't just make PWP 40 simple to use. We also made it simple to buy.

SMITH CORONA®
TOMORROW'S TECHNOLOGY AT YOUR TOUCH™

For more information on this product, write to Smith Corona Corporation, 65 Locust Avenue, New Canaan, CT 06840 or Smith Corona (Canada Ltd.), 440 Tapscott Road, Scarborough, Ontario, Canada M1B 1Y4.

The life busine drop-dead tech

In life insurance, the saying goes, the paperwork can kill you. The paperless office is every company's dream.

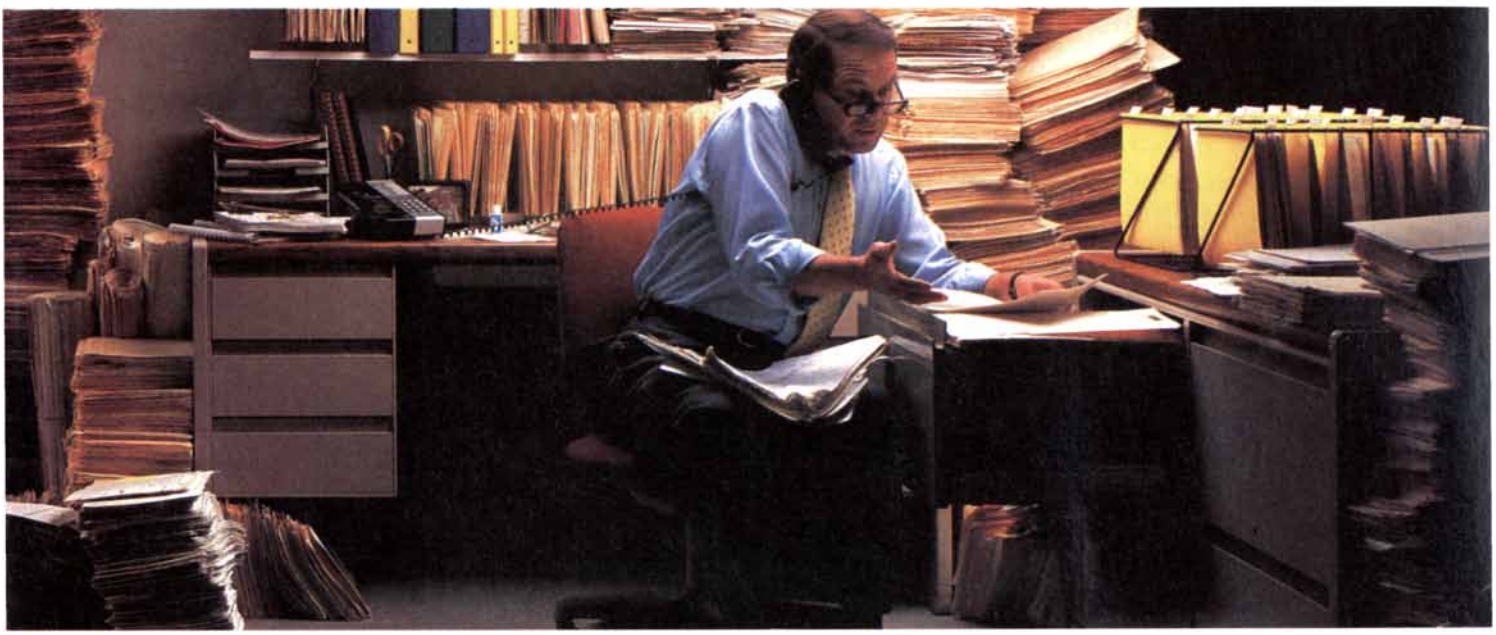
That's why lots of huge insurance companies are using NYNEX.

Recently, to simplify life for one of these colossi, NYNEX replaced acres of paper with an innovative new voice and data network, connecting their operations center with their headquarters. It makes their systems compatible.

And remarkably, it allows them to access their whole information network by plugging any work station into any phone jack.

This NYNEX technology isn't just convenient. It's saving them a lot of money.

ss needed some nology.



Which is always the best policy.

The NYNEX family of companies would like to work with you, too. We offer you everything from computer networks and software, to mobile phones, to the services of New England Telephone and New York Telephone.

Call us at 1-800-535-1535.

When we say the answer is NYNEX, it's not just an empty claim.

Need to communicate? Need to compute? The answer is

NYNEX

Do you have the true spirit of enterprise?

The Rolex Awards for Enterprise were inaugurated in 1976 with a single goal: "To encourage the spirit of enterprise in individuals throughout the world by acknowledging outstanding personal contributions made in selected categories of human endeavor." The qualities that the Awards set out to honor are the ones that have long been shown by Rolex and by Rolex owners.

Rolex now takes pride in announcing that 50,000 Swiss francs and an inscribed 18kt gold Rolex Chronometer will be awarded to the five self-motivated winners who have developed projects displaying the most outstanding spirit of enterprise.

AREAS OF ENTERPRISE.

Candidates may submit projects in three categories—Applied Sciences and Invention; Exploration and Discovery; and The Environment. The Selection Committee will be evaluating these projects in terms of their spirit of enterprise, originality and the likelihood of realization.

If your venture falls outside the committee's area of expertise, Rolex will call on specialists to appraise the merits of a particular proposal.

An illustrated Spirit of Enterprise book, summarizing the best projects submitted for the 1990 Awards, will be published concurrent with the Awards ceremony. The publicity given to projects in previous editions of the book has often led to additional support from a wide range of sources.

HOW TO APPLY.

Your entry must be submitted in English on an official application form and must reach the Secretariat before March 31, 1989. To obtain that form, and the rules and conditions for entry, write: The Secretariat, The Rolex Awards for Enterprise, P.O. Box 178, 1211 Geneva 26, Switzerland.

The Awards will be presented in Geneva at the end of April 1990.

If you possess originality, imagination and initiative, you will not let this opportunity pass you by.

ROLEX

THE 1990 SELECTION COMMITTEE

CHAIRMAN:

Mr. André J. Heiniger (Switzerland). Chief Executive Officer and Managing Director of Montres Rolex S.A., Geneva.

MEMBERS:

Dr. George V.B. Cochran (USA). Professor of Clinical Orthopaedics, Columbia University, New York. Past President of the Explorers Club.
Miss Fleur Cowles (Great Britain). Painter and authoress. Member of the World Wildlife Fund International Council.
Dr. Xavier Fructus (France). Specialist in hyperbaric physiology; Scientific Director of COMEX, Marseilles.
Professor Yoshimine Ikeda (Brazil). Oceanographer specializing in Antarctic research. Professor at the Oceanographic Institute of the University of São Paulo.
Mr. Kisho Kurokawa (Japan). Architect. President of Kisho Kurokawa Architect and Associates, and Director of the Institute of Social Engineering, Inc., Tokyo.
Mr. Hans Joachim Panitz (FRG). Telecommunications engineer. Payload Operations Manager at the German Aerospace Establishment—DFVLR, Cologne.
Professor Carlo Rubbia (Italy). Physicist and 1984 Nobel Laureate. Project Spokesman at the CERN (European Laboratory for Particle Physics).
Mr. Robert Sténuit (Belgium). Underwater archaeologist and author.

SCIENCE AND THE CITIZEN

Ignorance in Action

Politicians hear but do not heed scientists' advice on drug abuse

For the past six months or so Congress has heard testimony from a steady stream of sociologists, physicians and other investigators of drug abuse. The consensus among scientists was perhaps best expressed in a recent report by the American Medical Association, which recommended "moving away from an adversarial approach to a therapeutic model." (Although the media have spotlighted advocates of legalization, neither the AMA nor most other experts favor that step.)

Has the Government been listening to its scientific counsel? The adversarial approach has certainly prevailed during the Reagan Administration. According to the General Accounting Office, funds for interdicting the flow of illegal drugs into the U.S. and enforcing drug laws have increased almost fivefold in the past seven years to $2.5 billion this year; not counting AIDS-related spending, funds for preventive education, for treatment and rehabilitation of addicts and for research into the extent and causes of drug abuse have remained essentially constant at about $500 million.

In September the House of Representatives showed what it thought of the experts' advice by passing a $6 billion drug bill that provides more funds than ever for interdiction and law enforcement. Although some new funds are earmarked for prevention and treatment, the "therapeutic model" gets short shrift: the bill would make even casual users of illicit drugs ineligible for student loans, public housing or other Government assistance. It also would allow illegally collected evidence to be considered in trials of drug offenders if the police had shown "good faith." The Senate has a similar bill in the works.

"Politicians view science as irrelevant when it comes to dealing with drug abuse," notes Karst Besteman, a former deputy director of the National Institute on Drug Abuse who is now with the Alcohol and Drug Problems Association. "When they get in these feeding frenzies, they just worry about what the public wants." The public wanted a "tough" law-enforcement bill, Besteman remarks, and the House of Representatives provided it in spite of the fact that it represents a rejection of informed points of view.

22 PHYSICAL SCIENCES
30 BIOLOGICAL SCIENCES
36 MEDICINE
40 OVERVIEW

Besteman and other experts acknowledge that they do not know all the answers. Indeed, it is precisely because drug abuse is so poorly understood, they say, that more research is so desperately needed. Almost all such research is funded by the National Institute on Drug Abuse. NIDA has a 1989 budget of $253 million, about half of which was recently added for AIDS-related work.

Epidemiological data, in particular, are in short supply. There is only one widely accepted source of statistics on the extent of illegal drug use: an annual survey of high school seniors by the Institute for Social Research at the University of Michigan. Rhetoric about the so-called drug epidemic notwithstanding, the survey found that the use of cocaine and marijuana decreased last year and that use of illicit drugs in general has been declining for five years. Lloyd D. Johnston, who heads the survey, readily admits that it has an obvious flaw: it does not include high school dropouts. Johnston says he would like to expand his sample population; perhaps even more important, he would like to survey people enrolled in various prevention and treatment programs in order to evaluate their effectiveness.

Better knowledge of the biological effects of addiction are also needed, according to Mary Jeanne Kreek of Rockefeller University. Such research can help lead to such treatments as methadone, a nonintoxicating substitute for heroin. As many as 80 percent of the addicts enrolled in methadone-maintenance programs stay off heroin; program participants are better able to hold jobs and less likely to commit crimes, and since methadone is taken orally, their risk of contracting AIDS is greatly reduced. Investigators are now studying other drugs that may help in treating cocaine addicts. Studies of addiction may also alter the popular notion that addiction results solely from a lack of will or moral character. Kreek says her own work suggests that heroin addiction is linked to metabolic disorders.

Research can aid in preventing drug abuse by alerting the public to its various dangers. A host of new reports have detailed the ill effects of cocaine, which only six or seven years ago was widely thought to be relatively benign. One disturbing finding is that pregnant women who use cocaine, particularly in the smokable and highly addictive form called crack, may irreversibly damage the fetus they carry. Researchers may not always reach politically acceptable conclusions, however. Last summer members of the White House Conference for a Drug Free America, a so-called grass-roots group appointed by the president, criticized NIDA for not stating unequivocally that marijuana is addictive. If the agency has not produced conclusive evidence that marijuana is harmful, it is not for lack of trying, according to one NIDA official. "Never has so much money been spent trying to find something wrong with a drug and produced so few results," he says.

David F. Musto of Yale University, who described the history of illegal drugs in the U.S. in his 1973 book *The American Disease,* notes that political fashion has often unduly influenced the direction and funding of research on illegal drugs. The result is that investigators lack the foundation of reliable knowledge needed to deal effectively with new trends in drug abuse. "Research requires steady funding and an atmosphere of calm," Musto says. "Up to now it's been like a cork on the water." —*John Horgan*

Reactors Redux

A political fight develops over new military reactors

Battle lines are being drawn for a fight over the Department of Energy's plan to build two new reactors for producing the raw materials of nuclear weapons. The new facilities

would be the first large reactors, civilian or military, to be ordered in the U.S. in 15 years.

The department says the reactors are needed in order to guarantee the supply of tritium, an isotope of hydrogen that is made by irradiating lithium targets and serves to boost the explosive yield of warheads. Doubters such as Senator Ernest F. Hollings of South Carolina question whether the cost of two reactors—estimated at $6.8 billion—can be justified even under pessimistic assumptions about arms reductions in the future.

The need for some new capacity is widely acknowledged. Each year 5.5 percent of the tritium in a warhead inexorably decays and has to be replaced. Existing tritium-production reactors were not built to conform to modern safety standards, and they are showing signs of age.

At the Energy Department's only remaining tritium-production site, the Savannah River Plant in South Carolina, one aging heavy-water reactor was retired last year because of cracks in its casing; three other reactors at the site, which are now the country's only source of tritium, are frequently out of service for repairs or maintenance. When they are operating, they have to be run at less than half power for safety reasons.

A damning study by the National Research Council last year cast serious doubt on the reactors' future. It identified problems such as cracks caused by stress and corrosion in the reactors' cooling systems, found "significant uncertainties" about the ability of the reactors' confinement systems to prevent a release of radioactivity in a severe accident and criticized the Energy Department's management of the facility. Confidence in the reactors was further shaken in August when safety officials discovered that workers who were having problems restarting one of the reactors had tried to increase power rather than immediately shutting the reactor down. A few days later the same reactor had a probably unrelated but nonetheless worrisome power "spike."

Just before the incident in South Carolina the Energy Department announced plans for the two new production reactors. One large heavy-water reactor would be built as quickly as possible at the Savannah River site; construction might take a decade. At the same time planning would start for a so-called modular high-temperature gas-cooled plant, to be constructed at the Idaho National Engineering Laboratory. The department says the Savannah River reactor alone could satisfy the projected tritium requirements; the Idaho reactor would provide 50 percent surplus capacity.

The surplus would be available even if no further major arms-reduction agreements are reached with the Soviet Union. An arms agreement would reduce the demand for tritium and at the same time greatly increase the supply of tritium scavenged from retired warheads. David Albright, a researcher for the Federation of American Scientists, argues that a decision to build even one new reactor could safely be delayed until the outcome of the current Strategic Arms Reduction Talks becomes clear. Senator Edward M. Kennedy of Massachusetts, among others, believes excessive tritium-production capacity could imperil attempts to monitor tritium production as part of some future arms-control agreement.

At the opposite end of the spectrum is Senator James A. McClure of Idaho, a champion of the two-reactor plan. McClure says that prudence demands having diverse sources of tritium in case of an engineering failure, and he stresses the dangers of relying on the existing production reactors. Work on the new reactors should begin as soon as possible, he thinks, foreseeing technical hurdles. He believes that the Energy Department's Energy Research Advisory Board, which reported that the heavy-water reactor planned for South Carolina was the maturest production technology, underestimated the difficulties it will face. McClure thinks new designs for the core and lithium targets will be necessary.

McClure also argues that the second proposed tritium-production reactor could be a technology demonstrator for future "inherently safe" civilian power reactors. The probable design, by General Atomics of San Diego, employs a graphite moderator and a helium coolant; it does not require elaborate and failure-prone systems to prevent the release of radiation in an accident. Even if the cooling system fails, the reactor, built underground, is designed to lose heat naturally and stay at a safe temperature. The one existing reactor of similar design in the U.S., an aboveground power station at Fort St. Vrain, Colo., has had technical troubles, but General Atomics maintains that these could now be overcome.

At a time when a renaissance of civilian nuclear power is seen as a possible way of reducing reliance on fossil-fuel combustion, the notion of having the military foot the bill for testing a new reactor design has attracted some support. The Energy Department is expected to request funds for the new reactors in the Administration's budget for fiscal year 1990, which is scheduled to be sent to Congress in January. —*Tim Beardsley*

The Shroud of Turin

New findings confirm old ones: it is a 14th-century fake

In 1356 the French nobleman Geoffrey I de Charny proclaimed that he had found the long-sought burial shroud of Jesus: a linen sheet bearing the faint outline of a bearded man who had wounds on his forehead, torso, feet and hands. Although the bishop of Troyes promptly declared that the shroud was a fake painted by a local artist, the public preferred to believe de Charny's story. Today the cloth is housed in St. John's Cathedral in Turin, in northern Italy, where it still attracts crowds of awestruck worshipers.

Now, perhaps, the myth of the Shroud of Turin can finally be laid to rest. At the request of the archbishop of Turin, workers at the University of Oxford, the University of Arizona and the Federal Institute of Technology in Zurich have dated the cloth by measuring its carbon-14 levels. The British Museum supervised the research. The archbishop has delayed announcing the results, but they were leaked to the press in September. The dating shows that the flax in the cloth was grown in the time not of Christ but of de Charny.

Walter C. McCrone, a microscopist who specializes in authenticating art objects, told SCIENTIFIC AMERICAN that a Catholic official from Turin had somewhat sadly acknowledged to him that the leaks were accurate. That was no great surprise to McCrone, who heads his own research institute in Chicago. Almost 10 years ago he examined the shroud himself and concluded it was a 14th-century forgery.

In the late 1970's McCrone joined a group of about 30 other U.S. scientists—most of them from Federal laboratories—in analyzing samples of the shroud supplied by the cathedral. McCrone says he and others in the group were at odds from the start. The first time they met, he recalls, "almost to a man they had big crosses around their necks. I thought they might be biased."

Working independently, McCrone examined fibers and other materials

1983 Car and Driver Ten Best List.
1984 Car and Driver Ten Best List.
1985 Car and Driver Ten Best List.
1986 Car and Driver Ten Best List.
1987 Car and Driver Ten Best List.
1988 Car and Driver Ten Best List.

There's a pattern here somewhere.

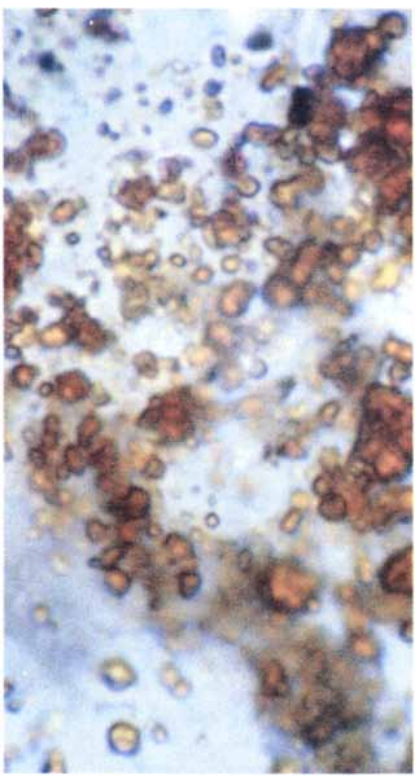
The 1989 Accord LXi.

HONDA

lifted from the surface of the cloth with sticky tape. Fibers from the region representing the figure of Christ, he found, were imbued with a pale, gelatin-based medium speckled with particles of red ocher. Fibers from the dark areas representing wounds were stained not with blood but with particles of vermilion.

The cloth had been painted, but when? Red ocher, which is made of a natural iron oxide found in the earth, has been used as a pigment for thousands of years. So has vermilion, which can be made from an ore of mercury called cinnabar. The vermilion on the cloth, however, was a synthetic form developed in the Middle Ages. Moreover, the practice of painting linen with gelatin-based temperas began in the late 13th century and was common in the 14th. McCrone concluded that the bishop of Troyes had been right: a 14th-century artist had forged the shroud.

McCrone says his results "were universally ignored." Church authorities chose to publicize the findings of other members of the U.S. group, who maintained that the figure of Christ seemed to have been produced by an intense burst of energy and that its wounds contained blood. Now that his work has finally been corroborated, McCrone notes that his satisfaction is tempered somewhat by sympathy for those people who believed in the shroud. "I wanted it to be real too," he confesses. —*J.H.*



***PARTICLES OF RED OCHER** in a gelatin-based medium tint the surface of the Shroud of Turin. The micrograph is by Walter C. McCrone.*

Why Warfare?

A broad study of preindustrial societies offers some answers

The diversity and intensity of opinions in a given field of inquiry are usually inversely proportional to the amount of objective data available to scholars. This theorem seems particularly applicable to studies of human warfare. Scholars often base elaborate theories on the behavior of a select group of societies, whether they are tribes living in rain forests or modern nation-states. Yet what is true of one group may not be true—or as true—of others.

Carol R. Ember of Hunter College in New York and Melvin Ember of Human Relations Area Files, Inc., in New Haven hope to provide scholars with more far-reaching and rigorous data. For the past five years the wife-and-husband team has been conducting a broad cross-cultural study funded by the National Science Foundation. Their goal is to search for statistical correlations between warfare as practiced around the world and various hypothetical causes.

The study's raw material is a mélange of reports on 186 different societies from every inhabited continent. The reports range from 18th-century observations of the Pawnees to recent ethnographic descriptions of the Masai. Almost all the societies are preindustrial and even "primitive" compared with modern nation-states, according to the Embers. Of those that still exist, all but a few have now been subjected—or "pacified"—by (supposedly) more advanced states. The Embers included only reports that predated such pacification.

To extract statistical data from these often anecdotal and highly subjective reports, the Embers depend on two "coders." Each reads the literature describing a given society and then independently gauges how often the society engaged in war and other practices and how susceptible it was to various environmental influences. If even one of the coders thinks the literature is too vague or if the two coders' answers do not overlap substantially, the society is excluded from the study.

Although the research is not complete, the Embers have reported some preliminary findings. Fear of deprivation seems to have been the most powerful motivating factor for making war. The victors in a war almost always took land, food, chattel or other economic resources from their enemies. Yet it was not so much chronic, predictable shortages that seemed to spur societies toward war as it was unpredictable disasters—such as drought, flood, pestilence or frost—leading to severe deprivation, according to the Embers.

The study has cast doubt on other hypotheses commonly advanced to explain warfare. Some students of human nature have proposed, for example, that males who are sexually frustrated are more likely to initiate wars; actually men in warlike societies seem to have engaged in more premarital and even extramarital sex than men in relatively peaceful societies. Other scholars think a society's propensity for warfare should be reflected in high rates of "interpersonal aggression" such as murder and theft; the Embers found no such correlation.

The study did find a correlation between warfare and child-rearing practices that encouraged boys to be aggressive and "tough," but the Embers suspect such practices were a consequence rather than a cause of frequent warfare. They note that even very warlike people, after they had been pacified by other societies, rarely continued to encourage their young males to be aggressive.

In the future the Embers intend to investigate, among other questions, whether warfare is correlated with such "population-control" practices as abortion or infanticide and whether societies known to distribute resources in times of scarcity are less likely to go to war. Eventually the Embers hope to determine whether correlations found in their study hold true for more complex societies, including modern nation-states.

Not all anthropologists accept the validity of the Embers' work. Some object to any cross-cultural comparisons, arguing that each society is unique and can be understood only on its own terms. Others contend that the Embers' method of quantifying historic and anthropological literature involves too much interpretation. Allen W. Johnson of the University of California at Los Angeles maintains that, given the importance of the subject, these objections are overly narrow-minded. "Do you throw out the data and never compare societies?"

BUY 50 SHARES OF STOCK,
TRACK YOUR PORTFOLIO,
GET TAX ADVICE,
SHOP NEIMAN MARCUS,
GET TICKETS TO HONOLULU,
SEE WHO WON THE BALL GAME,
ORDER BABY CLOTHES,
MAKE A LUNCH DATE,
CHECK YOUR BANK BALANCE,
GET A DOZEN EGGS
AND A GALLON OF MILK,
PAY YOUR BILLS,
READ YOUR HOROSCOPE

he asks. "Or do you do the best you can? I respect the Embers for what they're doing." —*J.H.*

PHYSICAL SCIENCES

Ice House

Did a growth spurt of mountain ranges initiate the ice ages?

Beginning about 2.5 million years ago vast sheets of ice pushed south from the Arctic and blanketed much of North America and Europe. Since then the ice has advanced and retreated every 100,000 years or so. The periodic recurrence of these ice ages has been linked to wobbles in the orbit of the earth known as the Milankovitch cycles, but these cycles predate the ice ages by millions if not billions of years. Why did the earth become susceptible to glaciation only recently?

Workers from the Lamont-Doherty Geological Observatory propose that a growth spurt of major mountain ranges over the past five million years may have triggered the glaciation. They suggest that the higher topography of the Rocky Mountains and the Himalayas could have disturbed jet streams in the Northern Hemisphere, diverting frigid arctic air south. Computer simulations lend credence to this hypothesis. The rapid uplifting might also have led, rather circuitously, to a decrease in the amount of carbon dioxide in the atmosphere and hence to a global cooling—a kind of negative greenhouse effect.

Elaborating on this latter hypothesis in *Geology,* Maureen E. Raymo, William F. Ruddiman and Philip N. Froelich note that higher elevations result in higher erosion rates. Lofty terrain attracts more precipitation, and water is more erosive spilling down a slope than on a flatland; moreover, uplifting continually exposes fresh rock to the weather. The weathering of highland rocks produces positively charged ions—including sodium, potassium, magnesium and calcium—that are carried by streams and rivers to the ocean. An increased influx of positive ions makes the oceans more alkaline and decreases the amount of carbon dioxide dissolved in the seawater, in part by locking it up in other carbon compounds. The net effect, since levels of carbon dioxide in the atmosphere and in the ocean are at equilibrium, is a decrease in the amount of atmospheric carbon dioxide.

Over very long time scales the depletion of atmospheric carbon dioxide caused by weathering is balanced by other mechanisms, such as the release of the gas from volcanic eruptions and metamorphic activity. But Raymo, Ruddiman and Froelich suggest that over the past five million years these restorative processes have not kept pace with the depletion caused by increased weathering. They cite numerous studies showing that during this period most of the earth's major mountain ranges have been uplifted at a faster rate than in the preceding five million years. The rate of uplift of the Himalayas and the Andes, in particular, has more than doubled. Core samples extracted from the ocean floor also indicate that by-products of weathering have been deposited in ocean sediments more rapidly over the past five million years.

Apparently millions of years passed before the uplifting cooled the earth's climate enough to bring on glaciation. Even though the same processes continue today, Raymo points out that this slow and subtle negative greenhouse effect is not likely to offset the positive greenhouse effect that is now in the news: the global warming expected to result from the buildup of atmospheric carbon dioxide initiated by fossil-fuel burning and other human activities. —*J.H.*

Plus Ça Change...

Once again: The gravitational constant is constant

The idea that Newton's gravitational constant G may actually change with time has fascinated physicists for a century. The Austrian physicist Ernst Mach stated that the forces acting on water in a spinning bucket, for example, are due to accelerations relative to the distant galaxies. This idea, known as Mach's principle, led naturally to several theories in which G varies as the universe expands.

A different type of "variable G" theory was proposed by P. A. M. Dirac in 1938. He noted that the ratio of the size of the universe to the size of an atom very nearly equals the ratio of the electrical force between an electron and proton to the gravitational force between the same two particles—and that both ratios are the huge number 10^{39}.

Everything in the two ratios except the size of the universe is ordinarily considered constant; it is far from ob-

READ YOUR HOROSCOPE, RUN A COMPANY, LEARN GEOGRAPHY, CATCH UP ON THE LATEST NEWS, FIND OUT IF IT'S RAINING,

WITHOUT LEAVING HERE.

Now when you need to get things done or find out what's going on in the world, we suggest you take a seat. With the PRODIGY℠ service in your home, it's just that easy.

PRODIGY℠, a new computer service from a partnership of IBM and Sears, is available now in selected cities and will be in many more next year.

Get your PRODIGY™ Start-up Kit, pull a chair up to your computer and soon you'll be able to shop, bank, get your groceries, play the stock market, get the advice of experts in over 30 fields, stay on top of the news, make travel plans and much more. All for a low monthly fee.

There may be no getting around all the things you have to do, but now, at least, you don't have to run all around town to do them.

PRODIGY℠
Interactive Personal Service

Anything you can do, you can do better.

The PRODIGY™ service and some of its features may not be available in all areas.
PRODIGY is a service mark and registered trademark of Prodigy Services Company.

vious why the universe should be just the right size to make the ratios equal. Their apparent equality may be a cosmic coincidence but if, for unknown fundamental reasons, the two ratios are in fact always equal, then it follows almost immediately that G should decrease as the universe expands. More recently, superstring theories have also predicted that the constants of nature, including G, may change as the universe ages.

In spite of the many predictions of a nonconstant constant, no experiment to date has found any evidence of variation. Astronomical observations of the orbital period of the moon (which would change if G changed) and Viking lander data on the orbital period of Mars show that any change in G must be at a rate of less than about 3×10^{-11} part per year.

A larger G in the past would also have made the abundance of helium formed during the big bang larger than the 24 percent predicted by the usual value of G. Yet astronomical observations show that helium makes up no more than 25 percent of the universe's mass. This restricts G's variability to a limit of about 2×10^{-11} part per year, or to considerably less in certain cosmological models; G could have changed at most by 20 percent since the big bang.

Now Thibault Damour, Gary W. Gibbons and Joseph H. Taylor report in *Physical Review Letters* that observations of the binary pulsar designated PSR 1913 + 16 give similar results: G can change by no more than $(1 \pm 2.3) \times 10^{-11}$ part per year (2.3 represents two standard deviations).

The binary pulsar, which consists of a neutron star orbiting around another compact object, has been studied intensively since its discovery in 1974. Its orbital period and the rate of change of the period are known to extremely high accuracy—the latter to about 13 decimal places. Einstein's general theory of relativity actually predicts that the period of the binary pulsar will change as the two objects emit gravitational radiation and gradually spiral into each other. The observed rate of change of the period is in excellent accord with relativity, which assumes G is constant.

If G varied slowly, however, the equations describing the pulsar's orbit would be altered and would predict a different rate of change in the orbital period. The small differences between the observations and relativity's predictions constitute the maximum allowed effect of a variable G. This leads to the limit quoted above; as the authors say, this limit is consistent with zero. Some things never change. —*Tony Rothman*

Wobbly Evidence

Doubts remain on "sightings" of planets near other stars

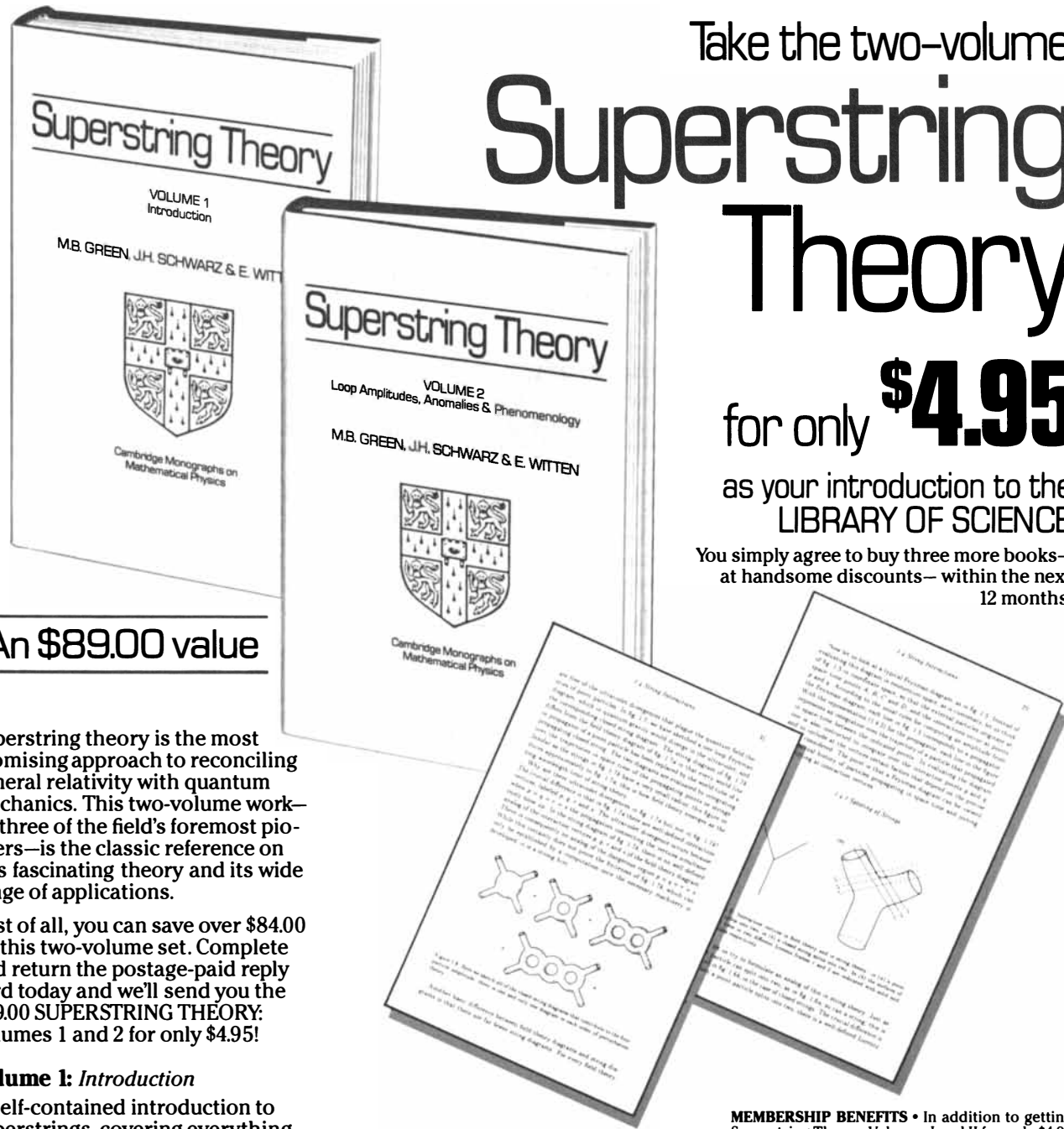
In the hunt for planetary systems other than our own, tantalizing reports have consistently led to disappointment. In the late 1950's Peter van de Kamp of Swarthmore College attributed wobbles he had observed in the position of Barnard's Star, the sun's second-nearest neighbor, to the gravitational tug of an unseen giant planet. Unfortunately the wobbles did not turn up in later, more precise observations. In 1985 workers from the University of Arizona and the National Optical Astronomy Observatories said they had actually spotted—by means of a complex imaging technique known as speckle interferometry—a planetlike object near another of the sun's neighbors. Again follow-up observations suggested that the original findings were erroneous.

Undaunted, groups headed by Bruce Campbell of the University of Victoria in British Columbia and by David W. Latham of the Harvard-Smithsonian Center for Astrophysics recently nominated two new candidates for planethood. These objects seem unlikely to vanish outright; the question is whether they are actually planets or more massive, starlike objects.

Like van de Kamp, Campbell and Latham search for wobbles in the positions of stars. But van de Kamp employed astrometry, in which the apparent distance between a nearby star and others so distant that their positions seem fixed is measured. Campbell and Latham instead examine the spectra of stars for Doppler shifts indicative of motion toward and away from the earth. This technique can monitor stars far beyond the range of astrometry, as long as they are bright.

Spectral analysis provides information about motion in only one dimension, however, and so it cannot indicate how the companion's orbital plane is oriented. Consequently the technique can establish only a lower limit for the mass of a companion. The lower limit follows if one assumes the orbital plane is parallel to the line of sight (edge on to viewers on the earth) so that the companion's full gravitational effect can be seen.

After seven years of monitoring a group of stars, Campbell has dis-



Take the two-volume Superstring Theory

for only $4.95

as your introduction to the LIBRARY OF SCIENCE

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

An $89.00 value

Superstring theory is the most promising approach to reconciling general relativity with quantum mechanics. This two-volume work—by three of the field's foremost pioneers—is the classic reference on this fascinating theory and its wide range of applications.

Best of all, you can save over $84.00 on this two-volume set. Complete and return the postage-paid reply card today and we'll send you the $89.00 SUPERSTRING THEORY: Volumes 1 and 2 for only $4.95!

Volume 1: *Introduction*

A self-contained introduction to superstrings, covering everything from an elementary treatment of the bosonic string to supersymmetry, internal quantum numbers, and gauge interactions. Evaluation of tree and one-loop scattering amplitudes, and hexagon anomolies and functional methods are also discussed. Publisher's price: $39.50.

Volume 2: *Phenomenology and Field Theory*

This volume surveys four-dimensional physics that arises by the compactification of six extra dimensions. It also explores string field theory in the light-cone gauge as well as in a covariant and gauge-invariant form. Publisher's price: $49.50.

The Library of Science is the nation's oldest and most respected science book club.

Dedicated to serving both professionals and serious amateurs, it surveys the entire field of published sources on scientific subjects to make the very finest works available to club memebers at substantial savings.

MEMBERSHIP BENEFITS • In addition to getting Superstring Theory, Volumes I and II 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 Science 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.

If reply card is missing, please write to The Library of Science, Dept. 2-CC9-81442, Riverside, NJ 08075, for membership information and an application.

Scientific American 11/88

cerned a tiny, complex fluctuation—a wobble on top of a wobble—in the spectrum of a star called Gamma Cephei. He proposes that Gamma Cephei is under the influence of both a planetlike companion that orbits it every 3.1 years and a distant, unseen star. Campbell estimates that the planet is at least 1.6 times as massive as Jupiter. Other workers say the minute Doppler shifts observed by Campbell might be the result of cyclic, convection-driven fluctuations on the surface of the star. Campbell replies that if convection—a universal process in stars—caused such Doppler shifts, he should have observed similar shifts in more of the stars he has been monitoring.

Latham has also been watching a group of stars for seven years, and he has detected a distinct periodic motion in one star, called HD 114762. The motion has an 84-day cycle, he says, four days shorter than the orbital period of Mercury. Latham says the companion, which is apparently whirling in an extremely tight orbit around HD 114762, is at least 10 times as massive as Jupiter. This lower limit places the object in the range of brown dwarfs, hypothetical objects more massive than a planet but too small to initiate the nuclear burning that fires stars. The object does have one advantage over Campbell's: Latham's spectral observations have been confirmed by another investigator, Michel Mayor of the Geneva Observatory.

It may be that other planetary systems abound. Campbell, Latham and Mayor all have "suggestive" data from as many as half of the stars they have observed. But convincing skeptical astronomers that any particular star has a planet will almost certainly require cross-checking spectral measurements with other techniques, such as speckle interferometry or astrometry, according to George D. Gatewood of the Allegheny Observatory. Gathering such corroborating data is likely to take years. Unless some breakthrough in imaging technology occurs, Gatewood adds, the public should not expect the discovery of earth-size planets to be announced in the foreseeable future. —*J.H.*

Holochrome

New techniques create the best full-color holographic images

For decades the stereoscopic images called holograms have imparted a remarkable impression of depth, but the color has been either monochromatic or distorted. Now investigators at the University of Oxford and the Massachusetts Institute of Technology have created the first holograms that faithfully reproduce the colors of real objects or computer images. The coloring techniques will greatly enhance the artistry and accuracy of holograms and should lead to improvements in medical imaging and computer-aided design.

At Oxford, Paul M. Hubel and Andrew A. Ward combined advanced laser technologies and film materials to create full-color holograms of real objects. Their strategy, first reported in the British journal *New Scientist,* was to create three holograms with three different colors of laser light and then combine the images into one full-color hologram.

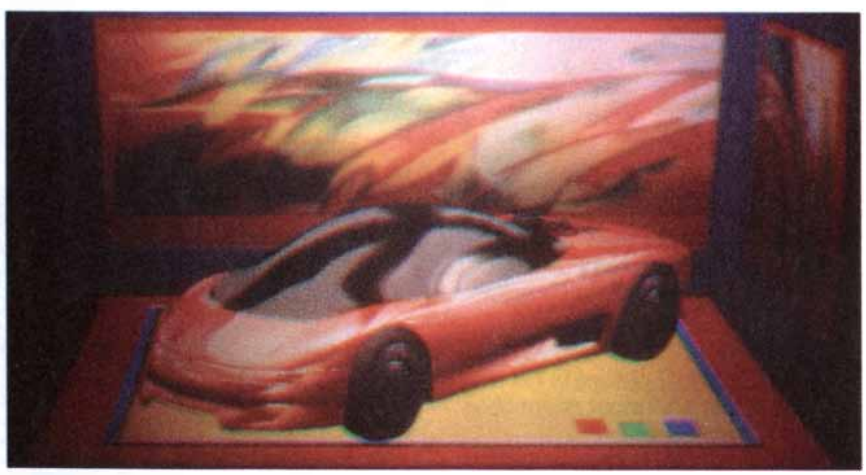
To create a hologram of a coffee mug containing colored pencils, for example, the Oxford investigators began with red laser light at a wavelength of 647 nanometers (billionths of a meter) and a film that was sensitive to red light. As in any hologram, the laser light was split into an object beam and a reference beam. The object beam was aimed at the mug and pencils so that red areas reflected the red light. The reference beam was directed so that it emerged from behind the plane of the film. The light waves from the reference beam and the waves from the reflected beam created an interference pattern, which was recorded in the film. A second exposure was made with green laser light at 528 nanometers and a third one on the same film with blue light at 458 nanometers. The two films were developed. When the films were glued together in the proper register and illuminated with white light, the hologram re-created all the original colors.

At M.I.T., Stephen A. Benton, Julie L. Walker, Wendy J. Plesniak and Michael A. Klug transform objects stored in computers into three-dimensional color images known as stereograms. The stereogram consists of many vertical holographic strips representing successive perspectives of an object; seeing two different perspectives, a viewer perceives depth.

To produce a stereogram of a car, for instance, the M.I.T. team first generated a three-dimensional image of the car on a computer screen and made a photographic slide of the image. The slide was inserted into a device that controls a film holder and a laser. The unexposed film was masked to expose a strip about a millimeter wide. The laser projected the slide onto the film, creating interference patterns like those that would be created if light were reflected from a real car instead of from a computer image. After the first exposure, the computer determined a second perspective of the car, the film was advanced to the next strip and the laser projected onto the film a second slide showing the second perspective. The process continued until the desired number of perspectives were recorded.

To create a color stereogram the

Holograms now accurately represent dimension and color for real objects and computer images



***FULL-COLOR STEREOGRAM** of a computer-designed car is composed of many holographic images that create an impression of depth. Stephen A. Benton and Julie L. Walker developed the process for generating stereograms at the Massachusetts Institute of Technology.*

Product of France. Made with fine cognac brandy. 80 proof.
For gift delivery anywhere call 1-800-CHEER-UP (except where prohibited by law)
©1985 Carillon Importers, Ltd., Teaneck, N.J.

STATE OF THE ART.

The refined shape of the Volvo 780 was carefully arrived at by Bertone, Italy's leading automotive designer.

But since the turbocharged 780 was engineered in Sweden, the attraction goes well beyond sophisticated looks.

Consider, for example, the Multi-link independent rear suspension. Unlike more conventional systems, Multi-link allows each tire to individually adjust to road conditions. So when you encounter uneven road surfaces only one tire, instead of four, is affected. The result is a suspension system that puts comfort *and* handling under one roof.

As one would expect, a car of the 780's class is equipped with a driver's-side Supplemental Restraint System (SRS).

When used in conjunction with our three-point seat belt, SRS is designed to increase protection to the driver in certain types of accidents.

For additional safety, the Volvo 780 is equipped with anti-lock braking (ABS) which continuously regulates the distribution of braking power. Consequently,

ABS allows you to brake with less chance of skidding or losing steering control. And that helps take the panic out of panic stops.

So see us soon to test drive the Volvo 780 personal luxury coupe. It's an engineering achievement even an artist can appreciate.

VOLVO

A car you can believe in.

computer separated the image of the car into basic color components: red, green and blue. A stereogram for the red component was made on the film. Stereograms for the green and then the blue components were made on the same piece of film with red laser light, but the film was chemically altered during the exposures to produce the colors. After the film was developed it revealed a full-color stereogram of the car. The M.I.T. team has applied its technique to create color holographic images of the heart from magnetic-resonance imaging and of bismuth atoms from a scanning tunneling microscope. —*Russell Ruthen*

BIOLOGICAL SCIENCES

Living with Your Self

Investigators trace the fate of self-reactive immune cells

One might expect the immune system to be the body's own worst enemy. Both *B* cells, which secrete antibodies, and *T* cells, which kill foreign cells and regulate the immune defense, recognize their targets by means of surface receptors that are produced in sufficient diversity to bind to almost any large molecule—including components of the body itself. Yet autoimmune reactions, which underlie such diseases as juvenile diabetes, are the exception rather than the rule. How are cells that are targeted to "self" molecules held in check?

Past work has shown that *B* or *T* cells that encounter their antigen, or molecular target, before they have matured become tolerant: they cannot respond to it later. Self-tolerance presumably develops because immature lymphocytes (*B* and *T* cells) bearing receptors for self antigens are certain to meet their targets. But the fate of the self-reactive cells has remained a puzzle: do they die or are they merely inactivated? Ingenious experimental strategies are now yielding an answer. Strangely, it seems to be different for the two classes of lymphocytes.

Workers faced two major challenges in studying self-tolerance. Self-reactive clones (clones are groups of genetically identical cells bearing the same antigen receptor) ordinarily lie hidden from observation among the many millions of *B*- and *T*-cell clones in the immune system. Moreover, until recently the only way to detect a self-reactive clone was by its activity: its ability to secrete antibodies or to activate other immune-system cells. That criterion cannot distinguish between cells that are absent and ones that are merely silent.

A group led by Christopher C. Goodnow and Antony Basten of the University of Sydney recently succeeded in tracing self-reactive *B* cells. The key to their success was a mouse strain genetically programmed to make just one kind of *B* cell, targeted to a single antigen: a novel self protein. To detect the *B* cells directly, the workers enlisted another antibody—a probe molecule that recognized the surface receptor of the self-reactive cells and tagged it with a fluorescent stain.

Writing in *Nature,* the investigators describe how they developed their self-reactive mice. They began by injecting foreign genes into the fertilized eggs of two mouse strains. One strain received the gene for hen-egg lysozyme (HEL), a nontoxic protein that is normally foreign to the mouse immune system. A gene for an antibody to HEL was injected into the germ cells of the other strain. In the offspring carrying the first gene, HEL became a self protein: nearly every cell produced it. In the other transgenic mice, from 60 to 90 percent of the *B* cells carried receptors for HEL, and they made high concentrations of anti-HEL antibody.

The workers then crossed the two strains. A proportion of the offspring inherited both foreign genes—the gene for the antibody and the gene for its protein target. In effect, the entire *B*-cell arsenal of these double-transgenic mice was programmed to be self-reactive. Yet, like normal mice, the double-transgenic mice were self-tolerant: their serum contained very low levels of anti-HEL antibody. What had happened to the self-reactive *B* cells? The probe antibody showed that most of the HEL-specific *B* cells were still present; the cells had merely been silenced.

To find out whether some other element of the immune system, such as so-called suppressor *T* cells, was keeping the *B* cells under temporary restraint, the workers tested the cells' response in isolation. They transplanted 100,000 *B* cells from the double-transgenic mice into mice whose own immune system had been destroyed with radiation. They then challenged the transplant recipients with HEL. The response of the transplanted cells was feeble. The *B* cells of the self-reactive mice had apparently undergone "a persistent functional change."

Analogous studies of *T* cells give a quite different result: self-reactive *T* cells, it seems, are eliminated rather than just silenced. The studies rely on an odd feature of the mouse immune system. Some mouse strains carry self proteins that act as master keys, activating many different clones of *T* cells whose receptors bear common features. As many as a fourth of the receptors generated during *T*-cell development are capable of reacting with these antigens. Two groups, one led by John W. Kappler and Philippa C. Marrack of the National Jewish Center for Immunology and Respiratory Medicine in Denver and the other by H. Robson MacDonald of the Ludwig Institute for Cancer Research in Epalinges, Switzerland, have described the fate of *T* cells that recognize these global antigens.

Equipped with antibody probes that single out the self-reactive receptors, the investigators found that immature cells carrying the receptors were abundant in the thymus gland, where *T* cells mature. Yet mature self-reactive cells were virtually absent from the thymus and the rest of the body. Self-reactive *T* cells, it appears, die out during the maturation process. A group led by Harald von Boehmer of the Basel Institute for Immunology recently observed the same pattern of cell death in transgenic mice engineered to make a single *T*-cell receptor that is specific for a self protein.

Just what kills or silences self-reactive lymphocytes? For *T* cells the Basel group reports a hint: cells that expressed the self-reactive receptor but lacked two surface proteins characteristic of *T* cells, CD4 and CD8, escaped deletion, which suggests that those markers also play a part in cell death. As for the silencing of *B* cells, the Sydney group found a clue in their antigen receptors. *B* cells make various classes of surface receptors for the same antigen. Goodnow and his colleagues discovered that receptors of one class, known as IgM, were unexpectedly scarce on the silent *B* cells of their double-transgenic mice.

The transgenic-mouse strategy is well suited to following up on these hints, because it enables experimenters to vary the characteristics of the receptor and the self antigen at will. To study the possible role of each receptor class in *B*-cell silencing, for example, the Goodnow group now plans to make transgenic strains capable of displaying only a single class of receptors on their self-reactive cells.

Why should the fate of self-reactive *B* and *T* cells differ in the first place? Marrack points out that the difference

TWA's Business Class. More cities. More flights. More often.

If you like flying business class—and who doesn't—TWA's Ambassador Class® was made for you. It's the only business class with over 700 domestic departures each week. And it's the only one that flies to more than 18 major U.S. business centers.

You'll enjoy all the amenities you've been missing on other airlines. A separate Ambassador Class cabin, so you can work or sleep, free of distractions. Seats so wide, with legroom so ample, they rival first class seats on many other airlines. Complimentary cocktails and headsets. Meals served on fine china and linen.

With TWA's Ambassador Class, you're treated to more service, more comfort, and more convenience. A lot more often.

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

TWA

What should every company demand from a computer system?

Growth.

Introducing the IBM Application System/400.

When your business is small, you can buy an IBM® Application System/400,™ and it will be just the right size.

Later on, you'll still be smiling.

That's because as your company grows, your Application System/400 can grow right along with you. And the investments you made at first–in software, training, and peripherals–will still be working for you.

That's what the IBM Application System/400 is all about. It comes from IBM's leadership with over a quarter million mid-size computer systems in place, and it does what growing companies have told us they want.

It lets you grow into what you need, without outgrowing what you've paid for.

Today: Solutions for your business, from the leader in business solutions.

Never before has a mid-size computer system been introduced with so much proven software

ready to go. Thousands of programs that run on IBM's System/36 and /38 can run on the IBM Application System/400.

Even better, any program you start with, you can stay with. No matter how big your Application System/400 gets, your software will always work.

Tomorrow: Protection for your investment as your business grows.

Most models of the IBM Application System/400 come rack-mounted like a stereo system. Components slide in and out, so you can upgrade right on the spot. You're not locked into any particular setup. Your system's only as big as you need, and making it bigger is easy.

What's more, the Application System/400 comes with a state-of-the-art education system, plus customer support that's unmatched in the industry.

For a free brochure, or to arrange for a detailed discussion about the Application System/400, call 1-800-IBM-2468, ext. 82.

IBM The Bigger Picture

is at least consistent with what is known about the immune response. *B* cells cannot secrete antibody efficiently without help from *T* cells that have recognized the same antigen. Since *T* cells are the linchpin of the immune response, Marrack suggests, one would expect them to be more tightly controlled than *B* cells. Goodnow adds that because antibody genes mutate rapidly, unlike the genes for *T*-cell receptors, it might pay to spare a self-reactive *B*-cell clone, since a mutation might eventually make the cells useful against some foreign antigen. —*Tim Appenzeller*

Masters of Mutation

Experiments suggest mutations are not always random after all

Few recent findings have attracted such widespread interest as two reports that attack a fundamental principle of modern biology: the idea that mutations arise at random.

Conventional wisdom holds that an organism cannot specifically produce mutations that are advantageous. Rather, some mutations happen purely by chance to be beneficial and so become evident. This idea emerged from experiments with bacteria done in the 1940's and 1950's. One now classic study was carried out in 1943 by Salvador E. Luria and Max Delbrück. They counted the numbers of mutant bacteria capable of resisting infection in several identical and rapidly growing cultures that were suddenly exposed to bacteriophages (bacterial viruses). The numbers of mutants fluctuated widely among the cultures, suggesting that the cultures included some lineages that had already been carrying mutations preadapting them to resist infection. (If the mutations had arisen in direct response to the bacteriophages, they would presumably have been more evenly distributed among the cultures.) The "fluctuation test" and other experiments demonstrated that mutations arise spontaneously.

John Cairns and his colleagues at the Harvard School of Public Health now point out in *Nature* that the classic experiments were not well designed to search for additional mutations that might occur in response to a new need. Cairns and his co-workers therefore repeated the classic test, but with a strain of the common bacterium *Escherichia coli* that cannot metabolize the sugar lactose. The investigators grew several cultures and then tested them for mutants that had reacquired the ability to grow on lactose and so could form colonies.

They found, in addition to the expected preexisting mutants, certain late-occurring ones that were evenly distributed among cultures, suggesting that they had indeed arisen in response to the sudden availability of lactose. The workers confirmed this by showing that these extra mutations could be delayed by delaying the addition of the sugar. Moreover, they found, the sugar's presence specifically elicited advantageous mutations conferring the ability to metabolize it. The classical view cannot explain these extra mutations. Cairns says "it is extraordinary that the conventional wisdom has been based on so little evidence."

In another experiment he and his collaborators showed that when one strain of *E. coli* is required to use lactose for growth, it can quickly delete a specific genetic fragment that blocks lactose metabolism. In a third experiment *E. coli* colonies exposed to lactose brought about two rare mutations within days that together activated a "cryptic" (normally unused) gene conferring the ability to metabolize lactose.

Barry G. Hall of the University of Connecticut has done related experiments. He finds that two mutations, normally rare, that together allow *E. coli* to utilize a nutrient called salicin occur in close succession and at a high rate in colonies of *E. coli* when they are fed a diet of little except salicin. His results will be described in the journal *Genetics*. Another experiment done by Hall indicates that the frequency of one beneficial mutation increases fiftyfold under strongly selective conditions.

The results "suggest that cells have some way of recognizing what would be an advantageous mutation and increasing the chance that it occurs," Hall says. He points out that bacteria seem to have many cryptic genes. Since starvation is presumably a common situation for bacteria, he thinks they might activate cryptic genes by directed mutation as the need for the genes arises.

The mechanism of this startling ability—if indeed the ability is confirmed—is for now completely unknown. It invites comparisons with the long discredited views of Jean-Baptiste Lamarck, the 19th-century biologist who held that organisms pass on to their offspring characteristics acquired during their lifetime (so that, for example, the son of a blacksmith who developed strong arms through his work would in turn be endowed with strong arms). Hall cautions that bacteria are quite different from higher organisms, however; a bacterium's genes are the very ones it passes on to its offspring, whereas multicellular organisms sequester their genetic legacy in distinct germ cells. Even if directed mutation is found to be widespread in bacteria, it may therefore be impossible in higher organisms. Still, Cairns and Hall agree that one would expect evolution to favor the preservation of such a useful trick. —*T.M.B.*

Acquired Taste

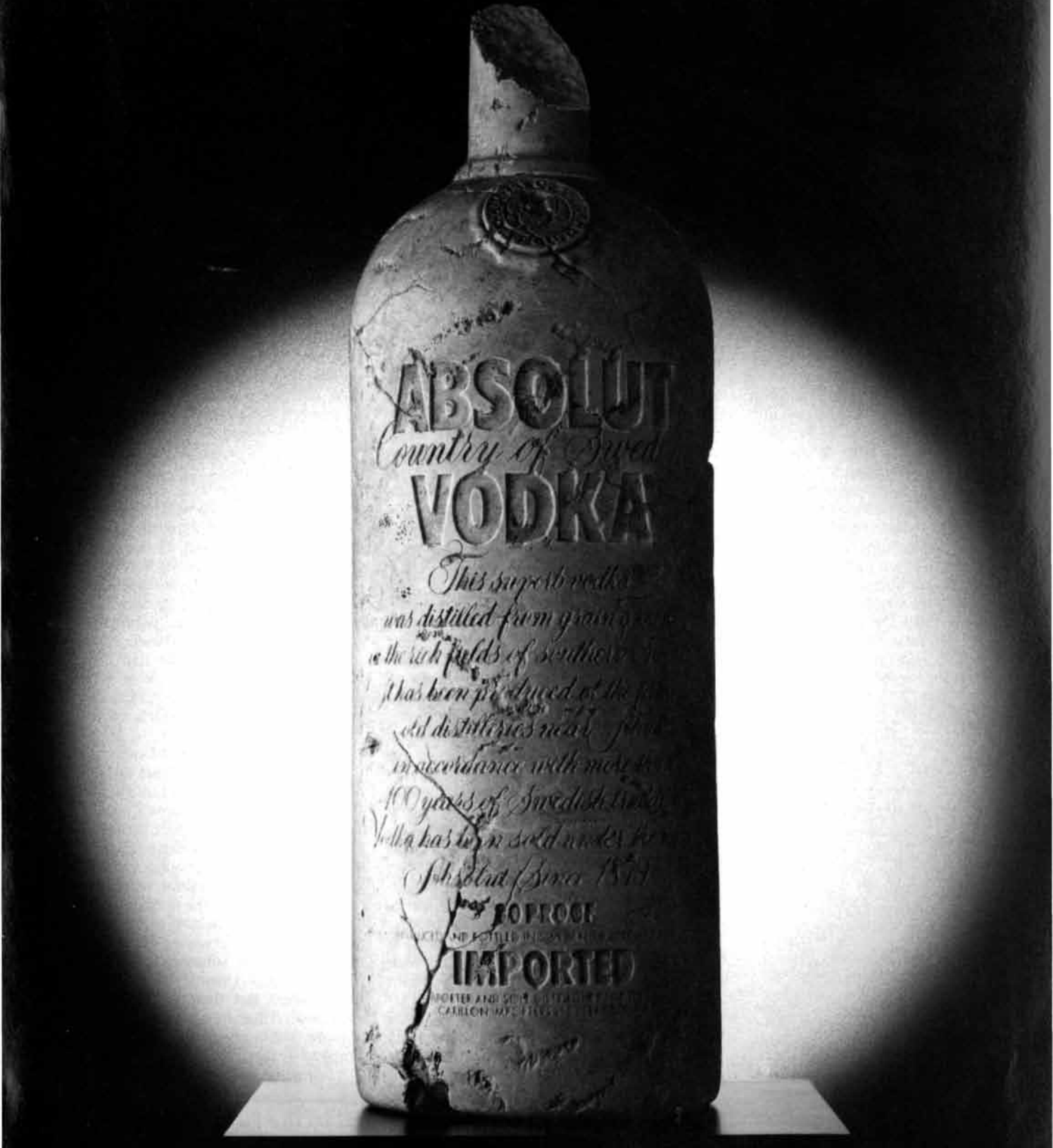
Now the pear thrips is eating New England's sugar maples

Fall foliage and maple syrup, two traditional glories of New England older than the Republic itself, are threatened by an insect no bigger than the head of a pin. Across much of Vermont, New Hampshire, Massachusetts and Connecticut and parts of New York and Pennsylvania, hundreds of thousands of maple trees, many of them more than 200 years old, are falling prey to *Taeniothrips inconsequens,* the minute insect commonly known as the pear thrips.

Normally found in small numbers on fruit trees, the pear thrips has recently acquired a taste for the sugar maple, *Acer saccharum*. The insect inflicts its damage early in the spring, when adults emerge from the soil and migrate upward to feed on the blossoms, leaf buds and young leaves of nearby maples. Telltale signs of infestation are shriveled, tattered leaves and bare branches, many of them in the lower crown of the tree.

Thrips, first associated with sugar maples about 10 years ago, have spread wildly in the past year, causing damage to an estimated half million acres of maple trees in Vermont alone. In some parts of the state soil samples reveal population densities as great as 50 million immature thrips per acre. Infested trees produce less sap than healthy ones and are unlikely to bear foliage as brilliant as in years past.

No sugar maples are believed to have died from thrips damage yet, and many that lost their leaves early in the season were able to produce a second set. It is the long-term stress imposed by the pear thrips that has many experts worried. Dale Bergdahl, a forest pathologist at the University of Vermont, speculates that the trees must have been stressed by unknown envi-



ABSOLUT ORIGINAL.

FOR GIFT DELIVERY OF ABSOLUT VODKA ANYWHERE, CALL 1-800-CHEER-UP (EXCEPT WHERE PROHIBITED BY LAW).
80 AND 100 PROOF VODKA/100% GRAIN NEUTRAL SPIRITS (ABSOLUT COUNTRY OF SWEDEN™) ©1988 CARILLON IMPORTERS LTD. TEANECK, NJ.

ronmental factors—perhaps excessive heat, reduced snow cover or atmospheric pollutants—to have become vulnerable to the pear thrips in the first place. The damage caused by the thrips will no doubt further weaken the trees, rendering them susceptible to an onslaught of opportunistic pests such as fungi and other insects. This secondary wave of infestation may be what ultimately sounds the death knell for the tree. "It is the long-term, progressive effects of infestation, the origins of which are poorly understood, that we fear the most," Bergdahl observes.

For that reason some maple-sugar producers are being asked to not tap their trees next year. The fewer leaves a tree has, the less sugar-laden sap it can produce. Because sap is essential to the tree's ability to produce leaves the following year, Bergdahl says "tapping a stressed sugar maple is comparable to taking blood from an anemic."

In July, Governor Madeleine M. Kunin of Vermont appointed a technical task force to investigate the pear-thrips problem. The members of the task force agree that little can be done until the biology and life cycle of the pear thrips is better understood. "Our primary goal," says Margaret Skinner of the University of Vermont, "is to develop a statewide sampling and monitoring system. We need to know where the thrips are, how they migrate and how their numbers fluctuate from year to year."

A group from the University of New Hampshire's Institute for the Study of Earth, Oceans and Space will compare infrared Landsat satellite tapes made four years ago (a healthy year for the sugar maple) with tapes made this year. The group hopes to pinpoint the destruction caused by the pear thrips to get a better idea of its spread.

Is the sugar maple likely to go the way of the American elm, all but disappearing from the landscape? No one knows, but it is certain that as early as next April a new generation of hungry thrips will emerge from the soil looking for budding maple leaves on which to feed. —*Laurie Burnham*

MEDICINE

Starvaholics?

Anorexics may be addicted to a starvation "high"

People who have anorexia nervosa typically lose a fourth of their weight or more by dieting and exercising and often by inducing vomiting or taking diuretics and laxatives. What can account for such extreme self-destructiveness? Most theories offer psychological explanations and say in part that anorexics—chiefly teen-age girls and young women—are perfectionists who become convinced they are fat and ugly.

Several investigators have suspected that even if a psychological problem accounts for the onset of anorexia, physiological mechanisms help to perpetuate the disorder in the estimated 50 percent of patients in whom it becomes chronic. One specific theory, advocated by Mary Ann Marrazzi and Elliot D. Luby of the Wayne State University School of Medicine, suggests that anorexia is perpetuated by an addiction to starvation. In particular, they say, prolonged starvation is thought to elicit the release of certain endogenous (self-produced) narcotic-like substances known as opioids. The opioids have several effects, including the slowing of the metabolism to conserve energy. They also induce an elation that could well reinforce anorexic behaviors, much as an alcoholic "high" reinforces the desire to drink. If Marrazzi and Luby are correct, their hypothesis may lead to new therapies.

The investigators began to consider that chronic anorexia might be an addiction to the body's own opioids because they knew that food deprivation leads to opioid release in many animals and that patients with anorexia nervosa had been found to have elevated levels of opioids in their cerebrospinal fluid. Moreover, anorexics resemble alcoholics and drug addicts in many ways. Anorexics are compulsive, often centering their lives on fasting and other anorexic behaviors just as drug addicts center their lives on getting their next "fix." They typically deny they have a problem or, if they do admit it, report feeling driven by an almost demonic force beyond their control. Anorexics also have to be forced into treatment when it is indicated, and many try to continue their self-destructive behavior even while hospitalized.

The opioid hypothesis has not yet been tested extensively, but Marrazzi and Luby have treated eight anorexic patients with opioid blockers and found that the medications helped six of them to gain weight. One long-term anorexic has remained at a normal weight for more than two years, they report, and has also abandoned the 10 to 30 laxative tablets she had been consuming daily. The workers have now undertaken a double-blind clinical trial, the results of which are expected in about a year.

An apparently paradoxical aspect of the theory is that endogenous opioids produce not only elation but also intense hunger, which normally overrides the elation and drives people to eat. Marrazzi and Luby suspect, however, that those who develop intransigent anorexia have an abnormal physiological response to the opioids—that they are not as sensitive to the hunger as other people would be, at least most of the time. Sometimes the hunger does break through and the patients binge. But then, in their constant search for an opioid high, they resume starving themselves, sometimes to death. —*Ricki Rusting*

All in the Family

Inheritance may have a role in most colorectal cancers

Almost all cases of cancer affecting the colon or rectum may be due to an inherited genetic susceptibility, according to a study recently published in the *New England Journal of Medicine*. The study's preliminary results suggest that such susceptibil-

128 formulas at the speed of light.

The Casio FX-5000F Formula Calculator. To save you the time and trouble of looking up or memorizing many of the most important mathematical and scientific formulas, we put 128 of them in our FX-5000F Formula Calculator. And you can call them up in a flash.

The formulas are numbered and cover the fields of *math, statistics, physics, electronics* and *mechanics.* Plus you can input 12 of your own. Just key-in the appropriate number and the formula you need appears instantly. The alpha numeric display with 10 digit mantissa plus 2 digit exponent is easy to read and scrolls to over 70 characters. Its two-line display shows both the formula and the answer simultaneously.

And it doesn't stop there. Once you've recalled the formula, the calculator prompts you to input the values of the variables and automatically calculates the result.

The FX-5000F's 675 steps of program memory allow you to create some pretty sizeable programs and store them until needed. While an instant replay feature lets you review and edit formulas at the touch of a button.

Adding to its usefulness are an additional 160 powerful scientific functions, for a combined total of 288 functions and formulas.

Get your hands on a Casio FX-5000F and flash through a few formulas. You'll find it very illuminating.

CASIO®
Where miracles never cease

Casio, Inc. Consumer Products Division: 570 Mt. Pleasant Avenue, Dover, NJ 07801
Casio Canada Ltd., 2100 Ellesmere Road, Suite 240, Scarborough, Ontario M1H3B7

ity is present in a third of the population; better methods of identifying the susceptible subgroup could lead to a sharp decrease in mortality from the disease in the future. In the U.S. colorectal cancers are the third leading cause of death from cancer; it is estimated that 147,000 new cases will be diagnosed in 1988.

One reason the new findings are so striking is that cancers of the colon and rectum have long been associated with environmental risk factors, particularly high dietary intake of fat and low intake of fiber. Previous studies have shown that close relatives of people with colorectal cancer have an increased risk of developing the same condition, but no clear pattern of genetic inheritance could be seen. Consequently it was thought the increased risk might be due to shared lifestyles in a family or to other environmental factors.

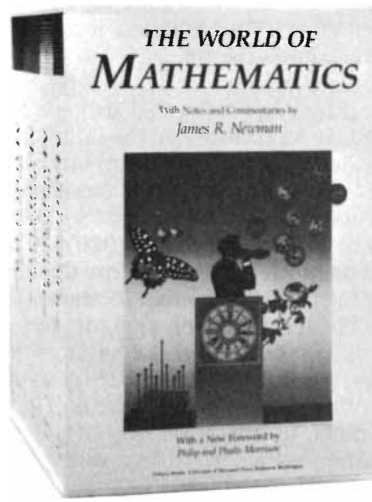
The current study was able to clarify the picture by establishing not only the frequency of colorectal cancer but also that of adenomatous polyps: benign growths thought to be the precursors of all, or almost all, colorectal cancers. After individuals with polyps or colorectal cancer were selected, their close genetic relatives were examined for the presence of either condition; spouses of the patients served as controls. When the results were analyzed, a clear pattern of inheritance emerged.

The most probable explanation for the pattern is that there exists a gene the inheritance of which renders an individual susceptible to developing polyps. The gene is "dominant": if a person has one copy of the gene, the susceptibility will be present. Computer modeling of the study sample (which consisted mainly of white Utah residents whose ancestry was British or northern European) suggests the gene is present in 32 percent of the population.

It is clear that this new "susceptibility" gene is not the only factor leading to the development of colorectal cancer. Only about 5 percent of adenomatous polyps take the next step and become transformed into cancers. According to Randall W. Burt of the University of Utah Medical Center, one of the study's authors, the other factors may well include interactions between the susceptibility gene and the environmental conditions that have long been implicated in the development of tumors of the colon and rectum.

The implications of the study for saving lives may ultimately be quite significant. Colorectal cancer is "very

A Classic Returns



"Many anthologies have been designed to introduce intelligent laypersons to the scope, power, significance, and beauty of mathematics, but none has come even close to Newman's in meeting this objective. THE WORLD OF MATHEMATICS is a superb collection."

Martin Gardner, contributing editor to Scientific American

Out of print for more than a generation, THE WORLD OF MATHEMATICS is, happily, available once again. This rich and spirited anthology presents an outstanding picture of the diversity, intellectual excitement, impact, even humor of mathematics through the centuries. James R. Newman has selected 133 classics of mathematics literature, from benchmark works that have shaped civilization to lesser-known works that explore the lighter side of mathematics.

There is something for everyone—no matter what your level of mathematical sophistication. You'll find Galileo's hard-to-find work on motion, Euler's famous topology puzzle, A.M. Turing's test for the presence of machine thought, Hermann Weyl's brilliant piece on symmetry, and Lewis Carroll's charming logic riddles. The essays are prefaced by Newman's perceptive commentaries. These delightful and astute introductions place the essays in historical context and make even the most complex mathematical concepts accessible to a wide range of readers.

THE WORLD OF MATHEMATICS. A timeless classic.

THE WORLD OF MATHEMATICS

James R. Newman

With a New Foreword by Philip and Phylis Morrison.

4-volumes, 2800 pages.
Each volume is 6 x 9.
Available in both hardcover and softcover boxed sets.

Tempus is an imprint of Microsoft Press.

THE WORLD OF MATHEMATICS includes selected writings from:

Archimedes ■ Daniel Bernoulli ■ George Boole ■ René Descartes ■ Clement V. Durell ■ J.B.S. Haldane ■ Edmund Halley ■ Aldous Huxley ■ John Maynard Keynes ■ Morris Kline ■ Philippe Le Corbeiller ■ Thomas Robert Malthus ■ Gregor Mendel ■ Isaac Newton ■ George Bernard Shaw ■ A.M. Turing ■ Plutarch ■ Henri Poincaré ■ Bertrand Russell ■ John von Neumann ■ Alfred North Whitehead

At your favorite bookseller. Or order directly from Tempus. To order, fill out the coupon below. Or you can place a credit card order by calling **800-638-3030** (In MD, call collect 824-7300). 8AM to 4PM EST. Refer to campaign MATH88.

YES, please send me:

_____ hardcover boxed sets at $99.95 each [86-96924] ..... $_____

_____ softcover boxed sets at $50.00 each [86-96932] ..... $_____

Sales tax: CA residents 6.5%; WA state residents 8.1%; MA and MD residents 5% ..... $_____

UPS & Handling: $7.50 per set ..... $_____

TOTAL ..... $_____

Name _____

Address _____

_____

City _____ State _____ ZIP _____

Payment:

☐ Check/Money Order ☐ MasterCard ☐ VISA ☐ AmEx

Credit Card No. _____ Exp. Date _____

MATH88

Send your order to: Tempus Books, P.O.Box 1532, Hagerstown, MD 21742. Order by November 25th to ensure holiday delivery.

treatable," Burt said, if it is caught early enough. Because the cancer produces few symptoms as it develops, however, currently only about 50 percent of colorectal cancers are diagnosed early enough for the patient's life to be saved through surgery. At present early detection is accomplished by regular screening for colonic polyps and blood in stools. Efforts at early detection are likely to be much more successful if screening programs can be directed at susceptible individuals. "One day," Burt said, "we hope to find a genetic marker to identify the susceptible population," a discovery that might greatly reduce mortality. —*John Benditt*

OVERVIEW

Developmental Dialectics

An intricate hierarchy of genes seems to organize the organism

Every cell in an organism, with few exceptions, carries copies of all that organism's genes. Somehow each cell is able to activate only the genes it needs: a muscle cell becomes elongated and produces myosin, whereas an epidermal cell becomes flattened and makes keratin.

DEVELOPMENTAL ROLE of the gene* Xhox1A *is demonstrated in a frog embryo. At the two-cell stage one cell was injected with an excess of the gene's RNA. Disorganized blocks of muscle* (left) *developed on the treated side.

How are the cells of a developing embryo instructed to assume the specialized forms and functions appropriate to their type and site? Although the nature of the molecules responsible for the signaling is still largely unknown, the techniques of molecular biology are beginning to reveal how genes direct the process.

Much research in the area has concentrated on the familiar fruit fly, *Drosophila melanogaster.* The adult fly is divided into some 19 segments along its length, some bearing characteristic structures such as wings, legs or antennae. Since the early 1970's it has been known that even before the ellipsoidal *Drosophila* embryo shows any visible specialization along its length, incipient cells in the embryo somehow become "determined," or programmed to become part of particular regions in the adult fly. This was learned either by transplanting fragments of the embryo from one place to another or by destroying fragments. Often such manipulations caused a structure to be misplaced or absent in the adult. It was clear that information about a cell's location is conveyed to it early in development.

How? A major clue to the nature of the genetic control of development was provided by the so-called homeotic genes. A fly carrying a mutation in such a gene characteristically has a segment bearing the wrong kind of structures. The classic example is the *Antennapedia* gene: a mutation in it puts legs on the head of a fly instead of antennae.

The study of homeotic mutants led Walter J. Gehring of the University of Basel and others to discover that an almost identical short sequence of DNA subunits, called the homeobox, is found in virtually all homeotic genes, of which dozens are now known. Significantly, the homeobox appears to specify a short stretch of protein shaped so that it binds to DNA. This suggested that proteins made by homeotic genes might control other genes by binding to critical regions of DNA that govern the time when those genes are activated. Homeotic genes thus seemed likely to be switches that turn on the array of genes needed by a particular segment.

This of course begs the question of how the borders of the segments are themselves established and how the right homeotic genes are activated in each segment. Recent experiments are starting to sketch the outlines of a system that appears to be able to do that job.

To begin with, several other classes of genes are known that affect development in particular ways. Mutations in most of these genes kill the embryo, but not before some abnormality can be seen. One class of genes that operates very early in development establishes a distinction between the anterior and posterior ends of the embryo. These positional genes are "maternal effect" genes: in the mother rather than in the embryo. Important genes of this type are *bicoid* and those in a complex called *oskar.* Females that have a defective *bicoid* gene produce embryos with no head and no thorax; females with mutant *oskar* genes produce embryos with no abdomen. Analysis of such mutations suggests maternal genes lay down substances that somehow "label" the two ends of the egg even before it is fertilized.

Other genes that affect embryonic development belong (like the homeotic genes) to the embryo itself, and they have diverse effects. Segmentation genes, for example, can determine how many segments develop. Some segmentation mutants, the so-called gap mutants, lack an entire group of adjacent segments. An example is *hunchback,* which lacks several head segments. Other segmentation genes result in "pair rule" mutants, in which alternate segments are affected. *Fushi tarazu,* which is Japanese for "not enough segments," is one pair-rule gene: mutations in the gene cause alternate segments to fuse with the one in front of them, so that the embryo has only half the proper number of segments.

Where in the embryo is which gene active? Techniques are now available that allow the investigator to label a particular developmental gene's RNA transcript and try to answer that question. With such techniques investigators including Diethard Tautz of the Max Planck Institute for Developmental Biology in Tübingen have begun to trace a complex web of positive and negative regulatory interactions between different types of segmentation genes, positional genes and homeotic genes. For example, the product of the gene *bicoid* activates *hunchback* but inhibits another gap gene, *Krüppel.* Moreover, some genes of the same general type inhibit one another.

The picture that emerges is of a control hierarchy, with genes at each level passing instructions down the chain. After maternal genes have established an anterior/posterior gradient in the embryo, different types of embryonic genes are activated in sequence, each reacting to the activity of the level above (and probably other

levels) and specifying in increasing detail how the cells within its domain should develop. The overall effect is to divide the embryo into progressively narrower subdivisions until each subdivision is only half a segment, or a few cells, in width.

It seems likely that each subdivision may eventually have a unique pattern of active segmentation genes; this pattern could in theory specify in combinatorial fashion which homeotic gene or genes to activate and could tell the cells which half of a segment they are in. Many of the genes in the regulatory cascade incorporate the homeobox sequence; others have another stretch of DNA encoding the finger domain, which also binds to DNA and is likely to operate in a similar manner. There are indications that a similar arrangement could specify pattern in the dorsal/ventral direction.

Until this year there was no direct proof that the product of a positional gene could actually specify the developmental fate of a cell; that was merely inferred from the observed patterns of gene activity in mutants. Wolfgang Driever and Christiane Nüsslein-Volhard, also at Tübingen, have now obtained such proof for the protein produced by *bicoid*. They found a way to manipulate the abundance and distribution of the protein product of *bicoid* in *Drosophila* embryos. When they did so, the developmental fates of cells were subsequently altered, distorting the embryos in predictable ways. The *bicoid* protein appears to be the first proved example of an early embryological morphogen: a substance that acts on cells directly and so determines their developmental fate.

Perhaps the only other established morphogen is retinoic acid, which operates at a much later stage and is known to affect the development of chicken wings. It is much harder to manipulate genes in vertebrates than in invertebrates, and so understanding of the genetic control of vertebrate development lags far behind what has been achieved in *Drosophila*. Yet William McGinnis of Yale University and others have identified homeobox sequences in many other animals that have a basically segmented body plan, including earthworms, frogs, chickens, mice and humans. It therefore seems likely that homeoboxes play an important role in many species.

Richard P. Harvey, now at the Walter and Eliza Hall Institute of Medical Research in Melbourne, and Douglas A. Melton of Harvard University recently obtained the first good evidence for a developmental effect of a vertebrate gene that contains a homeobox sequence. By injecting half of a developing embryo of the frog *Xenopus laevis* with the RNA transcript of a gene known as *Xhox1A*, they could disrupt the spatial pattern of developing blocks of muscle. The tissue appeared to be normally constituted otherwise. Melton has also found a gene that is active in the developing frog embryo in specific spatial areas, a pattern reminiscent of segmentation genes in *Drosophila*. These results strengthen confidence that mechanisms like those being uncovered in *Drosophila* may operate in vertebrates.

Although developmental cascades operating along these lines appear to be important in subdividing a growing embryo into segments and thereby providing "landmarks" for developing cells, there are likely to be other mechanisms as well that communicate information between cells over very short distances. A few tantalizing findings suggest that molecules resembling growth factors may play a role. If the resemblance is not coincidence, this immediately suggests how the cells might be communicating: growth factors bind to receptors on a cell's surface, and so initiate a chain of events that can include the alteration of gene expression. —*Tim Beardsley*

SCIENTIFIC AMERICAN November 1988 Volume 259 Number 5

Educating Poor Minority Children

Schools must win the support of parents and learn to respond flexibly and creatively to students' needs. A successful program developed in New Haven points the way

by James P. Comer

Thomas Jefferson and other advocates of free public schools believed fervently that an educated populace is the lifeblood of democracy. In their view the school clearly had a political purpose: to socialize children to become good citizens. Jefferson wrote, "I know no safe depository of the ultimate powers of the society but the people themselves; and if we think them not enlightened enough to exercise their control with a wholesome discretion, the remedy is not to take it from them, but to inform their discretion."

It is a long fall from this lofty ideal to the grim reality facing youths at the margins of today's society. Poor minority children are undereducated in disproportionate numbers across the country. Academically such children may lag behind the national average by up to two years. In large cities as many as 50 percent of minority children drop out of school. The failure to educate these children makes ever harder the task of rectifying economic and social inequities. Job opportunities increasingly reside in service and technology industries, but poor minority youths are the least likely to have the social and academic skills these jobs demand. Unless schools can find a way to educate them and bring them into the mainstream, all the problems associated with unemployment and alienation will escalate.

JAMES P. COMER is professor of child psychiatry and director of the School Development Program at Yale University's Child Study Center and associate dean of the Yale University School of Medicine. He received an A.B. at Indiana University, an M.D. from the Howard University College of Medicine and an M.P.H. from the University of Michigan School of Public Health. He writes and consults extensively on school improvement. In his most recent book, *Maggie's American Dream,* which will be published this month by New American Library, Comer writes about his mother, an impoverished black woman who grew up in the rural South.

The task seems overwhelming. And yet it can be done. In 1968 my colleagues and I at Yale University's Child Study Center started an intervention project at two inner-city schools in New Haven. Unlike many of the reforms that are now being tried or proposed, which focus on academic concerns such as teacher credentials and basic skills, our program promotes development and learning by building supportive bonds that draw together children, parents and school. By 1980 academic performance at the two New Haven schools had surpassed the national average, and truancy and disciplinary problems had declined markedly. We have now begun to duplicate that success at more than 50 schools around the country.

The perceptions underlying our approach are partly rooted in my own childhood. In 1939 I entered an elementary school in East Chicago, Ind., with three other black youngsters from a low-income community. The school was considered one of the best in the district; it was racially integrated and served the highest socioeconomic group in town. All four of us were from two-parent families, and our fathers made a living wage in the local steel mill. We were not burdened by any of the disadvantages—school segregation, inadequate schools, single-parent families, unemployment—commonly cited as causes of educational underachievement in poor black children. Yet in spite of the fact that we had similar intellectual potential, my three friends have had difficult lives: one died prematurely from alcoholism, a second spent a large part of his life in jail and a third has been in and out of mental institutions.

Why did my life turn out better? I think it was largely because my parents, unlike those of my friends, gave me the social skills and confidence that enabled me to take advantage of educational opportunities. For example, I became friendly with my third-grade teacher, with whom I would walk hand in hand to school every day. My parents took me to the library so that I could read many books. My three friends, however, never read books—which frustrated and angered their

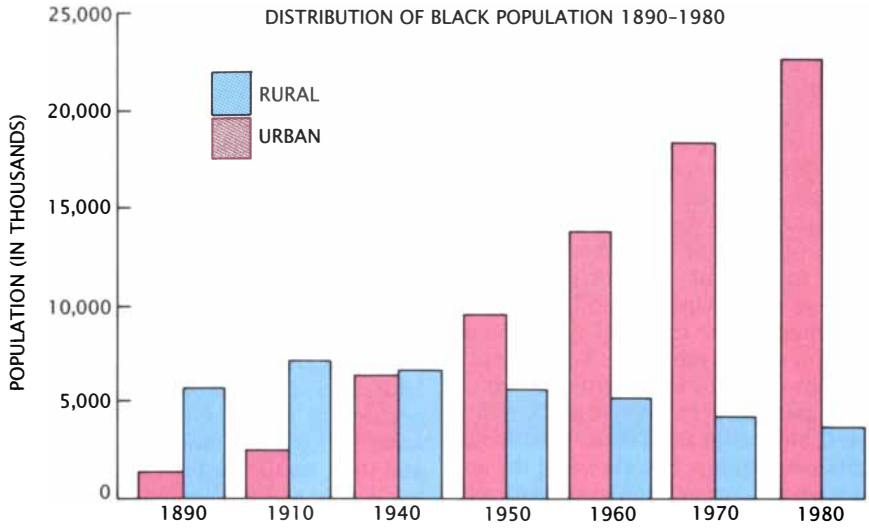
teachers. What the teachers did not realize was that their parents were afraid to go to the library; indeed, they were uncomfortable around white people in general and avoided them.

In the 1960's I began to speculate that the contrast between a child's experiences at home and those in school deeply affects the child's psychosocial development, and that this in turn shapes academic achievement. The contrast would be particularly sharp for poor minority children from families outside the mainstream. If my hunches were correct, then the failure to bridge the social and cultural gap between home and school may lie at the root of the poor academic performance of many of these children.

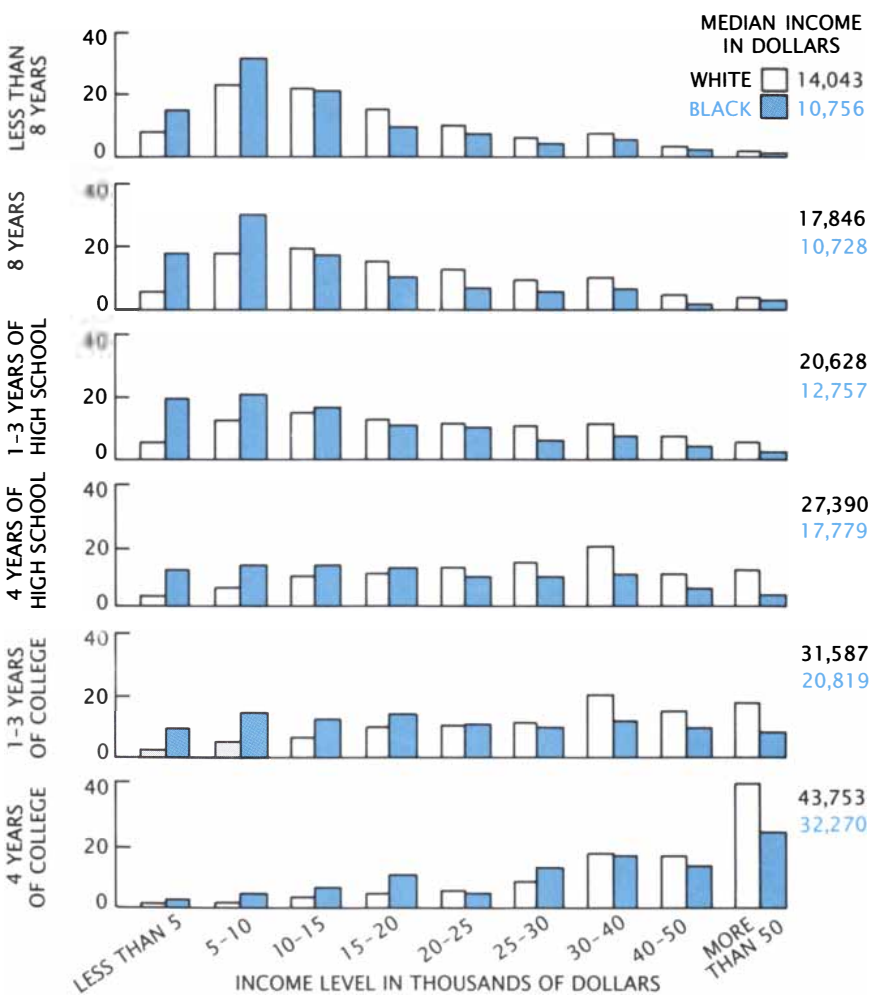
Yet current educational reforms deemphasize interpersonal factors and focus instead on instruction and curriculum. Such approaches reveal a blind spot: they assume that all children come from mainstream backgrounds and arrive at school equally well prepared to perform as the school expects them to. Reading, writing, arithmetic and science are delivered to students in much the same way as tires, windows and doors are attached to the frame of an automobile on an assembly line. Yet students do not come in standardized frames that passively receive what is delivered. Most educators do not challenge this assumption, however, and the approach has never been systematically evaluated or modified through direct experiments in schools.

In contrast, Albert J. Solnit and his colleagues at Yale's Child Study Center believed educational reformers should develop their theories by directly observing and intervening in schools over long periods of time. Solnit's ideas inspired the school-intervention research project that was begun by the center and the New Haven school system in 1968 and continued until 1980. I was asked to direct the project and to work with a social worker, a psychologist and a special-education teacher from the center. We decided to immerse ourselves in the schools to learn how they function and then, on the basis of our findings, to develop and implement a model for improving the schools. We were guided by our knowledge of public health,

SMILE REWARDS A JOB WELL DONE in teacher Nancy McManus' first-grade class at the Katharine Brennan School in New Haven, Conn. The school, one of two elementary schools in which the author and his colleagues intervened, serves a nearby low-income housing project that suffers from rampant joblessness and crime. The staff is alert to the special needs of developing children, particularly those from marginal homes, and works hard to get parents involved; 92 percent of the parents visited the school 10 times or more in the past year. The school is now academically one of the best in the city.



BLACK MIGRATION into urban communities (towns of more than 2,500 people) accelerated after World War II. The rural black population once greatly outnumbered the urban population, but the postwar economic boom led large numbers of blacks to move to the cities in search of jobs. Discrimination and lack of adequate education, however, denied many blacks access to the primary urban job markets.



EDUCATION LEVEL correlated to family income for whites and blacks shows that black incomes are at least a third less than the incomes of whites with equivalent schooling. For a given level of schooling completed by the family head, the graph indicates the percentage of families within each income bracket. Lack of education reduces the income of blacks more than that of whites, and black income rises more slowly with education than white income does. The data, for 1984, are from *Statistical Abstract of the United States 1987,* issued by the U.S. Department of Commerce.

human ecology, history and child development—and by common sense.

Our model evolved in two schools: the Martin Luther King, Jr., School, which had about 300 pupils from kindergarten through fourth grade, and the Katharine Brennan School, which had more than 350 pupils from kindergarten through fifth grade. The pupils were 99 percent black and almost all poor; more than 70 percent were from families receiving Aid to Families with Dependent Children. At the beginning of the project the pupils were ranked near the bottom in achievement and attendance among the 33 schools in the city. There were serious problems with attendance and discipline. The staffs were discouraged; their turnover rate was 25 percent. Parents were dejected, distrustful, angry and alienated.

Both staff and parents approached the first year of the project with high expectations. But because teachers and administrators could not agree on clear goals and strategies, we had a difficult school opening. Some new teachers tried to have open classrooms, but the children soon became uncontrollable. Teachers blamed the administration for not providing adequate resources, and parents became angry—angry enough to march on one of the schools. Needless to say, the students did not learn much.

We, on the other hand, learned a great deal. The spectacular deterioration of the schools illuminated their social dynamics, something that would otherwise have taken us many years to perceive. We learned, first of all, that both the schools and our project needed more structure; we established regular meetings so that the staff could coordinate plans and set goals. More important, our analysis of interactions among parents, staff and students revealed a basic problem underlying the schools' dismal academic and disciplinary record: the sociocultural misalignment between home and school. We developed a way to understand how such misalignments disrupt beneficial relations and how to overcome them in order to promote educational development.

Our understanding is based on the fact that a child develops a strong emotional bond to competent caretakers (usually parents) that enables them to help the child develop. Many kinds of development, in social, psychological, emotional, moral, linguistic and cognitive areas, are critical to future academic learning. The attitudes, values and behavior

of the family and its social network strongly affect such development.

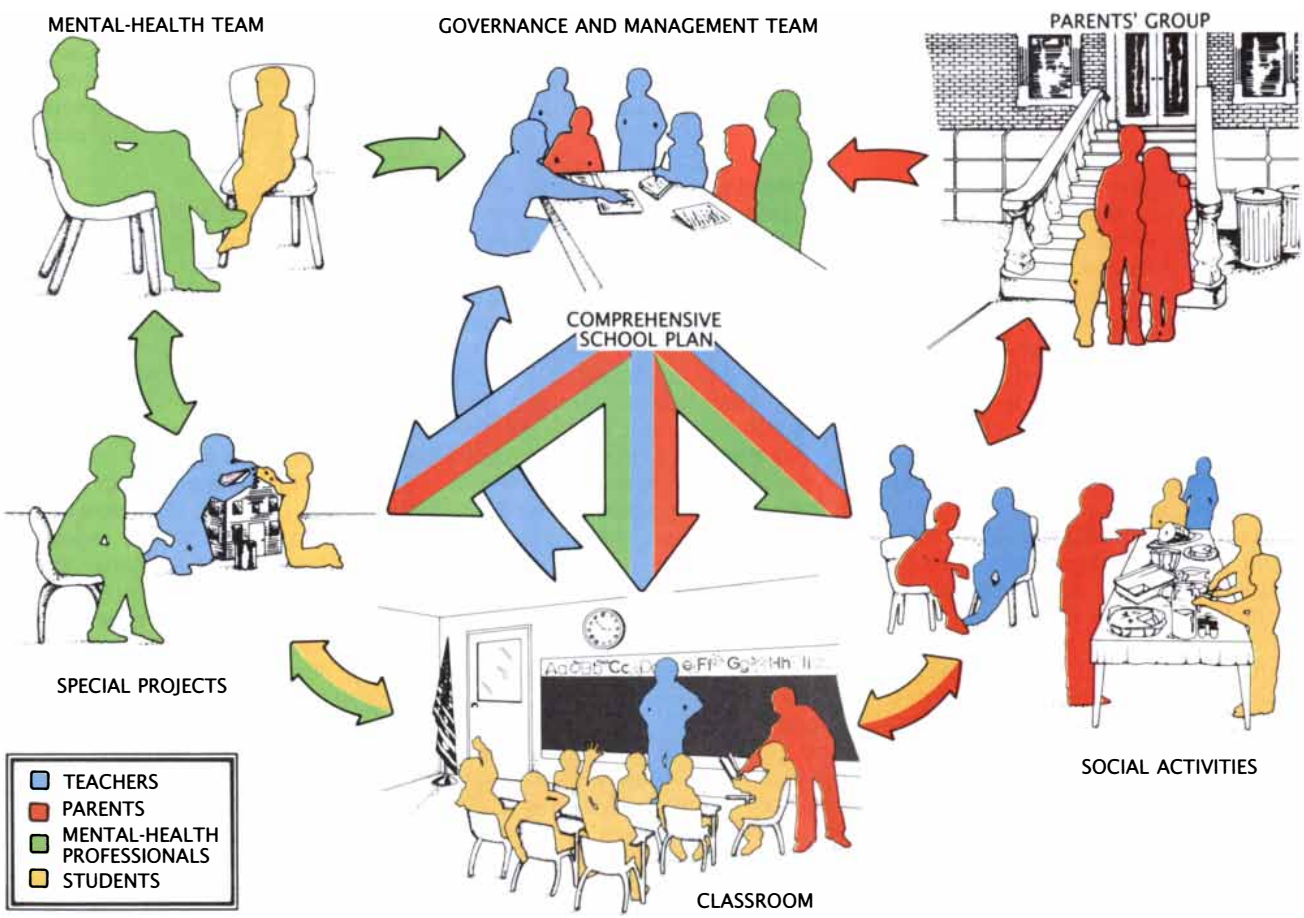
A child whose development meshes with the mainstream values encountered at school will be prepared to achieve at the level of his or her ability. In addition the meshing of home and school fosters further development: when a child's social skills are considered appropriate by the teacher, they elicit positive reactions. A bond develops between the child and the teacher, who can now join in supporting the overall development of the child.

A child from a poor, marginal family, in contrast, is likely to enter school without adequate preparation. The child may arrive without ever having learned such social skills as negotiation and compromise. A child who is expected to read at school may come from a home where no one reads and may never have heard a parent read bedtime stories. The child's language skills may be underdeveloped or nonstandard. Expectations at home and at school may be radically at odds. For example, in some families a child who does not fight back will be punished. And yet the same behavior will get the child into trouble at school.

Such lack of development or development that is at odds with the mainstream occurs disproportionately often among children from the minority groups that have had the most traumatic experiences in this society: Native Americans, Hispanics and blacks. The religious, political, economic and social institutions that had organized and stabilized their communities have suffered severe discontinuity and destruction. Furthermore, these groups have been excluded from educational, economic and political opportunity. These themes are particularly vivid in the black experience.

Blacks arrived in this country forcibly uprooted from their own culture, and they had another culture—that of slavery—imposed on them. Slavery was a state of enforced dependency and inferiority, which offered no future. The dominant Anglo culture, in contrast, placed a high value on independence and personal advancement. The dominant culture devalued the imposed black culture, and many blacks in turn developed a negative self-image. After the abolition of slavery, widespread discrimination denied blacks access to education and to the political and economic mainstream. Yet in spite of these psychological and social handicaps many poor black families, particularly in rural areas, were able to develop strong religious and cultural support systems and to function reasonably well.

After World War II opportunities for rural work diminished and many black families migrated to cities, but as a result of discrimination they were largely shut out of the primary job markets. Moreover, urban jobs de-



SCHEMATIC of the school-intervention program shows its key components and the relations among them. A governance and management team, consisting of the principal, parents, teachers and a mental-health worker, develops a comprehensive school plan covering academics, social activities and special programs, such as a Discovery Room for children who have lost interest in learning. Social activities, such as potluck suppers, teach children social skills and enable parents to meet teachers. Some parents become teachers' aides. The mental-health team assigns a member to work with a child who is having difficulty. It also tries to prevent behavior problems by recommending changes in school procedures. By reducing behavior problems and improving relations with parents, the program creates a school climate conducive to learning.

MENTAL-HEALTH TEAM of the Katharine Brennan School meets weekly under the direction of principal Dietria Wells (*center*). Members track the progress of students who are having problems and discuss intervention strategies, which often involve the children's families. Participants include staff member Bridget Hardy (*left*), speech pathologist Judith Campbell (*right*) and psychologist Karen McCleu (*far right*).

manded a higher level of education than rural ones, and blacks, undereducated in prewar years, were at a disadvantage. At the same time, they experienced severe stress resulting from the loss of supportive communities. For all these reasons, many black families began to function less well and could not provide their children with preschool experiences that would enable them to succeed in school.

Furthermore, blacks were able to achieve mainstream success only in limited professional areas. Thus they could not gain a significant share of political, economic and social power in the larger society and thereby help to advance socially marginal blacks. With time, marginal blacks came to resent mainstream blacks and whites for being unable—and apparently unwilling—to help them, and they defensively rejected the mainstream.

In spite of their alienation from the mainstream, many poor black parents still look to the school as their hope—indeed, their only hope—for the future, even though at the same time they expect the school to fail them and their children as other mainstream institutions have. And in fact the schools often do fail them. Typical schools, with their hierarchical and authoritarian structure, cannot give underdeveloped or differently developed students the skills and experiences that will enable them to fulfill expectations at the school. Instead such students are labeled "bad," unmotivated or stupid. Staff people punish the children and hold low expectations for them, often blaming the students, their parents and their communities for the problems. Parents, for their part, take the problems as a personal failure or as evidence of animosity and rejection by the mainstream. They lose hope and confidence and become less supportive of the school. Some parents, ashamed of their speech, dress or failure to hold jobs, become defensive and hostile, avoiding contact with the school staff.

The result is a high degree of mutual distrust between home and school. A black first-grade teacher in an inner-city school with a nearly all-black student body recalled explaining classroom rules on the first day. When she finished, a six-year-old raised his hand and said, "Teacher, my mama said I don't have to do anything you say." Fortunately this teacher understood the underlying problem, but most teachers would have reacted angrily, whereupon any chance of gaining parental cooperation would have quickly evaporated. This degree of alienation between home and school makes it difficult to nurture a bond between child and teacher that can support development and learning.

The consequences of alienation become most apparent when these children reach the age of about eight. Around this age they are expected to progress academically at a rate that begins to exceed their level of development. In addition the children begin to understand how they and their families differ in income, education and sometimes race and style from other people in the school. At this age, moreover, children seek to decrease their dependence on adults and on the approval of adults.

Unable to achieve in school, these children begin to see academic success as unattainable, and so they protect themselves by deciding school is unimportant. Many seek a sense of adequacy, belonging and self-affirmation in nonmainstream groups that do not value academic achievement. Such children are at risk for dropping out, teen-age pregnancy, drug abuse and crime. On the other hand, the decision to pursue academic achievement and to join the mainstream also exacts a heavy price: such a choice means rejecting the culture of one's parents and social group.

Our analysis of the two New Haven schools suggested that the key to academic achievement is to promote psychological development in students, which encourages bonding to the school. Doing so requires fostering positive interaction between parents and school staff, a task for which most staff people are not trained. Such changes cannot be mandated or sustained from outside the school. Our task, then, was to create a strategy that would overcome the staff's resistance to change, instill in them a working understanding of child development and enable them to improve relations with parents.

From our experience during the first difficult year it was obvious that we would make no progress until we had reduced the destructive interactions among parents, teachers and administrators and given cohesiveness and direction to the schools' management and teaching. To this end we created in each school a governance and management team of about a dozen people led by the principal and made up of elected parents and teachers, a mental-health specialist and a member of the nonprofessional support staff—all the adults who had a stake in the outcome. The teams decided issues ranging from the schools' aca-

demic and social program to changes in school procedures that seemed to engender behavior problems.

Several rules guided these teams. First, team members had to recognize the authority of the principal but, equally important, the principal could not push through decisions without weighing the concerns of the team members. Second, we agreed to focus efforts on problem solving and not waste time and energy in placing blame. Third, we made decisions by consensus rather than by vote; this promoted cooperation by reducing the harmful tendency of groups to polarize into "winners" and "losers."

The teams were not fully accepted at first, nor were they immediately effective, because we at the center were viewed as outsiders (from Yale, to boot, which working-class people in New Haven have always regarded with suspicion). But as we helped the principals to see that power sharing increased their own ability to manage the school, and as teachers and administrators benefited, the staff became more willing to apply our expertise in social and behavioral sciences to every aspect of the school.

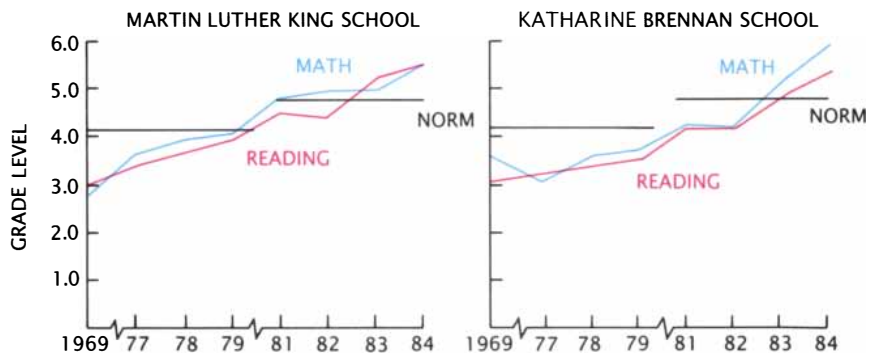
We invited parents from among the group that had protested against us in the first year to join the team. With their input we developed a program that involved parents at three levels: shaping policy through their representatives on the governance and management team, participating in activities supporting the school program, and attending school events.

At one point about a dozen parents worked as classroom assistants and formed the core of the parents' group. (They were paid the minimum wage.) Parents and staff sponsored activities such as potluck suppers, book fairs and graduation ceremonies. These social gatherings fostered good relations between parents and staff, so that when a child was having problems, the staff could discuss the matter with the parents without eliciting defensive reactions. As a result the school climate and student behavior improved, and more parents began to attend school activities.

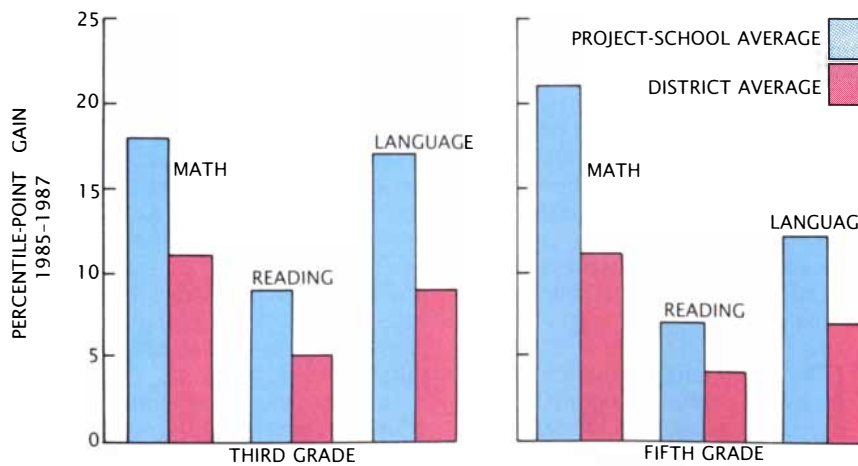
In a typical school, students who have emotional, learning or behavioral difficulties are seen by the school's psychologist, social worker or special-education teacher, who all work independently of one another. In our project, however, they worked as a team. We found this to be more efficient; the team would discuss each case and assign one member to it. The team approach also made it easier to detect patterns of troublesome behavior and to determine whether some aspect of the school was making them worse. Through its delegate on the governance and management team, the mental-health group recommended changes in school policies and practices so that students' developmental needs would be served better and behavior problems prevented.

These actions reduced the sense of failure, the feelings of anger and the loss of confidence that can lead to problem behavior among students. For example, an eight-year-old who was transferred into King from another school was taken directly to the classroom. He panicked, kicked the teacher in the leg and ran out. Usually such a child is punished. If nothing is done to reduce the child's anxiety, the cycle is often repeated until the child is labeled disturbed and referred for treatment. Our mental-health team helped the school staff to understand that the child's anxiety was a natural reaction to being thrust among strangers, and, together with the staff, we developed an orientation program to introduce transfer students and their parents to the school.

In the course of the 12 years we spent in the New Haven schools, other programs emerged in response to students' needs. In one school, children were kept with the same teacher for two years. A Discovery Room enabled "turned off" children to form a trusting relationship with an adult and, through play, rediscover an interest in learning. A Crisis Room provided a



FOURTH GRADERS at the two New Haven schools taking part in Yale University's Child Study Center's intervention program registered steady gains in achievement-test scores from 1969 through 1984. The graphs show mean scores on the Iowa Test of Basic Skills in reading (*red*) and mathematics (*blue*); scores in 1969 are for the Metropolitan Achievement Test. From 1969 through 1979 the tests were given in the fall (when the norm is a score of 4.2); from 1981 through 1984 the tests were given in the spring (norm score 4.8). Scores have stayed near the 1984 levels since then.



AVERAGE PERCENTILE GAINS on California Achievement Test scores from 1985 through 1987 were larger for 10 mainly black schools in Prince Georges County, Md., that use the Comer program than they were for the school district as a whole. Test scores of black students still lag behind those of white students, but the gap is narrowing. The school district, the 15th-largest in the U.S., has 105,000 students, 62 percent of whom are black. The schools using the program are more than 90 percent black; they receive extra staff and funds because they have been hard to integrate.

refuge for children who were "out of control." We discovered that this behavior could often be traced to a traumatic home experience, and so staff members helped the children to handle their feelings and regain a sense of control. With each intervention the staff became increasingly sensitive to the concerns of developing children and to the fact that behavior problems result mainly from unmet needs rather than from willful badness—and that actions can be taken to meet these needs.

By 1975 the program was clearly having an effect. Behavioral problems had declined, relations between parents and staff had improved and the intelligence of the children had become manifest. In that year we drew up a formal School Development Program based on the key ingredients of our success: the governance team, the parents' program and the mental-health team.

Having established a way to achieve and maintain a smoothly functioning school, we decided to see whether the school might also play a role in redressing the problem of social misalignment. We argued that it should be possible to teach our nonmainstream students the social skills that are expected of them in school, and that the acquisition of these skills would help them to succeed academically.

Staff and parents devised a curriculum of social skills, with instruction in the subjects children would need to know: politics and government, business and economics, health and nutrition, and spiritual and leisure activity. The staff chose specialists to help develop the program. Children learned how to write invitations and thank-you notes, how to serve as hosts, how the body functions, how to write checks, how to plan concerts, and so on. Each activity combined basic academic skills with social skills and an appreciation of the arts. These activities were an immediate and dramatic success. Students, parents and staff alike all felt a surge of excitement and a growing sense of participating in the mainstream.

The intervention program in New Haven produced significant academic gains. The students had once ranked lowest in achievement among the 33 elementary schools in the city, but by 1979, without any change in the socioeconomic makeup of the schools, students in the fourth grade had caught up to their grade level. By 1984 pupils in the fourth grade in the two schools ranked third- and fourth-highest on the Iowa Test of Basic Skills. By the early 1980's attendance rates at King were either first or second in the city. There have been no serious behavior problems at either school in more than a decade.

In 1980 our group left the schools. The program was fully integrated into the normal practices of the staff, who continued to carry it out. In the same year we began to develop a way to apply our program in other schools. We left intact the key elements of our success in New Haven—the governance and management team, the parents' program and the mental-health team, along with our operating rules—while allowing specific social and academic activities to vary with the needs of a particular school. In a sense the program enables school personnel to engage in a "clinical practice": armed with theories of child development and education, together with observations of children and school systems, they can diagnose problems in the school and develop solutions.

The Prince Georges County, Md., and Benton Harbor, Mich., school districts, which serve mainly low-income black children, have been using the program for several years; they have achieved successes on a par with those of the two New Haven schools. The program is being introduced to all New Haven schools, as well as to three other districts: Norfolk, Va., Lee County, Ark. (both also serving mainly poor black children), and Leavenworth, Kans. The program is now being implemented in more than 50 schools around the country, including two middle schools and one high school.

All the money and effort expended for educational reform will have only limited benefits—particularly for poor minority children—as long as the underlying developmental and social issues remain unaddressed. Yet most teachers and administrators are not trained to organize and manage schools in ways that support the overall development of students. Nor does their training enable them to analyze, much less solve, the social-misalignment problems of children from outside the mainstream.

The first step toward improving the education of these children, then, is to induce teachers' colleges and schools of education to focus on student development. Teachers who invest time in training will have an incentive to use what they have learned. The efforts of individuals will not be enough; the entire staff of a school must embrace new ways of thinking.

School districts, state and local governments and school boards must actively support these changes. They must recognize that students' social development is as important to society as their academic ability. They must select, certify and reward teachers and administrators who are skilled in nurturing the development of students. They must evaluate schools by their ability not only to produce high test scores but also to prepare students to assume adult responsibilities. And they must provide necessary funds.

The Federal Government must play a leading role in bringing about national educational reforms that can prepare young people to be effective and responsible citizens. Besides appropriating funds and establishing programs, the Government must facilitate the interaction of state and local government, educational authorities and private interests—foundations, businesses, colleges and universities.

To pull all of this together, I believe a National Academy of Education is needed. Its purpose would be to set national priorities, assess current research in education, learn how to implement approaches that work, identify areas for further study and allocate resources effectively. Such an academy must be free from the pressures of political expediency, and the interests of researchers must be balanced against those of educators. It could be created largely from existing Federal programs and governed by those who have a stake in education: educators, parents, government and business. A National Academy of Education could spearhead a process of change that is geared above all to the needs of children and to the national interest. It could set a timetable and move forward at a rational pace that recognizes the urgent need for reform and at the same time is guided by knowledge and common sense.

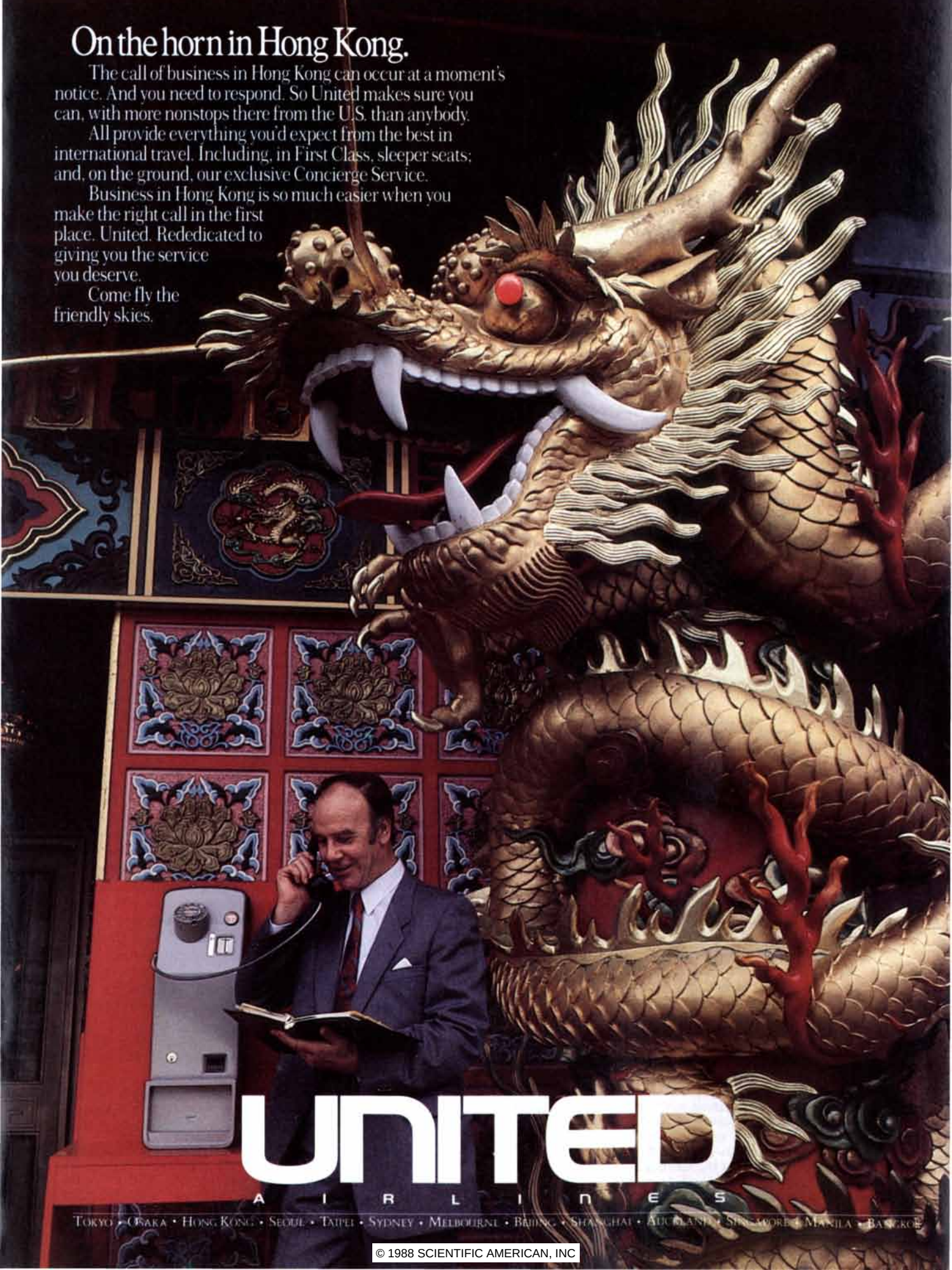
FURTHER READING

SCHOOL POWER: IMPLICATIONS OF AN INTERVENTION PROJECT. James P. Comer. The Free Press, 1980.

FAMILY LIFE AND SCHOOL ACHIEVEMENT: WHY POOR BLACK CHILDREN SUCCEED OR FAIL. Reginald M. Clark. University of Chicago Press, 1983.

GOOD SCHOOLS: WHAT RESEARCH SAYS ABOUT IMPROVING STUDENT ACHIEVEMENT. Willis D. Hawley and Susan J. Rosenholtz in *Peabody Journal of Education,* Vol. 61, No. 4; Summer, 1984.

WHY THE CURRENT WAVE OF SCHOOL REFORM WILL FAIL. John E. Chubb in *The Public Interest,* Vol. 90, pages 28–49; Winter, 1988.



Quasi-periodic Oscillations in Celestial X-Ray Sources

Not quite periodic, not quite random fluctuations in X-ray intensity provide clues to the nature of extremely bright X-ray sources loosely clustered near the center of our galaxy

by Michiel van der Klis

Much of what we know about X-ray sources in our galaxy derives from regularities in the way the intensity of their radiation varies with time. Cyclic increases and decreases in intensity can indicate that a star is orbiting the X-ray source, eclipsing it periodically. X-ray pulses uniformly spaced less than a few seconds apart, on the other hand, imply that the source is a high-density object, known as a neutron star, that happens to be strongly magnetized and spinning rapidly as well. Intermittent bursts of X rays also indicate the presence of a neutron star, but one on whose surface enormous thermonuclear explosions are taking place.

Analyzing the finer details of these kinds of regularities enables astronomers to learn about the physics of galactic X-ray sources. Yet even after a quarter century of observations the brightest X-ray sources in the nighttime sky, most of which are found in the central bulge of our galaxy, did not appear to exhibit any regular phenomena. Hence while the secrets of other, fainter X-ray sources were slowly being unraveled, the bright galactic-bulge sources remained mysterious.

It is for this reason my colleagues and I were both surprised and excited when we began to realize, in late 1984 and early 1985, that our observations of one of the brightest galactic-bulge sources, GX5-1, had revealed a peculiar type of regular variation in its X-ray intensity. (GX5-1 is named for its position in the sky: five degrees east and one degree south of the galactic center.) The observations were made with a satellite of the European Space Agency, the X-ray observatory *EXOSAT,* as part of a search program planned and carried out by a group consisting of Fred Jansen of the Laboratory for Space Research in Leiden, Jan van Paradijs and Edward P. J. van den Heuvel of the University of Amsterdam, Walter H. G. Lewin of the Massachusetts Institute of Technology, Joachim Trümper and Mirek Sztajno of the Max Planck Institute for Extraterrestrial Physics in Munich and me.

The variations we observed in the intensity of GX5-1's X-ray emissions were particularly surprising because they were not strictly periodic: in any given observation the interval between X-ray-intensity peaks was likely to be close to .03 second, but often it was a bit longer or shorter. Such not quite periodic fluctuations are generally called quasi-periodic oscillations (QPO's). If an analogy is drawn between the oscillations of GX5-1 and common sound waves, our surprise can be described in more familiar terms: Instead of "sounding" like a pure tone of a given frequency, the variations in fact "sounded" more like breaking glass. Our excitement was heightened by the fact that the oscillations were very rapid. That meant that their origin was probably quite close to a compact celestial object, where intense gravitational forces make things happen very fast.

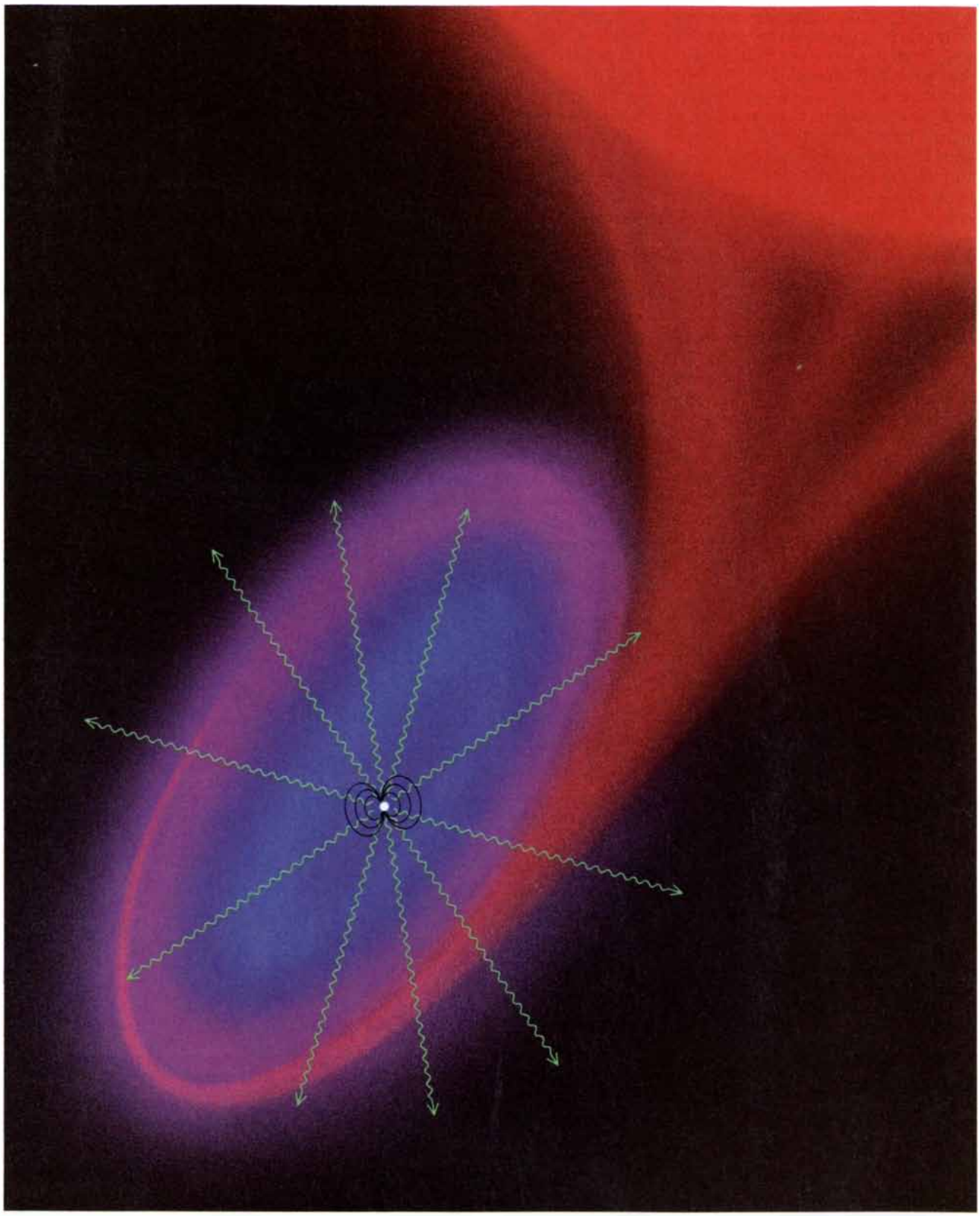
Why were we searching for rapid oscillations in bright galactic-bulge X-ray sources in the first place? To answer that question, I need to review some of what the 25 years of X-ray astronomy that predated our observations had revealed.

With only a few exceptions, bright X-ray sources in our galaxy are thought to be binary, or double-star, systems consisting of a gravitationally collapsed star (either a neutron star or a black hole) in orbit around a more or less normal star. Such collapsed stars represent the final stage of evolution for stars several times as massive as the sun. That stage is reached when not enough heat can be generated in the star to counterbalance the inward force of gravity.

As its name implies, a neutron star consists largely of closely packed neutrons. It typically has a mass between one and two times that of the sun compressed into a sphere only 20 kilometers in diameter. (In contrast, the sun's mass occupies a sphere more than a million kilometers across.) A black hole is the extremest form of gravitational collapse; one with a mass five times that of the sun would have an effective diameter of only about 30 kilometers. The combination of small size and large mass means that an enormously strong gravitational field surrounds either a neutron star or a black hole. It is this intense gravitational field that is responsible for the strong X-ray emissions of these binary systems.

In most bright X-ray binaries the collapsed star revolves so closely around its normal companion that its gravity can peel off plasma, or ionized matter, directly from the outer layers of the companion. The plasma swirls around the collapsed star, forming

MICHIEL VAN DER KLIS is on the staff of the *EXOSAT* Observatory at the European Space Agency's European Space Research and Technology Centre (ESTEC) in Noordwijk, The Netherlands. He studied at the University of Amsterdam, earning his Ph.D. in astronomy in 1983. After a postdoctoral position at the university, he received a research fellowship at ESTEC. He took his current position there in 1986. For the discovery of quasi-periodic oscillations in celestial X-ray sources, van der Klis was awarded the 1987 Bruno Rossi Prize of the American Astronomical Society.



BINARY SYSTEMS consisting of a neutron star (*white*) and a more or less normal companion star (*red*) are thought to account for most X-ray sources in the nighttime sky. A neutron star is formed when a star several times more massive than the sun collapses under its own weight. The enormously strong gravitational field of the collapsed star draws plasma, or ionized matter, from its companion (which can be as much as several million times larger) into a so-called accretion disk. From the accretion disk the plasma spirals down to the surface of the neutron star, reaching a speed approaching that of light as it hits the star's surface. On impact as much as 20 percent of the plasma's mass can be converted into energy, mostly in the form of X rays (*green*). During its formation a neutron star can be endowed with a very strong magnetic field (*black lines*) that may take about 100 million years to decay. Hence the strength of a neutron star's magnetic field is an indicator of its age.

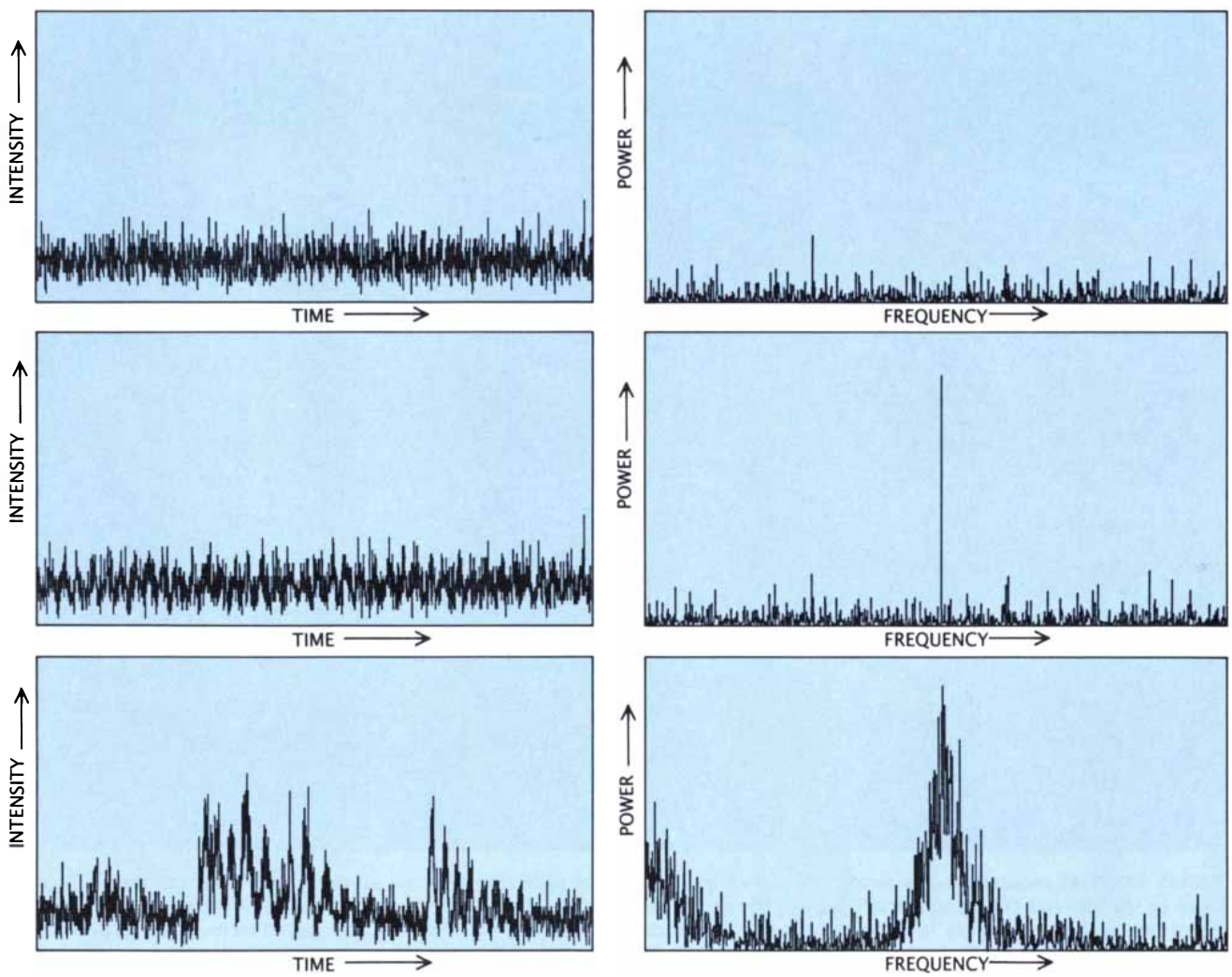
a rotating disk, called an accretion disk, that is centered on the collapsed star. From the accretion disk the plasma ultimately spirals inward onto the star's surface. In spiraling to the surface the matter is accelerated to speeds approaching the speed of light and is heated to temperatures close to 100 million degrees Kelvin. The result is the release of huge amounts of energy in the form of X rays.

Accretion onto a compact object is actually a highly efficient way of producing radiation. If the object is a neutron star, as much as 20 percent of the mass of the plasma can be converted into energy (in accordance with Einstein's famous relation, $E = mc^2$). If the object is a black hole, the efficiency can be as high as 40 percent. To put such efficiencies into perspective, consider that the most energetic process on the earth, a thermonuclear explosion, has a mass-to-energy conversion efficiency of less than 1 percent. A bright X-ray source such as GX5-1 accretes about a trillion tons of plasma per second and emits as much energy in the form of X rays in one second as the sun does in all wavelength bands in several days.

It is important to distinguish between "young" X-ray binary systems (which have existed for less than 10 million years) and "old" binary systems (which have existed for more than a few hundred million years), since the infalling plasma interacts differently with the neutron star in the two cases. Young X-ray binary systems are found close to the galactic spiral arms, where stars are born. The noncompact stars in these systems are giant, blue-white stars that have a mass more than 10 times that of the sun. For this reason the systems are also known as massive X-ray binaries.

Old X-ray binary systems, among which are all QPO sources, can be found almost everywhere in our galaxy, not just near regions of star formation. Because massive stars have a relatively short lifetime, the noncompact stars in the old systems must have low mass (that is, they must be less massive than the sun). This has actually been verified for a few old X-ray binary systems that can be observed with optical telescopes. Old systems are therefore also known as low-mass X-ray binaries.

During the gravitational collapse that results in the birth of a neutron star, dynamo effects can generate very strong magnetic fields (about 10^{12} gauss, more than a trillion times that of the earth). Such a magnetic field will disrupt the motion of the plasma in the accretion disk thousands of kilometers above the surface of the neutron star. Indeed, within a region known as the magnetosphere the plasma is constrained to move along the lines of the magnet-



FOURIER ANALYSIS enables astronomers to break down the variability in intensity of a celestial object's X-ray signal (*left*) into its many frequency components in a so-called power spectrum (*right*). In this way X-ray astronomers can readily distinguish between background X-ray "noise" (*top*), periodic oscillations (*middle*) and quasi-periodic oscillations, or QPO's (*bottom*).

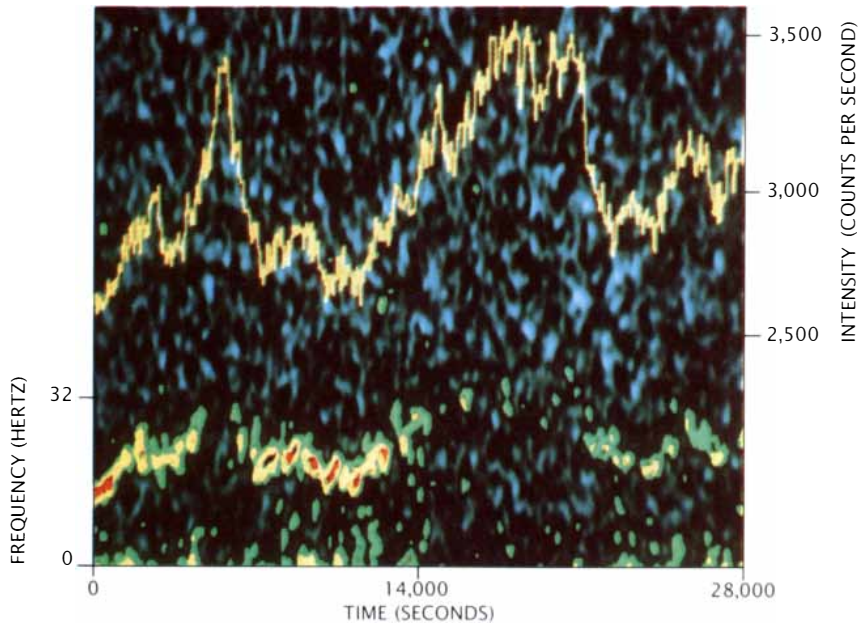
ic field, forcing the plasma to fall onto the neutron star near its magnetic poles. As a result two "hot spots" form on the surface of the star (one for each magnetic pole), and it is from these that most of the X rays are emitted. If these hot spots do not lie on the rotation axis of the star, then two wide beams of X rays emanating from the hot spots are swept around with each rotation of the star. To an observer on the earth such a neutron star, known as an X-ray pulsar, would manifest itself as regular pulses of X rays, much as a rotating lighthouse beacon is seen as a pulsating light. Hence whenever such X-ray pulsations are detected, they are taken to indicate the presence of a rotating, strongly magnetic neutron star in a binary system.

Apart from generating an enormous quantity of X rays, the accretion of matter onto a neutron star has another important consequence. As the plasma hits the surface of the star it transfers its motion to the star, causing the star to spin ever faster in the same direction as the circulating plasma. The star stops gaining angular momentum only when it is spinning as fast as the plasma is orbiting. It is then said to be spinning at its equilibrium period. If the star were to spin any faster than its equilibrium period, it would actually begin to hurl matter into space and thereby slow down its rate of spin.

If the magnetic field of the neutron star is strong enough for a magnetosphere to form, the orbital velocity of the plasma just outside the magnetosphere determines the equilibrium period. The reason is that inside the magnetosphere the plasma is no longer in a free orbit but is attached to the field lines, which in turn are rigidly connected to the star. For a magnetospheric radius of about several thousand kilometers (typical of an X-ray pulsar in a young binary system) the equilibrium period is on the order of several seconds. Indeed, many X-ray pulsars with such periods have been observed.

Most astrophysicists think that a neutron star's strong magnetic field does not persist forever but gradually decays. For this reason a neutron star in an old binary system is expected to have a weak magnetic field, if any, and at most a very small magnetosphere. (In this respect X-ray binaries containing an old neutron star are similar to those containing black holes: it is impossible for a black hole to have a strong magnetic field.)

Because there is no prominent magnetosphere enveloping a neutron star in an old binary system, the inner part of its accretion disk extends nearly all the way down to the surface of the star. In such a case plasma orbiting just above the stellar surface determines the equilibrium period. Because the orbital velocity of the circulating plasma increases as its orbital radius decreases, the equilibrium periods in the old systems are therefore predicted to be extremely short—on the order of one millisecond.



DYNAMIC POWER SPECTRUM of GX5-1 shows how the intensity of its X-radiation (*top trace*) correlates with the frequencies of its QPO's (*bottom trace*). When the X-ray intensity is low, distinct QPO's with frequencies near 20 hertz (cycles per second) are evident. (The strength of the QPO's is indicated by the color of the bottom trace.) When the X-ray intensity rises, the QPO frequency also rises until the intensity peaks, at which point the QPO's seem to disappear. The reason for their apparent disappearance is that the power of the QPO's has become so thinly spread among a range of frequencies that it is lost against the background X-ray noise.

Hence, according to the generally accepted picture of X-ray binaries I have painted, one expects to find very rapidly spinning neutron stars in old X-ray binaries. My colleagues and I were searching for just such millisecond rotation periods in X-ray sources when we made our observations of GX5-1. As it turns out, we did not observe the predicted rapid spins directly; in fact, they still have not been detected.

This is not altogether surprising, however, since the weak magnetic field that makes the rapid rotation possible also keeps astronomers from detecting the rotation. Without a strong magnetic field there is no channeling of the plasma to the magnetic poles; there are no rotating hot spots and consequently no distinct X-ray pulsations. And so although the pulsations from old systems are predicted to be rapid, they are also predicted to be weak and indistinct.

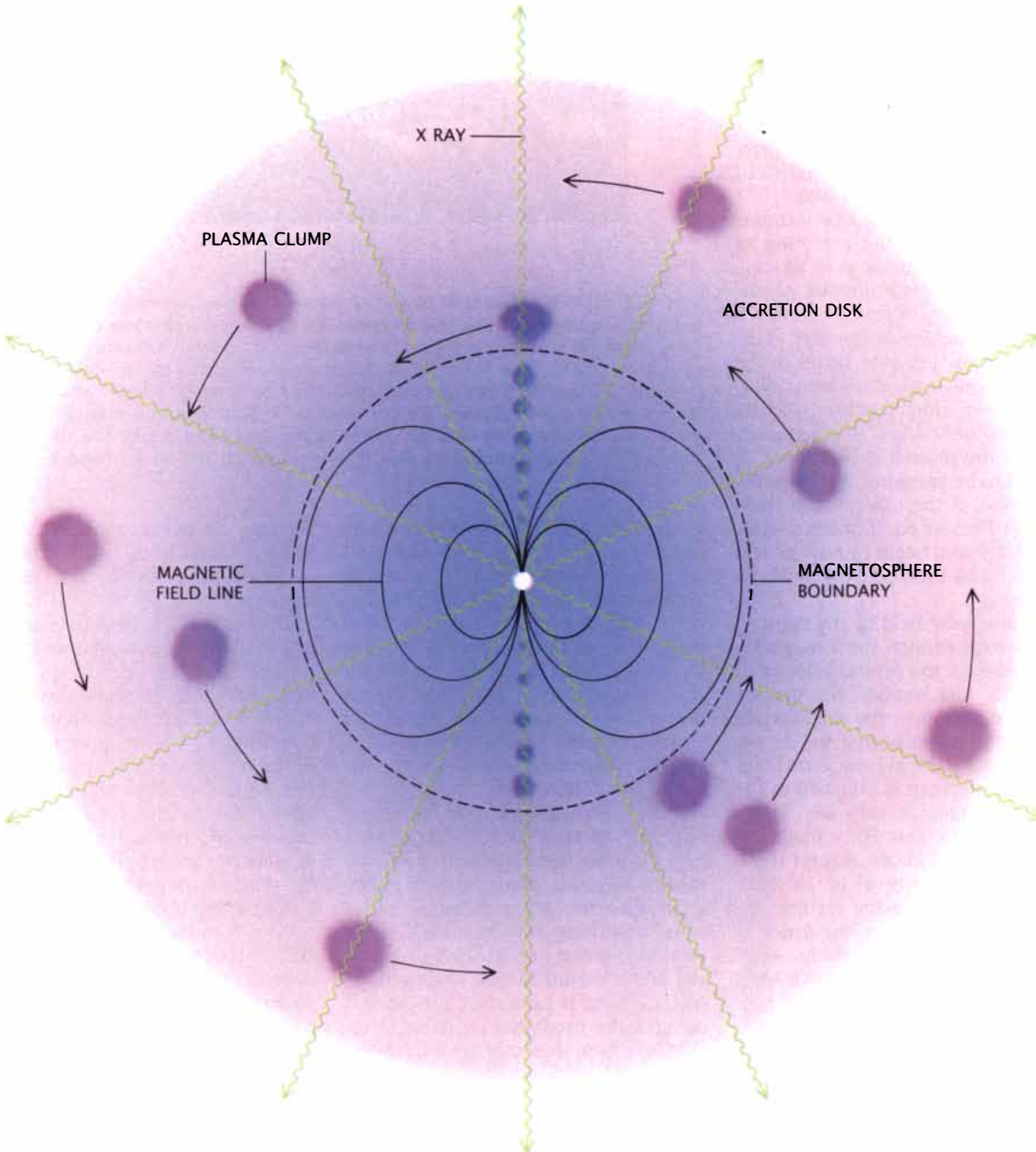
The best way to look for such rapid but weak pulsations against the background X-ray "noise" is to apply an old and proved technique called Fourier analysis. The technique enables one to construct a so-called power spectrum of the variations in X-ray intensity [*see illustration on opposite page*]. A power spectrum provides, for each of a large number of frequencies, an estimate of the strength of all the signals that fluctuate within a given small range (called a bin) about that frequency.

If the X-radiation emanating from a certain part of the sky varies periodically in intensity, the variation will stand out in a power spectrum as a narrow spike in the bin that contains the variation's fixed frequency. The longer the observation lasts, the more clearly the spike stands out above the noise in the surrounding bins. Whereas the noise is spread out among all the bins, the entire power of the periodic signal is concentrated in one bin. For this reason it is generally advantageous to make the frequency bins of the power spectrum as narrow as possible, because squeezing the power of a periodic signal into a narrow bin results in a higher spike. This makes it

possible for even a very weak but regular pulsation to be detected against the background-noise level.

Yet if an X-ray signal is not periodic but quasi-periodic, its power spectrum will have not a narrow spike but a broad peak, since its power is divided among many frequency bins. In this case narrow frequency bins are actually counterproductive: the more frequency bins the power spectrum has, the more thinly a quasi-periodic signal's power is spread among the bins and the more likely it is to get lost in the background noise. Indeed, one of the reasons it took X-ray astronomers 25 years to find QPO's may be that most of them were expecting to find periodic signals and therefore searched for spikes in narrow power-spectrum frequency bins.

Fortunately my colleagues and I had by chance arranged things so that the power-spectrum bins were relatively wide, and the quasi-periodic variability of GX5-1 was conspicuous to us right from the first observation. Having found evidence for a broad peak in the power spectrum, we then used a "dynamic" power spectrum, which enabled us to follow the evolution of the peak as a function of time. Such a spectrum showed that the peak of



CLUMPS OF PLASMA in the accretion disk surrounding an old neutron star (one having a relatively weak magnetic field) are the source of "fast" QPO's, according to the magnetospheric beat-frequency model. Every time a clump passes an "entrance" to the magnetosphere (the region in which plasma is constrained to move along the magnetic field lines anchored in the star), part of it manages to fall to the star's surface, emitting a flash of X rays when it impinges on the surface. In this way a train of gradually fading X-ray pulses is generated whose frequency is equal to the difference between the orbital frequency of the plasma clumps and the rotational frequency of the magnetosphere—their beat frequency. The resulting X-ray signal has a broad QPO peak in its power spectrum. Many plasma clumps could contribute simultaneously to the total X-ray signal emanating from the vicinity of the neutron star, but the signal would have a similar power spectrum.

The difference between a child behind the wheel of a play car and a drunk behind the wheel of a real car is that only one of them is funny.

Keeping drunk drivers off the roads is serious business. To that end, we've developed an enzyme called Alcohol Oxidase that makes sobriety testing more accurate than ever. And it makes it faster, because it can be administered and analyzed right there on the spot.

Alcohol Oxidase manufacturing is only one of the breakthroughs in biotechnology we've patented. Through our Provesta Corporation subsidiary, we're making high-protein food supplements for creatures as small as a fish and as big as a horse.

For the human animal, biotechnology has led us to research in medicines that could break up blood clots, stop heart attacks in progress, and reproduce the body's natural cancer-fighting agents. And we eliminated the headaches of pain reliever manufacturers with research and development advances that make their new formulations possible.

Whether it's keeping the roads safe for our children or paving new roads to health, we're in the driver's seat and headed for the future.

For more information write to Greg Derrick, Phillips Petroleum, 16D-4 Phillips Bldg., Bartlesville, OK 74004.

Phillips Petroleum Company
Performing to meet
the challenge of change.

PHILLIPS 66

IT'S SOBERING TO THINK THAT KIDS TAKE DRIVING MORE SERIOUSLY THAN SOME ADULTS.

GX5-1's X-ray signal was shifting over a large range in frequency, between 20 and 40 hertz (cycles per second).

Although the high average frequency of the QPO's implies that their origin must be near the surface of a collapsed star, we knew immediately that we were not seeing a neutron star's spin directly: the star's rotation rate would have to be changing by a factor of two within a few hours—an impossibility. But if the oscillations were not directly generated by the rotation of a neutron star, what could be causing them?

A clue came from a further examination of GX5-1's dynamic power spectrum: it appeared that whenever the oscillations increased in frequency, the X rays became brighter. The converse also applied: whenever the oscillations slowed, the X-ray intensity dropped. Such behavior suggested that a magnetosphere might be responsible for the QPO's.

The size of a neutron star's magnetosphere is determined by two opposing forces. Whereas the plasma spiraling inward from the accretion disk tends to compress the magnetosphere, magnetic forces tend to expand it. As a result an increase in the rate at which matter accretes onto the neutron star has two effects. First, more gravitational energy is released, so that the X rays emanating from the star become brighter. Second, the material forces gain ground on the magnetic ones, so that the magnetosphere shrinks and allows the freely orbiting plasma just outside it to move in smaller, faster orbits. Hence if the oscillations are caused by the orbital motion of plasma just outside a neutron star's magnetosphere, one would expect their frequency to increase as the X rays become brighter, which is what we observed.

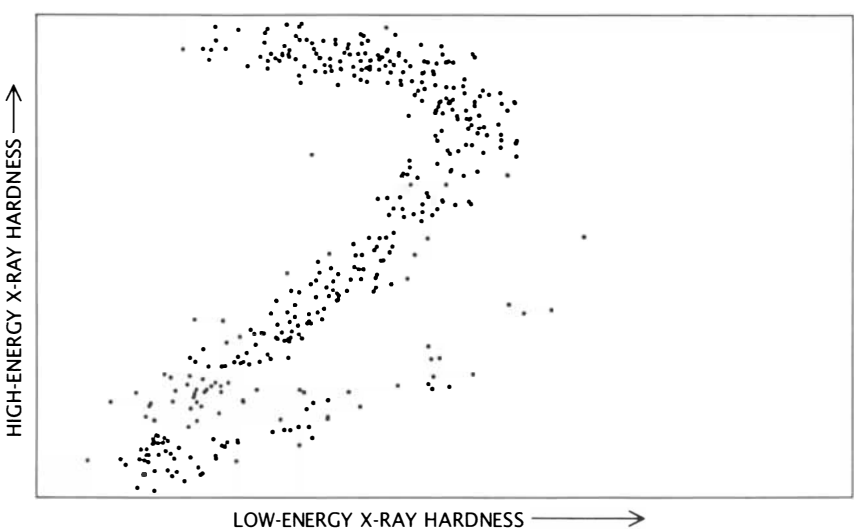
On closer scrutiny, however, this hypothesis seemed untenable: the precise relation between X-ray intensity and QPO frequency could not be reproduced in calculations. But then M. Ali Alpar of the University of Illinois at Urbana-Champaign and Jacob Shaham of Columbia University put forward a marvelous idea: the observed frequency of the oscillations did not have to be equal to the orbital frequency of the matter just outside the magnetosphere; it could be equal to the difference between this frequency and the (unknown) spin frequency of the neutron star.

This concept can be understood by considering the hands of a clock. The frequency of the large hand is one cycle per hour and that of the small hand is one-twelfth of a cycle per hour. The frequency with which the large hand passes the small hand—their "beat" frequency—is given by the difference between the two frequencies: 11/12 times per hour, or about once every hour and five minutes. In the Alpar-Shaham model the QPO frequency is given by the beat frequency between the matter circulating in the accretion disk and the magnetosphere, which corotates with the neutron star. In essence the beat frequency is the frequency with which a given plasma particle passes a given magnetic field line in the spinning magnetosphere. The Alpar-Shaham model could account nicely for the observed relation between QPO frequency and X-ray intensity.

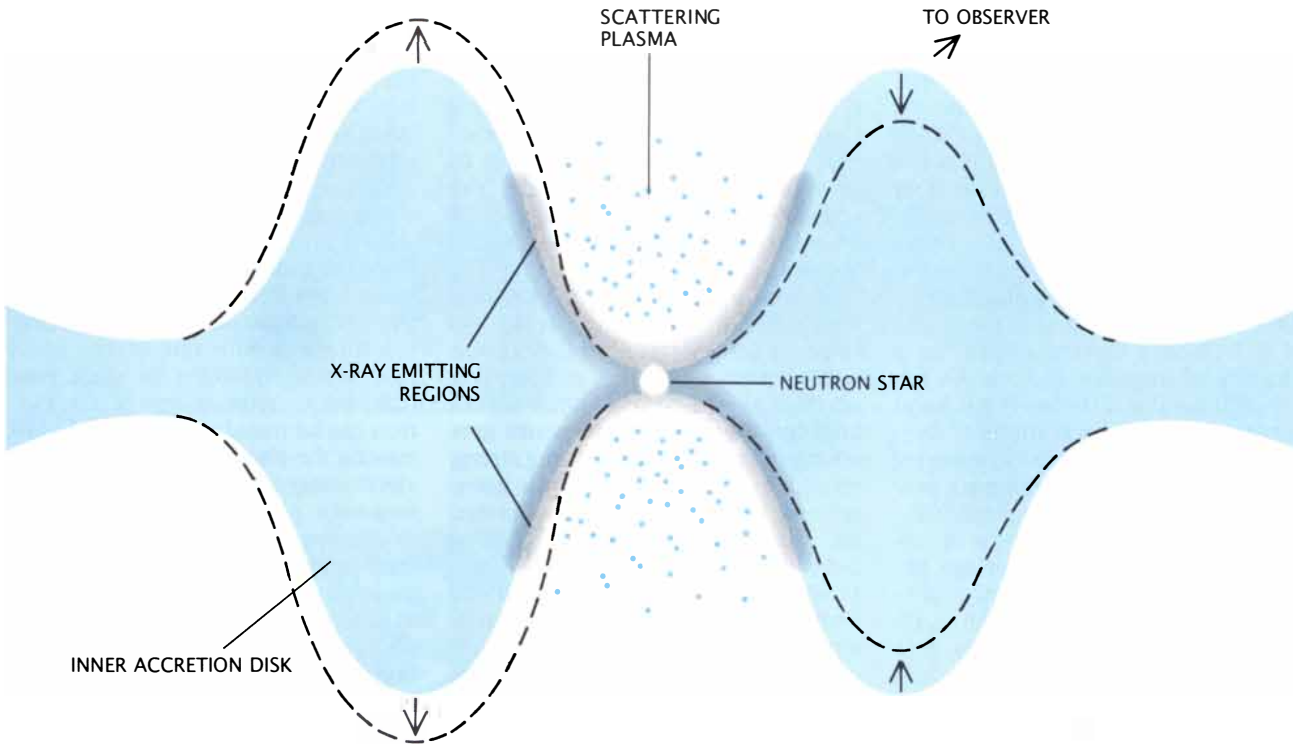
From the model it could then be inferred that the spin period of the neutron star at the heart of GX5-1 would have to be about 10 milliseconds and its magnetic field strength would have to be about a billion gauss, 1,000 times weaker than that of a young neutron star. The height of the magnetosphere above the stellar surface would then be a mere 60 kilometers. All of this was compatible with existing ideas about old neutron stars, and also with the suspicion that radio pulsars that have millisecond periods evolve from low-mass X-ray binaries [see "The Oldest Pulsars in the Universe," by Jacob Shaham; SCIENTIFIC AMERICAN, February, 1987].

The Alpar-Shaham model assumes that some kind of interaction between the accreting matter and the magnetosphere causes the QPO's. What kind of interaction could it be? One answer to the question was worked out by Frederick K. Lamb and Noriaki Shibazaki at Urbana-Champaign together with Alpar and Shaham. Suppose the plasma orbiting just outside the magnetosphere is not homogeneous but clumpy. Suppose additionally that the plasma can enter the magnetosphere more easily in some places than in others. For example, entry could be easier near the neutron star's magnetic poles. Then every time a clump of plasma passes by a magnetic pole, part of it will enter the magnetosphere, fall to the neutron star and cause a small flash of X rays. The same thing happens every time that clump passes the same pole—which occurs once every beat period—until the plasma clump is exhausted.

The resulting X-ray signal is a quickly fading series of small flashes, each flash separated by one beat period. Such a pulse train is a quasi-periodic signal whose power spectrum contains two distinct peaks: one peak is caused by the individual flashes in the train and another (at zero frequency) is caused by the train as a whole. A zero-frequency peak is in fact a feature of GX5-1's power spectrum. Many plasma clumps can contribute to the total signal at the same time, causing many overlapping pulse trains, but the



SPECTRAL STATES OF AN X-RAY SOURCE can be distinguished by plotting its radiation's hardness (the ratio of the number of high-energy X-ray photons to low-energy ones) as a function of the radiation's total intensity or, as is shown here for Cygnus X-2, by plotting the hardness in one X-ray spectral band as a function of the hardness in another band. The diagram reveals three branches connected in the form of a Z that correspond to the three different spectral states Cygnus X-2 adopts. Fast QPO's are observed when its X-radiation falls in the top horizontal branch, whereas "slow" QPO's are observed when its X-radiation falls in the middle diagonal branch.



EXPANSIONS AND CONTRACTIONS of a puffed-up, doughnut-shaped inner accretion disk explain slow QPO's in the so-called obscuration model. The plasma's motion intermittently reveals the X-ray-producing regions near the neutron star.

signal's power spectrum nonetheless would retain the same basic form.

From the beginning there was an obvious potential difficulty with this model, however. If at the heart of GX5-1 there really is a neutron star spinning at a rate of hundreds of revolutions per second and with a magnetic field strong enough to form a magnetosphere, then why is an extremely fast X-ray pulsar not observed? After all, a magnetosphere ought to channel the plasma surrounding the star to its magnetic poles, where it ought to form hot spots. For many astrophysicists that problem was reason enough to be skeptical about any magnetospheric QPO model.

Yet the lack of a detectable X-ray pulsar is not quite as devastating to the model as it may appear to be. The neutron star in GX5-1 has a magnetic field that is much weaker than the field of a young neutron star. Hence its hot spots can be much bigger than those of young neutron stars in binary systems—possibly even covering most of the star's surface—so that the "lighthouse effect" of the hot spots is much diminished. In addition large amounts of plasma surrounding the magnetosphere may smear out whatever X-ray pulsations there are by scattering the X-ray photons several times.

Actually a greater challenge for the magnetospheric beat-frequency model came as other QPO sources were found by means of *EXOSAT*. These new sources exhibited oscillations in X-ray intensity with a bewildering variety of properties, often quite different from those of GX5-1. In many cases the new QPO's had rather low frequencies of between five and seven hertz. Moreover, instead of increasing in conjunction with a brightening of the source's X-radiation, the frequencies either remained the same, decreased slightly or varied erratically. In addition their power spectrums did not always have a peak near zero frequency, as is seen in the power spectrum of GX5-1 and as is predicted by the magnetospheric beat-frequency model.

Many astrophysicists now think that these properties are representative of a second, "slow" type of QPO. They base their judgment mainly on the fact that each type of QPO has been found to be associated with a characteristic spectral state, or distribution of X-ray-photon energy. Spectral states can be clearly distinguished by plotting a source's X-ray "hardness" (the ratio of the number of high-energy X-ray photons to low-energy ones) as a function of its X-ray intensity.

It had already been known long before QPO's were discovered that when the measured spectral data of GX5-1 and a very similar source named Cygnus X-2 were plotted on such hardness-intensity diagrams, they fell into two distinct branches, one horizontal and one diagonal, that merged at the upper right-hand corner of the diagram. With further QPO observations, it became clear that "fast" QPO's were seen only when the X-ray source was emitting radiation in the horizontal, upper branch and slow QPO's were seen only when the source's radiation was in the diagonal branch.

Unfortunately there was a (literally) glaring exception to this rule: Scorpius X-1, the brightest X-ray source in the nighttime sky. The source, which was revealed to be a QPO source in May, 1985, by John Middleditch and William C. Priedhorsky of the Los Alamos National Laboratory, exhibited behavior that superficially resembled that of other QPO sources: in one spectral state it exhibited slow QPO's and in another state it exhibited faster ones whose frequency (between 10 and 20 hertz) was correlated with X-ray intensity. Nevertheless, when it was considered in more detail, Scorpius X-1 was decidedly different from other QPO sources. In particular its two spectral-state branches seemed unique: they joined at the lower left-hand corner

rather than the upper right-hand corner of the hardness-intensity diagram.

The key insight that resolved this confusing situation was provided by Günther Hasinger of the Max Planck Institute for Extraterrestrial Physics. He proposed that the spectral-state branches in the hardness-intensity diagrams for Scorpius X-1, GX5-1 and Cygnus X-2 were in fact parts of the same overall pattern consisting of three branches arranged in the shape of a Z. Such a pattern allows for a juncture of branches in both the upper right-hand and the lower left-hand corners. Further observations of Cygnus X-2, as well as other X-ray sources, beautifully confirmed Hasinger's prediction [*see illustration on page 56*].

More than likely the QPO's of between 10 and 20 hertz that are observed when Scorpius X-1 is in a spectral state in the bottom branch of its hardness-intensity diagram are just another form of slow QPO's. One reason to think so is that slow QPO's are observed to gradually transform into "third branch" QPO's as the spectral state of Scorpius X-1 rounds the corner from the diagonal branch into the third, bottom branch. Another reason is that the third-branch QPO's, like the slow QPO's, do not have a zero-frequency power peak.

Most of the QPO sources known to date fit into the Z spectral-state pattern, although here too there are some exceptions. One particularly pathological exception is the X-ray source called the Rapid Burster. Employing the Japanese satellite *Hakucho,* Y. Tawara of Nagoya University and his colleagues found in 1982 slow oscillations with a frequency of about two hertz in this source. In addition to not conforming to the Z pattern, the Rapid Burster's oscillations differ from slow QPO's in other respects as well.

Although ways can be found to fit most of the discrepant properties of slow QPO's into the magnetospheric beat-frequency model, many astrophysicists now think slow QPO's require an entirely different explanation. A large number of alternative models have been proposed that could explain them, some of which do not rely on magnetospheric effects at all. One such model, conceptually very simple, was examined by Luigi Stella and Nick E. White of the *EXOSAT* observatory team and me in an attempt to explain the slow QPO's of Scorpius X-1. The model takes as its premise two key properties of all known QPO sources: the accretion disk comes close to the surface of the compact object, and the X-radiation is exceedingly bright. (Actually it remains to be seen whether the second property is really characteristic of all QPO sources or is just an illusion created by the fact that QPO's are easier to detect in bright X-ray sources.) The coincidence of the two properties (if true) has important consequences for QPO sources.

The enormous number of X-ray photons emitted in the vicinity of the compact star exert a considerable outward pressure on the surrounding plasma. As the X-ray source shines brighter, the radiation pressure gets stronger. Indeed, it can become strong enough to overcome the intense gravitational force drawing in the plasma; the X-ray photons can literally blow away the plasma. The X-ray brightness at which this happens is named the Eddington limit, for the British astronomer Arthur S. Eddington. For a typical neutron star the limit corresponds to about 100,000 times the luminosity of the sun. Nearly all QPO sources have X-ray luminosities that are close to their Eddington limits; theoretically they cannot get any brighter, since the radiation pressure would then prevent accretion.

Yet even luminosities just below the Eddington limit have important consequences. In particular, it is believed such luminosities can generate enough radiation pressure to cause the innermost region of the accretion disk to swell up, so that it looks like a doughnut [*see illustration on preceding page*]. Depending on one's viewing angle, such a plasma doughnut can partially or totally obscure the X rays emanating from its central hole, where the compact star would be. If the plasma doughnut were to expand and contract in a direction parallel to the axis of rotation, it could thereby produce QPO's, since it might alternately expose and obscure the X-ray source to an observer. Although the properties of such puffed-up accretion disks are not well understood, it can be shown from elementary principles of accretion-disk physics that such a plasma motion would have the correct frequencies to account for slow QPO's.

How might astronomers probe the inner regions of bright, low-mass X-ray binary systems to test the magnetospheric beat-frequency model, the plasma-doughnut model and others? One way is to measure the time lags between the arrival of high-energy and low-energy X rays from QPO sources. Hasinger has recorded such time lags by measuring the delay between corresponding oscillations in high- and low-energy X-radiation coming from Cygnus X-2. Similar time lags have also been recorded in the QPO's of GX5-1. The difference in arrival time is small—only a few milliseconds—but quite significant.

It is very likely that the time lags are caused by a process called inverse Compton scattering, which takes place when photons are scattered by an extremely hot plasma. Every time a photon interacts with one of the many fast-moving electrons in such plasmas, some of the energy of the electron can be transferred to the photon, moving the photon's energy along the electromagnetic spectrum. As a consequence, photons that are scattered many times end up as high-energy X rays, whereas those that happen to undergo only a few collisions remain at low energies. Moreover, because the photons that end up as high-energy X rays have been scattered many more times than those that remain at low energy, they take a longer time to emerge from the plasma. (It is this smearing out of photon arrival times by inverse Compton scattering that may be responsible for the fact that no millisecond X-ray pulsars have yet been detected.)

The extent of the time lag depends on several important factors, such as the scattering plasma's temperature, density and size. Hence it is possible to obtain direct observational information about the conditions in the innermost regions of low-mass X-ray binaries by recording and analyzing the lags between the high- and the low-energy X rays associated with a particular oscillation.

This is one example of how QPO's are opening windows onto one of the obscurest regions of high-energy astrophysics. We have only just begun to explore the views they offer us.

FURTHER READING

INTENSITY-DEPENDENT QUASI-PERIODIC OSCILLATIONS IN THE X-RAY FLUX OF GX5-1. M. van der Klis, F. Jansen, J. van Paradijs, W. H. G. Lewin, E. P. J. van den Heuvel, J. E. Trümper and M. Sztajno in *Nature,* Vol. 316, No. 6025, pages 225-230; July 18, 1985.

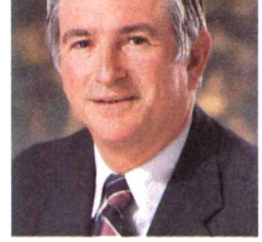
THE PHYSICS OF ACCRETION ONTO COMPACT OBJECTS. In *Lecture Notes in Physics,* Vol. 266; 1986.

THE PHYSICS OF COMPACT OBJECTS: THEORY VS. OBSERVATIONS. In *Advances in Space Research,* Vol. 8, No. 2-3; 1988.

A REVIEW OF QUASI-PERIODIC OSCILLATIONS IN LOW-MASS X-RAY BINARIES. W. H. G. Lewin, J. van Paradijs and M. van der Klis in *Space Science Reviews,* Vol. 46, pages 273-377; 1988.

A message from John V. Roach

Global competitiveness in an age of technology begins in the classroom...



Chairman of the Board
Chief Executive Officer
Tandy Corporation

...America's Challenge

Innovation, quality and competitiveness are at the center of today's debate on America's future. We have the opportunity for a Renaissance in competitiveness in the world marketplace because of the weakness of the dollar, technology advances in production, and greater sensitivity of managers and workers alike in the need for dedication to competitiveness.

A vital element in future competitiveness is the dedication and excellence of students and teachers. The complexities of today's technology requires engineers and scientists with strong foundations and high levels of intellectual curiosity. In order to help meet our country's challenge, we are going to provide additional recognition and a greater incentive to those high school students and teachers who are willing to make the sacrifices today, so we can be more competitive tomorrow in a world where technology is a major driving force.

For students and teachers demonstrating excellence in high school math, computer science and the sciences, Tandy Corporation wants to reward, honor and encourage you to become Tandy Technology Scholars. We know our future and America's competitiveness is dependent on those who attain excellence in rigorous disciplines.

We hope everyone will join us in encouraging even greater numbers of individuals to become Tandy Technology Scholars. The thousands chosen this year will be recognized in both public and private ways and hundreds will receive scholarships or cash stipends.

America needs a renewed dedication to pride in product, pride in quality and pride in country which can only be achieved through a serious commitment to educational excellence.

This school year's performance counts!

The top two percent of students completing their junior year in May of 1989 and recommended teachers of science, math or computer science for the 1988-89 school years at qualifying and participating public and private high schools will be the initial awards recipients. All qualifications, selection criteria and selections will be made by Texas Christian University and its national committee of education leaders. Parents, students and educators should alert your principal to enroll when your school receives the enrollment form in early 1989.

TANDY TECHNOLOGY SCHOLARS

P.O. Box 32897, TCU Station, Fort Worth, Texas 76129

What Makes a Tumor Cell Metastatic?

The cells that spread cancer throughout the body have distinctive molecular traits. Already cancer researchers know enough about those traits to convert malignant cells into benign ones

by Michael Feldman and Lea Eisenbach

The main clinical problem of cancer is metastasis, the migration in the blood and lymph system of tumor cells that give rise to tumors elsewhere in the body. If cancer cells did not metastasize, the surgical removal of a tumor would cure a cancer patient. Instead, by the time the "primary" tumor is detected cells have already escaped from it and settled in other organs, sowing the seeds for "secondary" tumors there. Thus begins a race against malignancy that the patient rarely wins.

Cancer would no longer be such a formidable enemy if metastasis could be controlled. And indeed, the understanding of metastasis is progressing rapidly. Fifteen years ago, when we started our research at the Weizmann Institute of Science in Israel, it was already clear that some tumors are more metastatic than others. It was also apparent that the cells in any given tumor differ in their ability to metastasize: some cells are much better at it than others. These facts constituted the basis of investigations in which we demonstrated that certain characteristics distinguish cells capable of metastasis from those that cannot metastasize.

Our findings have yielded clues to just how cells metastasize—and how they can be prevented from doing so. Recently we applied the principles drawn from our research to achieve successful immunization against the metastatic spread of lung cancer in mice. Such experiments may lay the groundwork for a "cancer vaccine."

There is plenty more work to be done. Characterizations of metastatic attributes are far from complete, and it still is not clear to what extent our findings, which are based on experiments with mice, can be applied to human beings. We are also investigating the connection between metastasis and oncogenes (genes whose activation is involved in tumor formation), for it is not known whether the genes that spur uncontrolled growth are also responsible for cancer's spread. We hope our investigations will help to answer some related questions, such as why some tumors are more metastatic than others, and why brain tumors do not metastasize even though tumors in other organs can spread to the brain. Of course, our work is just a small part of the enormous, ongoing effort to determine the molecular basis of malignancy.

The earliest indication that the cells in a single tumor might differ in their ability to metastasize came from studies done in the 1970's by I. Joshua Fidler and Margaret L. Kripke, who were then at the National Cancer Institute. Fidler and Kripke were using mice to study melanoma, a skin cancer that tends to metastasize to the lungs. To test the metastatic competence of individual cells, they collected cells from a tumor, separated them from one another and grew a "clone" from each cell. (All the cells in a cloned population are identical because they are all descended from a single cell.)

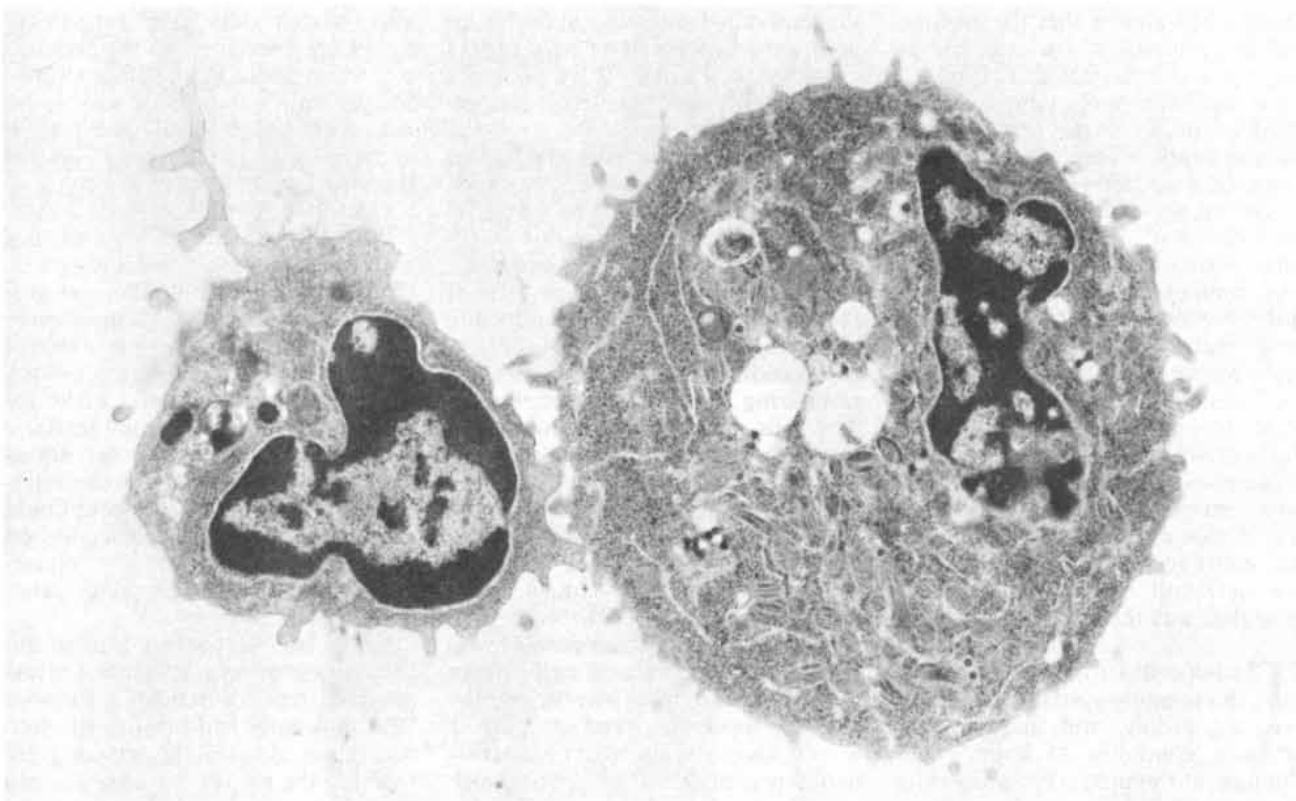
Fidler and Kripke transplanted cells from the clones into healthy mice and then looked for lung metastases. They found that certain clones produced many more metastases than others, and some did not generate any lung tumors at all [*see illustration on pages 64 and 65*]. It seemed that metastasis was a specialized job.

In retrospect, it is not surprising that not all tumor cells can metastasize; indeed, it is surprising that any of them can. A metastasizing cell must break loose from its parent tumor, invade the matrix between cells and penetrate the membrane of a blood vessel. It must survive its passage in the bloodstream, where hostile cells of the immune system patrol, and emerge from the bloodstream at a favorable spot. After lodging in the surrounding tissue it must induce the growth of new blood vessels to nurture the nascent tumor [*see bottom illustration on pages 62 and 63*].

Each of those steps is probably controlled by a different molecular system; a failure in any one of the systems would most likely render a tumor cell incapable of metastasizing. Hence the molecular differences between metastasizing and nonmetastasizing cells could be many, or they could be few. We began our search for the distinguishing features by focusing on enzymes required for the initial steps of tumor-cell dissemination.

One such enzyme is called colla-

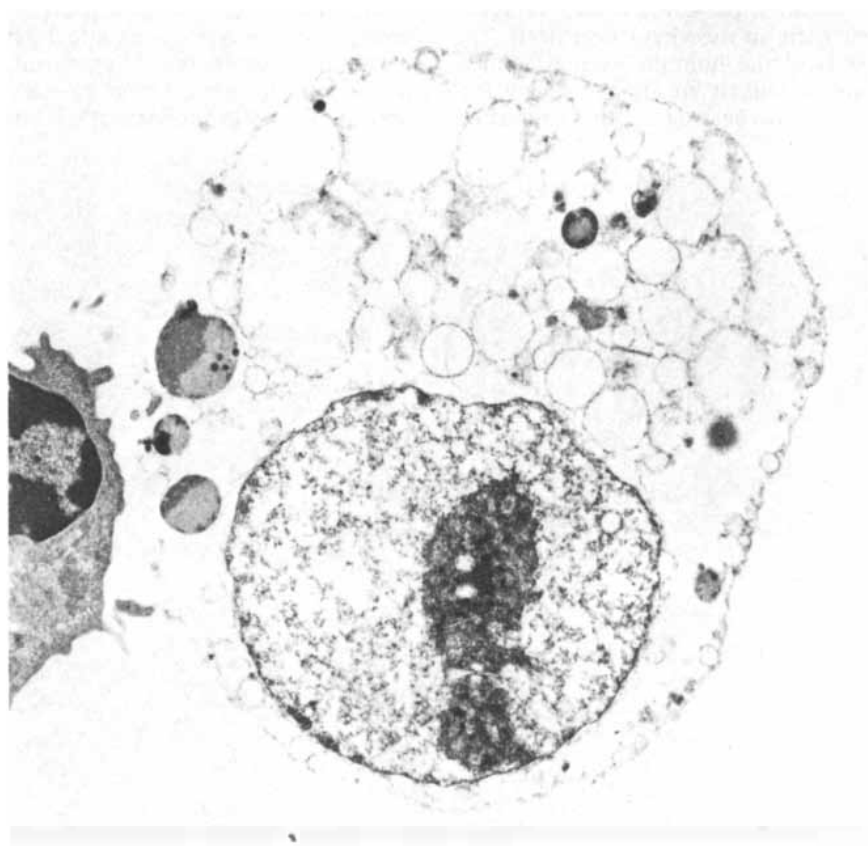
MICHAEL FELDMAN and LEA EISENBACH work together at the Weizmann Institute of Science in Rehovot, Israel. Feldman, who is Klutznik Professor of Developmental Biology and head of the department of cell biology, has been at Weizmann since 1955. He holds a Ph.D. (1953) from the Hebrew University of Jerusalem and has spent sabbaticals at the University of California at Berkeley, the Stanford University Medical Center and the National Cancer Institute. Feldman was elected a member of the Israel Academy of Sciences and Humanities in 1977; his research on cancer and immunology won him a Griffuel Prize in 1984. Eisenbach joined Feldman's laboratory in 1980 and is now staff associate scientist in the department of cell biology. She was awarded a Ph.D. from the Feinberg Graduate School at Weizmann in 1979 and did postdoctoral work for two years at the University of Wisconsin at Madison before returning to Israel.



genase type IV, which eats through blood-vessel walls by degrading a major structural component called collagen IV. Earlier studies had shown that benign and malignant tumors differ in their ability to make and secrete collagenase. We wondered whether cells in a single malignant tumor also produce different amounts of the enzyme.

In these experiments and most subsequent ones we used two common mouse tumors known as the 3LL (for "Lewis lung") carcinoma and the T10 sarcoma. Carcinomas and sarcomas are two types of solid tumors: a carcinoma is a tumor of the epithelial tissue, such as the breast, colon or lung, and a sarcoma is a tumor of the connective tissue. We cloned cells from the two tumors and tested clones, selected at random, for metastatic competence. At the same time we measured collagenase synthesis and secretion by each of the clones. We found that levels of collagenase production were comparable in metastatic and nonmetastatic clones. Apparently both kinds of cells can carry out the steps in tumor dissemination that are controlled by collagenase.

We then compared, in clones of metastatic and nonmetastatic cells, the amounts of another enzyme that was involved in blood-vessel penetration.



KILLER CELL ATTACKS and destroys a cancerous cell. The killer cell (*top left*) recognizes and attaches to the tumor cell by means of unique molecules on the tumor cell's surface. Then the killer cell produces a substance that makes the tumor cell burst (*bottom*). In the body many cells that detach from tumors meet this fate, but some cells have certain surface molecules that enable them to evade immune detection. These cells stand a much better chance of surviving the metastatic journey.

Studies had shown that the enzyme, called plasminogen activator, has an important role in metastasis. Indeed, we learned that metastatic clones do produce more of the enzyme than nonmetastatic clones. The disparity, however, does not seem to influence metastatic potential, since both clones are capable of invading the intercellular matrix. Later it became clear why: nonmetastatic cells somehow induce the accumulation of blood cells called macrophages, which secrete copious amounts of plasminogen activator. Thus nonmetastatic cells can penetrate blood-vessel walls by "borrowing" enzymes from macrophages.

Our work with cellular enzymes led us to suspect that the interactions that dictate a cell's ability to metastasize must include other elements in the metastatic cascade. One obvious candidate was the immune system.

Tumor cells encounter the cells of the immune system while growing locally, and immune surveillance continues to hound cells throughout the metastatic process: in the intercellular matrix, during penetration of the blood vessels and particularly in the circulation itself. The cells of the immune system cooperate to launch an attack on any foreign body bearing distinctive surface markers called antigens. On the "front line" are white blood cells known as cytotoxic, or "killer," *T* lymphocytes whose job it is to eliminate antigen-bearing cells.

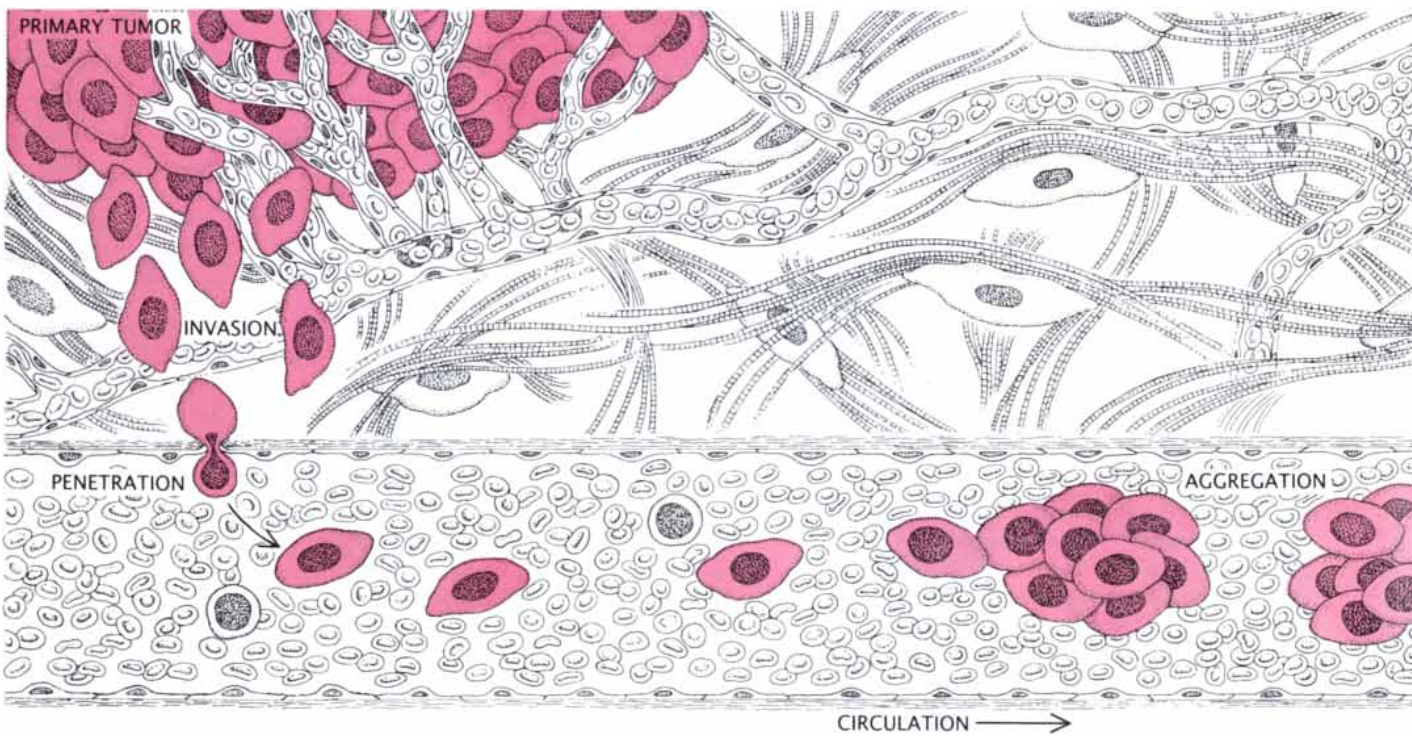
Even though cancer cells are derived from the body's own tissue, they start to generate unique "tumor antigens" when they become cancerous. A tumor cell thus undergoes a transition from an entity the immune system recognizes as "self" to an entity the immune system regards as foreign. We considered the possibility that metastasizing and nonmetastasizing tumor cells might differ in the degree to which killer cells recognize them as foreign.

Our first impulse was to compare the immunogenicity of tumor antigens from metastatic and nonmetastatic clones, but these antigens have not yet been isolated or characterized. An alternative strategy was to examine another group of cell-surface molecules that take part in routine immune recognition and are part of what is known as the major histocompatibility complex, or MHC. The choice turned out to be fortunate indeed.

MHC molecules are glycoproteins (proteins with sugar chains attached) that act like a biochemical fingerprint: their composition is different in every individual. They are the signal the immune system calls "self." Foreign antigens are presented to the immune system anchored to an MHC molecule. The immune system "sees" the combination of MHC molecule and foreign antigen—"self" and "other"—rather than the foreign antigen alone.

MHC molecules are divided into two classes, but only the class I molecules are directly involved in killer-cell interactions. These molecules are displayed on the surface of almost every cell in the body. They serve as a kind of docking base for killer lymphocytes, which possess receptors specific for particular MHC-antigen combinations. In organ transplants they can act as antigens themselves and evoke rejection of the transplanted tissue. Could differences in the MHC molecules on cancer cells impinge on killer-cell recognition and as a result on metastatic competence?

In the late 1970's this was an unusual question to ask, because it was assumed that an individual's class I MHC molecules had functionally similar, if not identical, immunizing potency. In the mouse, for example, two genes constitute the class I MHC's: these genes and their products are called H-2K and H-2D. Most cells express both H-2K and H-2D in varying amounts. Except for a few scattered reports, there was no reason to sus-



METASTASIS is the establishment of a new, "secondary" tumor (*right*) by a cancer cell that travels through the bloodstream from a "primary" tumor (*left*). The metastasizing cell must first invade the matrix surrounding the primary tumor and then penetrate blood-vessel walls to enter the circulation. In the bloodstream the cell seeks out a suitable location for the formation of a new tumor; along the way it may aggregate with other metastasizing tumor cells. When the cell reaches a tar-

pect any distinction between the two gene products.

One of the reports, however, was particularly interesting. In it Henry N. Claman of the University of Colorado Health Sciences Center suggested that H-2D, but not H-2K, might be involved in the development of immunological tolerance—the failure of the immune system to react against a particular antigen. Perhaps, we thought, the MHC molecules of a single individual do vary in their immunogenicity; perhaps H-2D has a suppressive effect on immune recognition. In that case a tumor antigen associated with H-2K would draw killer cells, whereas a tumor antigen attached to H-2D would remain invisible to them. Perhaps the degree of expression of the H-2K and H-2D genes in a tumor cell had some bearing on the cell's ability to complete metastasis.

We decided to test that hypothesis with our mouse 3LL carcinoma. We had already developed from previous experiments a highly metastatic clone called *D*122 and a virtually nonmetastatic clone called *A*9. We discovered that, consistent with our speculation, *D*122 cells had a high density of H-2D but almost no H-2K on their surface, and *A*9 cells displayed H-2K and H-2D in roughly equal amounts.

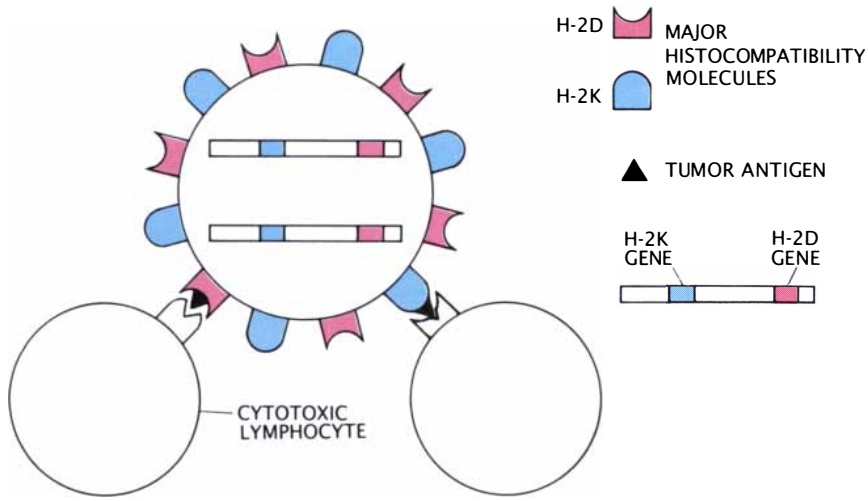
Do the differences in H-2K and H-2D molecules correspond specifically to differences in immunogenicity? We could examine this question by injecting *D*122 and *A*9 cells into healthy mice and observing their immune reaction. To be certain the reaction was prompted by the tumor antigen-MHC complex, and not simply by recognition of a foreign MHC, we had to use mice whose MHC's were identical with those expressed by the tumor cells. Such so-called syngeneic mice are produced by intensive inbreeding; their MHC "fingerprints" are all the same.

The tumor cells we injected had been exposed to a dose of X rays so that they would not form tumors. Two weeks after injection we collected killer lymphocytes from the mice and tested the interactions of the killer cells with tumor cells in culture. Sure enough, animals injected with the metastatic *D*122 cells showed very weak immune responses; they generated few killer cells that attacked the cultured tumor cells. The *A*9 cells, on the other hand, elicited a strong immune response: animals injected with them generated killer cells that destroyed both *A*9 and (to a lesser extent) *D*122 cells in culture.

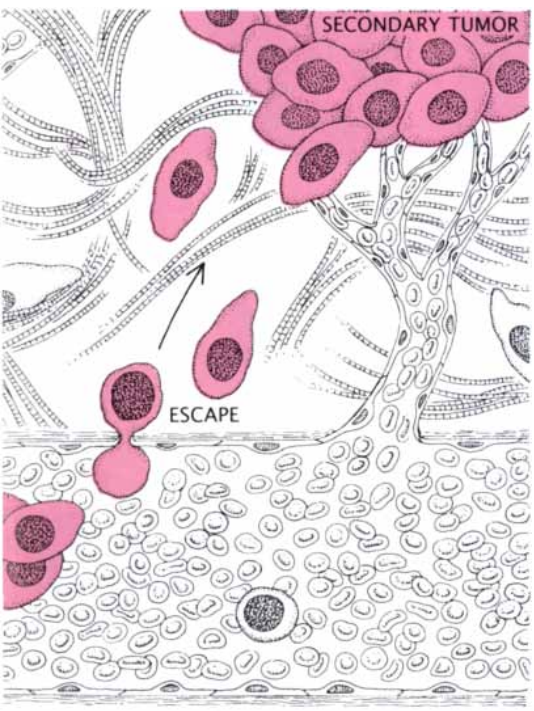
We tested more than 90 3LL clones to see whether the correlation between MHC type and metastatic competence held. It did, and we learned more: that the ratio of H-2K to H-2D on the surface of a tumor cell, rather than the absolute density of H-2K, was at the root of the correlation. We deduced that the H-2D–antigen complex is not simply immunologically inert: it might actually suppress the immune response, as Claman's work had suggested.

Having established the correlation, our next step was to determine whether we could claim a cause-and-effect relation. We could explore this possibility by altering the relative expression of the H-2K and H-2D genes in cultured tumor cells and then looking to see whether or not the metastatic potential of the cells had changed. Initially we altered gene expression by culturing cells with substances that were known to turn on several genes, not just those for the MHC. Our results were encouraging. Tumor cells cultured with retinoic acid, for example, showed increased expression of H-2D, and after seven days in culture with retinoic acid the ordinarily nonmetastasizing *A*9 cells were capable of generating large masses of lung metastases in syngeneic mice.

Treatment with the various kinds of interferon yielded results that were consistent with our hypothesis but that were surprising nonetheless. It seems that the different interferons—alpha, beta and gamma—affect different genes. Alpha- and beta-interferons administered in tandem switch on the H-2D gene more than the H-2K gene, whereas gamma-interferon switches on the H-2K gene more than the H-2D gene. Consequently the H-2K/H-2D ratio in cells treated with alpha- and beta-interferons dropped, and both



SURFACE MOLECULES of the major histocompatibility complex (MHC) identify a cell as "self" and aid in presenting antigens, the distinctive markers the immune system recognizes as foreign. In mice two types of MHC molecules are made by two different genes: H-2K and H-2D. (Mice, like human beings, have two copies of each gene type.) Cytotoxic, or "killer," *T* lymphocytes specific for each MHC type are stimulated to destroy a cell when they interact with MHC molecules that display foreign antigens. Cells produce tumor antigens when they become cancerous, but the H-2D MHC molecule may interfere with the recognition of such antigens by killer cells.



get site, it escapes from the blood vessel and lodges in the adjacent tissue. There it proliferates and induces the growth of a network of nurturing blood vessels.

*A*9 and *D*122 cells became more metastatic. Gamma-interferon increased the amount of H-2K in relation to H-2D, and treated *D*122 cells gave rise to a much smaller mass of metastases.

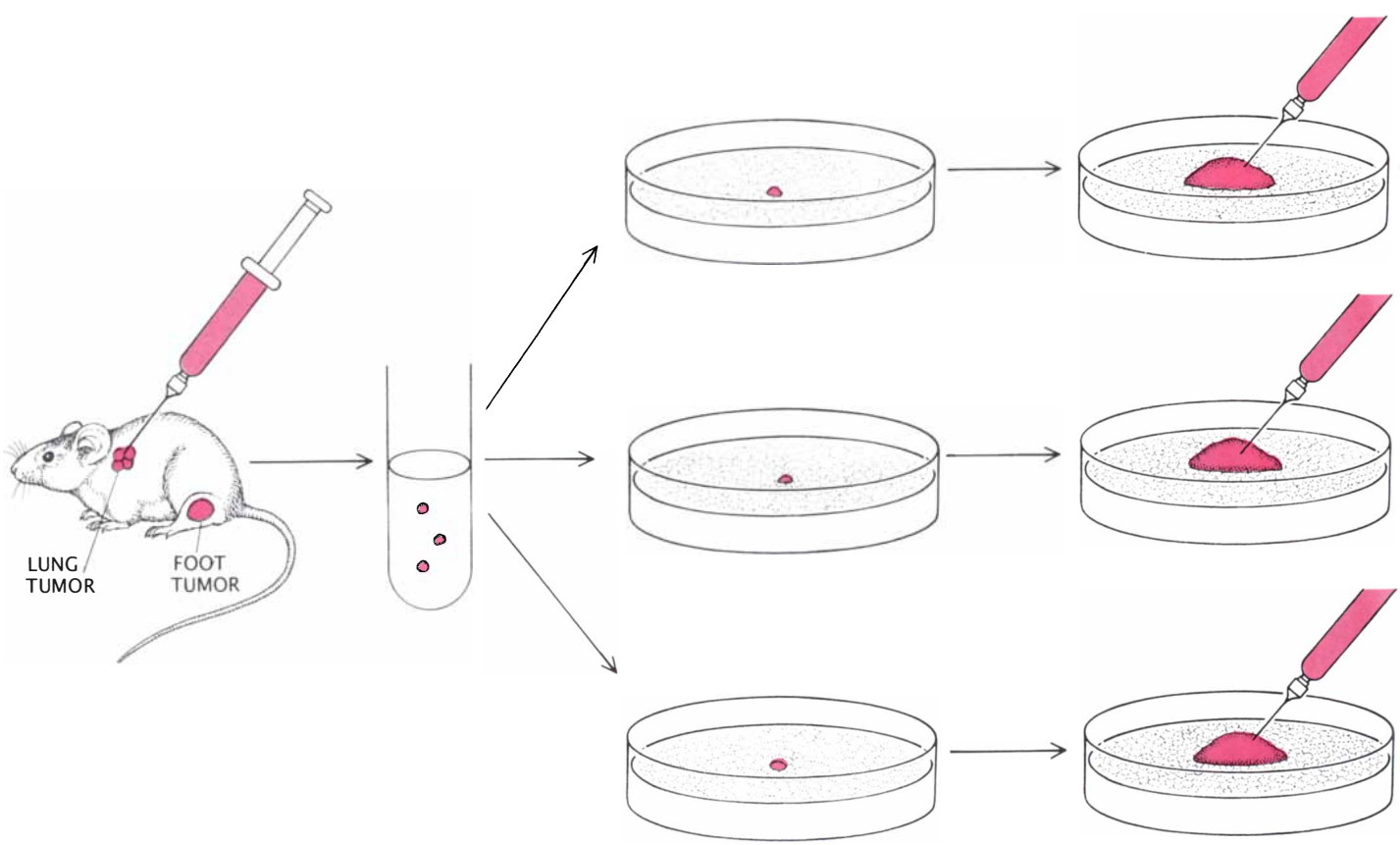
We still could not be sure that the changes in metastatic competence we observed were the result specifically of altered MHC ratios, because the effects of retinoic acid and interferon are not confined to MHC genes. Fortunately the techniques of genetic engineering allow more precise manipulation of gene expression. In particular, the H-2K gene had been isolated from normal mouse cells, in which it is abundantly expressed, and could be transferred to the *D*122 clone, in which there is virtually no expression of the native H-2K gene. Any difference in *D*122 metastatic potency following such a transfer could be attributed to the activity of the H-2K import.

A graduate student in our laboratory, Daniel Plaksin, performed such a transfer, which is known as transfection, with circular pieces of DNA called plasmids [*see illustration on page 68*]. The plasmids act as vectors for carrying the normal cell's H-2K gene into the *D*122 cell's nucleus, where they insert the gene into the *D*122 cell's DNA. In fact, we found that such H-2K "transfectants" often harbor as many as six copies of the imported H-2K gene. The level of H-2K expression in these cells exceeds that of H-2D; indeed, the H-2K/H-2D ratio is higher in the *D*122 transfectants than it is even in the nonmetastatic *A*9 clone.

When we transplanted the transfected clones into syngeneic mice, we observed that tumors developed at the site of transplantation, but in seven out of nine cases the load of lung metastases was small or nonexistent. Here was unequivocal proof that alteration of the expression of MHC molecules can convert a highly metastatic clone into a clone with almost no metastatic potential.

We repeated our experiments with H-2K genes from several other mouse strains and got similar results. Anxious to see if our finding had relevance for other tumor types, we also extended our investigation to the T10 sarcoma. With Shulamit Katzav, a graduate student in our laboratory, Shraga Segal, who was then in our laboratory and is now at the Ben Gurion University of the Negev, and Günther J. Hämmerling of the German Center for Cancer Research in Heidelberg, we carried out experiments in the sarcoma analogous to those we had done with the 3LL clones. The results were virtually the same: we found we were able to abolish the metastatic properties of tumor cells by gene transfer.

Immediately our thoughts turned to the possibility of using H-2K transfectants, which are highly immunogenic, to immunize mice against tumor metastasis. Was it feasible? In the course of ordinary malignancy the immune system seems to strike down metastasizing cells with a high H-2K/H-2D ratio while leaving cells with low levels of H-2K expression alone. If an animal's immune response were powerful enough, however, even tumor cells



TESTS FOR METASTATIC COMPETENCE indicate that the cells in a single tumor may differ in their ability to metastasize. Cells taken from a mouse lung tumor were separated and cultured individually; each gave rise to a population of identical cells called a clone. Cells from each clone were transplanted into mice. All the clones gave rise to a primary tumor at

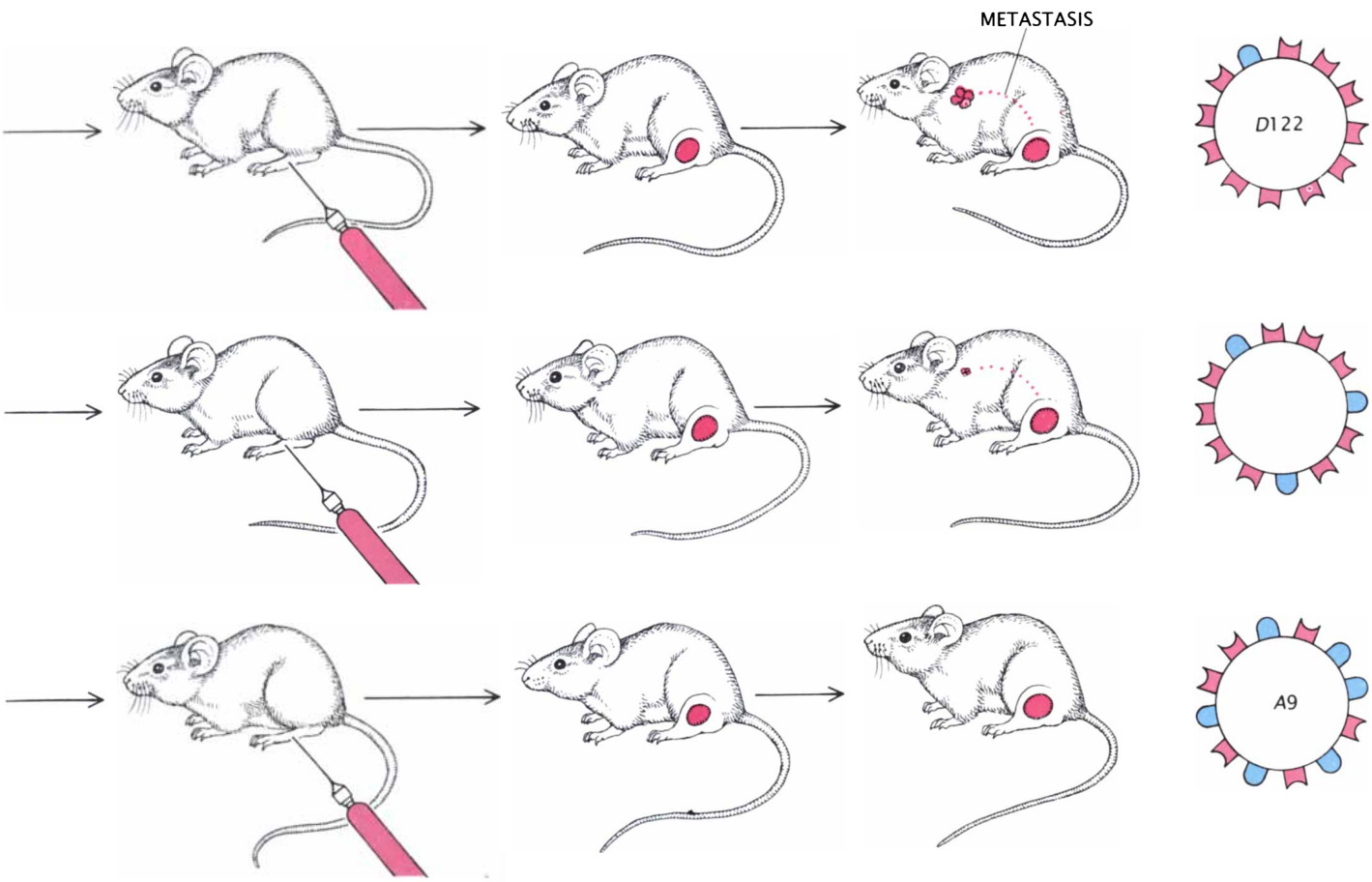
with a low density of H-2K on their surface might become susceptible.

Plaksin and Cochava Gelber, another graduate student in our laboratory, first immunized mice with H-2K-transfected *D*122 cells that had been irradiated to render them inactive. Then they transplanted the potent metastasizers from the *D*122 clone into the mice. Mice that were not immunized acquired heavy loads of lung metastases, as usual, but the metastases did not appear in immunized mice. As a control, some animals were "immunized" with inactivated *D*122 cells; following transplantation with normal *D*122 cells, these animals exhibited sizable loads of lung metastases [*see illustration on page 85*].

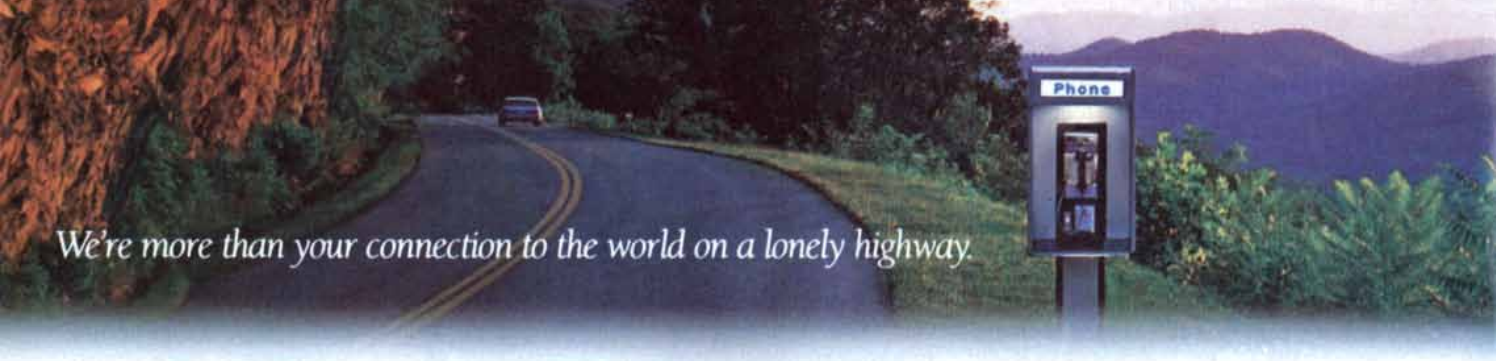
We could, then, prevent metastasis by immunization with H-2K transfectants. Could a similar approach halt metastasis once a tumor has been detected? Plaksin transplanted *D*122 cells into healthy mice; a week later, when a primary tumor had appeared, he began weekly inoculations of some of the mice with H-2K-transfected cells. After three weeks the amount of lung metastases in the immunized animals had dropped significantly compared with unimmunized control mice. Theoretically the same strategy could work for immunotherapy in humans. To ensure that the cells used for therapy are histocompatible, they would be taken from a patient's primary tumor and transfected.

It is worth noting that none of our immunizations succeeded in thwarting the growth of the primary tumor. It appears that by the time the tumor is large enough for the immune system to "notice" it, it is also large enough to resist the immune reaction, even when the immune system has been prepared by immunization. In experiments carried out by Hilliard Festenstein and his colleagues at the London Hospital Medical College, however, cells from a mouse lymphoma could not even establish a local tumor when they were transfected with H-2K before being transplanted. The transfected cells could, on the other hand, grow in mice whose immune systems had been suppressed by irradiation. It seems the immune system may even be able to defeat primary tumor growth in certain instances.

The preceding experiments demonstrate that MHC molecules are at least one of the determinants of metastatic competence, but are they the only determinant? We could answer this question easily enough by comparing the amount of metastasis that occurs in immunosuppressed mice transplanted with metastatic and with nonmetastatic clones. If ratios of MHC molecules are the only factor involved in metastatic potency, the metastatic loads in these mice should be the same. In fact, nonmetastatic clones do generate lung metastases in immunesuppressed mice, but the load of metastases generated in such animals by metastatic clones is still higher. Hence some other property or properties must weigh into the determination of metastatic competence. To focus our search for those properties, we decided to examine the products of oncogenes.



the site of the transplant, but they generated widely varying amounts of metastases in the lungs. The authors found that a clone's metastatic potential is related to the ratio of H-2K to H-2D MHC molecules on the cell surface; the higher the ratio, the less metastatic the clone (*right*). They named one of the most metastatic clones *D*122 and one of the least metastatic *A*9.



Most of is unheard

With a first name like Bell, it's no surprise Bell Atlantic is in the telephone business. But what is surprising is how much more than telephone service Bell Atlantic provides businesses around the world.

Did you know, for example, thousands of businesses come to us every year to help them find better ways to manage their information? What they've discovered is Bell Atlantic has the people and the technological expertise to design voice and data networks to meet their individual needs.

Here's something else you should know about us. We're the leading independent computer-maintenance organization in the world. To our over 45,000 customers we're known as Sorbus.SM And we service over 3,000 makes and models of computers and

LINKING UP AT WESTINGHOUSE.

When Westinghouse needed high-capacity voice and data transmission services to connect its 23 different business units worldwide, Bell Atlantic came to their assistance. We provided the local fiber network and high-capacity links necessary to make their internal communications system reliable, speedy, and cost-efficient.

THANK YOU, PAINEWEBBER.

When PaineWebber needed a computer-maintenance source that could increase their computer uptime without letting them down, they sought out Bell Atlantic's Sorbus. Today, Sorbus maintains over 500 IBM terminals in PaineWebber's metropolitan Manhattan offices alone, including printers needed for buy-and-sell orders and late-breaking news. Because Sorbus has a dedicated field engineer, on-site maintenance, in most instances, is performed in less than one hour.

DATA COMMUNICATIONS SM TELECOMMUNICATIONS SERVICES SM

peripherals, from mainframes to modems. In fact, we service more IBM computers than anybody...except, of course, IBM.

And if you think all you can lease from us are phone lines, think about this: our leasing and financing companies can supply your business with anything from desks to desktop computers, corporate helicopters to commercial airliners.

To make all of these unheard-of things happen, Bell Atlantic has over 79,000 unsung heroes. They are our employees, each trained to help you communicate better so you can conduct business better. To hear more about all of the unheard-of things they can do for your business, call us anytime, 24 hours a day.

1-800-444-8838

what we do of.

GETTING AMERICAN OFF THE GROUND.
American Airlines relies on Bell Atlantic for something just as vital as its phone service —airplanes. Through Bell Atlantic Capital Corporation's TriCon Leasing subsidiary, American leases DC-9s and Boeing 767s. Fact is, even if they needed something other than a plane, like a plain new desk, they could get it through Bell Atlantic Capital Corporation.

Bell Atlantic
We're More Than Just Talk.℠

MOBILE COMMUNICATIONS℠ COMPUTER MAINTENANCE℠ FINANCING & LEASING℠

Oncogenes are genes normal cells contain that, when activated by mutation or translocation, contribute to tumor growth, or tumorigenesis. In their natural state the genes are called proto-oncogenes and their expression is often an important part of normal cell activity. The existence of oncogenes was first postulated in the late 1970's, when Robert A. Weinberg of the Massachusetts Institute of Technology showed that transfecting normal cells with the DNA from tumor cells can "transform" the normal cells into cancerous ones. Since then about 40 specific oncogenes have been identified and isolated, but it is still not known to what extent or in what way oncogenes control the acquisition of cancerous properties or—more important from our point of view—if they control the manifestation of metastasis.

One way to approach the issue is to ask: Is metastatic competence linked to transforming efficiency? If it is, the DNA from metastatic and nonmetastatic clones should differ in its ability to transform normal cells. We carried out such DNA-transfer experiments with clones of various metastasizing potencies from both the 3LL carcinoma and the T10 sarcoma lines. We found no correlation between metastatic and transforming proficiency. We infer that the transformation of normal cells to cancer cells and the progression from a benign to a metastatic tumor type are distinct processes that are not dependent on the same oncogenes.

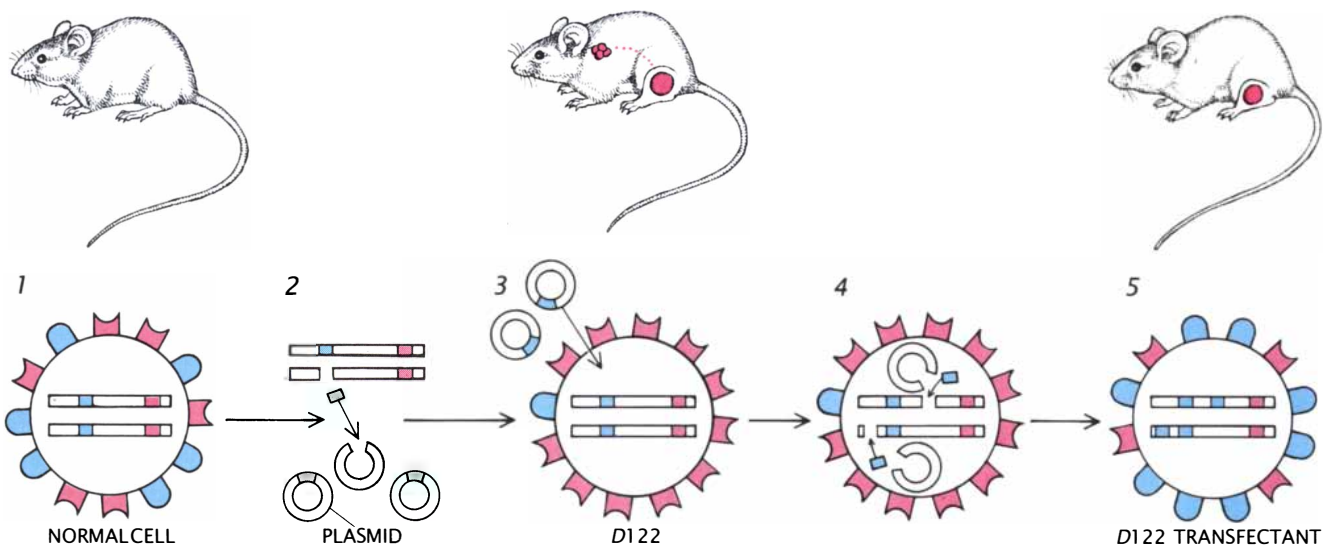
It remained possible, however, that metastatic competence could be determined by oncogenes that are not involved in the initial transformation. We thought it would help our inquiry to know which, if any, of the known oncogenes are operating in our mouse tumor cells. We screened clones from both the carcinoma and the sarcoma with probes that detect many of the known oncogenes. Several of them turned up; one, called *myc,* was present in the carcinoma cells in unusually large numbers (as many as 60 copies per cell), but the amplification is comparable in both the metastatic and the nonmetastatic clones and is therefore probably not connected with metastatic potency.

Another one is more intriguing. To our surprise, the oncogene, known as *fos,* was expressed in the low- and nonmetastasizing clones, but not in the highly metastatic cells. In that respect it resembled the H-2K gene. We wondered whether the *fos* gene could have something to do with the regulation of MHC genes. With Joshua Barzilay, who was then in our laboratory, we showed that the increase in H-2K production that gamma-interferon elicits in the metastatic *D*122 clone is preceded by activation of the *fos* gene. Later another member of our group, Gil Kushtai, transfected *D*122 cells with the mouse *fos* gene and its viral and human analogues. In every case the "dormant" H-2K gene in the *D*122 cells was expressed and H-2K molecules appeared on the cell surface; again, H-2K expression began after the transfected *fos* gene became activated. The *fos* gene, then, seems to switch on expression of the *D*122 cells' native H-2K MHC gene. We recently learned that, as one would expect, the ability to metastasize also decreases significantly when *D*122 cells are transfected with the *fos* gene.

We observed the same connection between *fos* activation and MHC expression in human leukemia cells. *Fos* does not seem to be one of the oncogenes involved in human leukemia, and so in these cells it is not ordinarily activated. But leukemia cells in culture can be induced with various substances to resume the characteristics and growth patterns of normal blood cells, and when they do, the profile of MHC molecules on their surface can change. The maturation to macrophages, for example, has been associated with an increasing density of MHC molecules on the cell surface. The maturation to granulocytes is associated with the disappearance of MHC antigens.

Could *fos* activity be correlated with these changes in the MHC molecules? The simple answer is yes. When we induced human leukemia cells to mature into macrophages, the *fos* gene was expressed and, following that, MHC antigens appeared on the surface of the cells. On the other hand, *fos* was not activated when we induced the cells to mature into granulocytes.

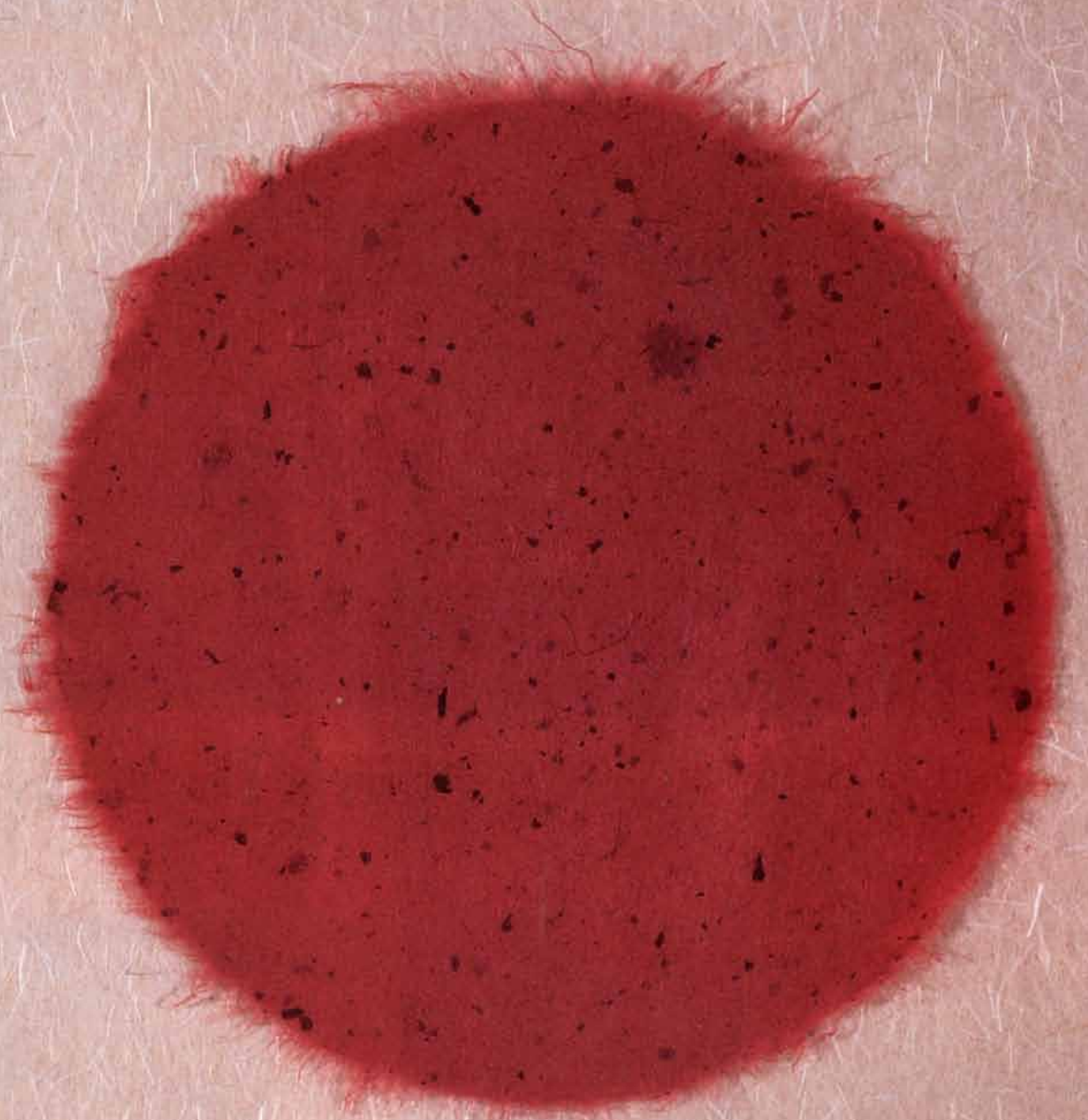
It looks as though our conclusions about the role of the *fos* gene in controlling MHC expression in mouse carcinoma have more general import.



TRANSFECTION of highly metastatic *D*122 cells with the H-2K gene from normal mouse cells reduces the *D*122 cells' ability to metastasize. To achieve transfection the H-2K gene is put into plasmids, bits of circular DNA that are often used to transfer genes (*1–2*). The plasmids insert several copies of the H-2K gene into the genetic material of each of the *D*122 cells (*3–4*). As a result the transfected cells express more H-2K than H-2D and also exhibit a decline in metastatic competence (*5*).

SPECIAL ADVERTISING SECTION

JAPANESE TECHNOLOGY TODAY



One Measure Of A Great Theater Is The

Size Of Its Screen.

With our 60-inch big screen TV, you can literally turn your living room into a screening room. For the Mitsubishi dealer nearest you, call (800) 556-1234, ext. 245. In California, (800) 441-2345, ext. 245. © 1988 Mitsubishi Electric Sales America, Inc.

MITSUBISHI®

JAPAN: AT THE FRONTIERS OF TECHNOLOGY

Introduction by Seishi Katayose

During the 40 years after the end of the war, Japan made remarkable progress. Today, Japan is called an economic giant and is a member of the first-class nations of the world. What has supported this economic development is technology.

The reason that numerous Japanese products including automobiles, home appliances, office automation equipment, cameras, watches and semiconductors have come to be so popular on the world market is that these products are the fruit of advanced technology: Japanese technology for producing high-quality products inexpensively and in large quantities is superior to that of other nations. This manufacturing prowess, however, has been based on technology developed in Europe and the U.S.

As the position of Japan rises in international society, it must take actions and fulfill obligations appropriate for the position. Other countries expect this of Japan, and we Japanese are aware of the fact. The environment surrounding Japan, however, is a severe one including a large trade surplus, appreciation of the yen, launching of offensives by NIE's, as well as technological friction that has arisen from these economic issues.

In such an environment, what course should Japan take to meet these obligations?

In the area of scientific technology, in particular, we must stop relying on European and American technologies now that we have reached a position of equivalence with those countries as a member of the industrialized world.

The following trends are evident in Japanese scientific technology.

(1) Emphasis has been placed on technologies for private use.

(2) Private-sector share of research and development is extremely high (80.5 percent in 1985).

(3) Flow of research funds is relatively fixed. There is very little change in who pays research expenses and who uses them.

(4) While the level of industrial technology is extremely high, the level of fundamental research is relatively low.

Taking these characteristics of Japanese scientific technology into consideration, the Japanese government in March, 1986, established the "Scientific Technology Policy Guideline," through which it decided to promote advancements in Japanese scientific technology with the basic principle of promoting the development of innovative scientific technology, the development of scientific technology that is in harmony with society and the development of scientific technology that emphasizes international characteristics.

To realize the promotion of creative and unique scientific technology as stated in this guideline, we must stop implementing other nations' technologies, begin aggressive fundamental research with a broad scope and carry out independent technological development. To this end, the Japanese government has already begun to put efforts into reinforcing fundamental research and an international research structure by establishing the "Creative Scientific Technology Promotion System" and the "International Frontier Research System."

Private businesses that carry out a majority of research and development in Japan are showing signs of putting efforts into research and development from a long-range perspective. They are striving to develop technologies that will be the foundation of their future instead of merely coping with such current economic situations as trade friction and the appreciation of the yen. One indicator of this effort is the increase in research and development expenditures, giving rise to the second boom in the construction of research centers.

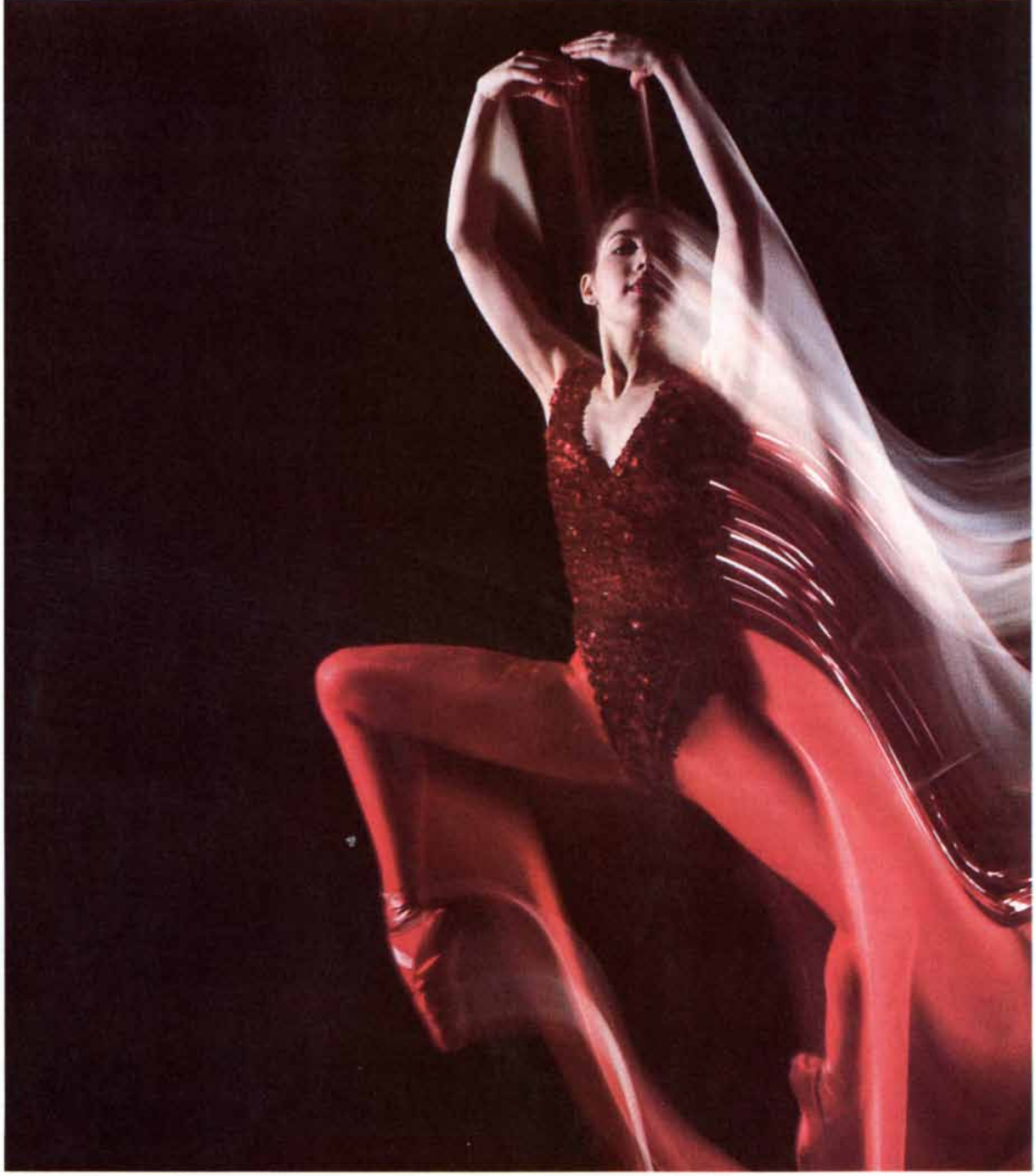
Total research expenditures in 1990 are forecast to be ¥8.1 trillion, or 3.19 percent of the national income of Japan. The total number of researchers this year is expected to reach 406,000. When these figures are compared in the international arena, both total research expenditure and total number of researchers are second only to the U.S. In 1955, total research expenditures in Japan were approximately ¥40 billion, about 13 percent of the total research expenditure of England at the time, minuscule when compared to the expenditures of the U.S., West Germany or France. For the next 30 years, however, Japan's advancement in this area was spectacular.

Of the now enormous amount spent on research, approximately 80 percent is spent by private businesses, an investment which can be credited for the rapid progress of Japanese business. According to the "Survey of Research Activities by Private Businesses in Fiscal Year 1987," prepared by the Science and Technology Agency in May, 1987 to investigate recent trends in business, most businesses increased research and development investments in spite of a decline in sales caused by the appreciation of the yen from fiscal year 1985 to 1986. This increase in research and development investment was most evident in communications, electronics, electrical instruments, electrical machinery and pharmaceutical industries.

As evident from these investments, Japanese businesses feel that aggressive research and development is essential for them to survive. In order to assure a bright future, it is necessary for Japanese businesses to switch from the traditional concept of improving existing technologies to innovating technologies as European and American firms are doing. The concept of placing priority on creative technology is finally being emphasized in Japanese companies, and it is about to become a custom.

Let us now state a few examples to show such a movement in businesses.

Seishi Katayose is managing editor of *Saiensu,* the Japanese-language edition of SCIENTIFIC AMERICAN and Japan's leading science publication.



The difficult it does automatically. The impossible takes a few more seconds.

With the Nikon N8008, things that routinely used to drive photographers up the wall are now matters of mere routine.

Difficult lighting situations, for example.

Nikon's five-segment Matrix Metering can solve problems like severe backlight, harsh contrast and deep shadows, all automatically.

To get a perfectly exposed picture, all you have to do is push the shutter button.

On the other hand, some things aren't that easy. Take something like synchronized, cybernetic, rear-curtain fill-flash (what it took to produce the picture above). To accomplish that, you actually have to do something: like push an extra button or two.

The point is, highly creative pictures that used to be difficult to impossible for anyone less than a professional photographer are now within your grasp.

So look for this symbol, and let a Nikon Advanced Systems Dealer put something else in your grasp. The slightly miraculous N8008.

Nikon.
We take the world's greatest pictures.®

One year membership in the Nikon USA Club is free with every SLR purchase when you submit the Nikon USA Club application. For further information on these products, write: Nikon Consumer Relations Dept. N20, 19601 Hamilton Ave., Torrance, CA 90502-1309. © Nikon Inc. 1988.

The Nikon N8008 and SB-24 Speedlight.

A NEW ERA

by John Hartley

Superconductivity

Ever since Georg Bednorz and Alex Muller of IBM Zurich discovered the high-temperature superconducting compounds based on yttrium and barium, Japanese researchers have been pushing the frontiers forward. Indeed, researchers at several universities and laboratories were working on similar materials before that announcement. When it came to the next range of compounds, based on bismuth, Japanese researchers were equal first.

Owing to their enormous potential, these high-temperature or high T_c materials have excited researchers around the world. Since superconductive materials transmit electric current without resistance, they open opportunities for ultrahigh-speed computers, cheap superconductive magnets and highly efficient distribution of electric power.

Currently in use are superconductive materials in magnets for nuclear magnetic resonance (NMR) and spectroscopy or body scanning. Other products under development include superconductive linear trains and synchrotron rings for advanced lithography and medical work. Further down the road are electric power distribution, and the use of superconductive linear motors in a host of other applications from conveyors in factories to commuter buses and ships.

Long before high-T_c superconductive materials had been discovered, Japanese researchers were busy trying to exploit the advantages of existing superconductive magnets. Toshiba, for example, is among the leaders in NMR for medical work, but some of the other projects are far more dramatic and far-reaching, and would benefit from the new materials. Whether those materials are used or not, these projects demonstrate that innovation in Japan is not new.

Most exciting of these is the magnetic-levitation (mag-lev) train developed by Japan Railways (JR). Plans to build ships powered by superconductive magnets in such a way that propellers are unnecessary are also being developed.

JR's mag-lev project faces competition from a system that does not rely on superconductivity, and which was developed by Japan Air Lines, and now being commercialized by a new company called HSST. HSST is due to start operating a 200 kilometers per hour service in Las Vegas.

JR's system is the most ambitious with a planned cruising speed of 500 kilometers per hour. Looking well into the next century, the researchers at JR see the possibility of a train traveling at 800 kilometers per hour, although it will need to run through a partial vacuum in tunnels to reduce air resistance sufficiently. Pipe dreams, maybe, but without them innovation would be in a sorry state.

Researchers at JR and at the companies building the prototype trains have needed long vision—some of them have been working on the project for 20 years, and have gradually pushed up the speed and reliability of the system to the stage where it is a practical proposition. JR's first mag-lev prototype established a world record of 517 kilometers per hour in 1979, while the latest model, the ML-002 which is a genuine passenger carrier, has already attained 350 kilometers per hour.

The newest of JR's mag-lev vehicles, the MS-002, is a prototype of a real train, with air conditioning, comfortable seats and a lightweight body. "To reduce weight, we developed a new manufacturing technique," explained Morishiga Hattori, senior engineer at Hitachi's Kasado works, where rolling stock is made.

Mag-lev trains will offer the fastest city-center to city-center journey between Tokyo and Osaka. "We can do it in 75 minutes by mag-lev train," said Koumei Miyairi, an engineer involved with the project from Hitachi. "It takes 120–130 minutes by air, and 190 minutes by shinkansen."

Why were superconductive linear motors adopted? This was considered necessary because earthquakes could easily cause local damage to the track.

In addition, the heavy snowfalls in the north of Japan could cause a build-up of snow several centimeters deep. The combination of high lift and a speed of 500 kilometers per hour led to the need for superconductivity.

JR is confident that even with helium refrigeration the mag-lev trains will be cost effective. However, Miyairi said that with liquid nitrogen refrigeration, running costs for the system could be reduced by about 95 percent. Clearly, the long years of research, likely to culminate in practical use this century, will pay off owing to the development of the high T_c materials.

Neural Networks and Artificial Intelligence

Although there have been numerous reports about the development of artificial intelligence (AI) in Japan, it must be admitted that until recently the work would hardly pass off as AI elsewhere. Now that barrier is being broken, NEC is developing neural networks—networks that resemble the cells and connections in the brain, but on a small scale.

All the Japanese computer manufacturers are working on AI, and there is also the government-sponsored fifth-generation computer project, being carried out at an institute specially set up for the project: Institute for New Generation Computer Technology. ICOT developed a version of Prolog for parallel computers, and is now working on a parallel computer with 1,000 processors for logic programming.

Yasuo Kato, director, and head of NEC's computer system research laboratory, says that the budget for the fifth-generation project is too small. "They are concentrating on the kernel for Prolog and extended Prolog. The aim is to stimulate other projects," he said.

Neural networks are quite a different approach from the work at ICOT, and are much closer to artificial intelligence than the combination of a huge knowledge database and some form of inference machine. Researchers in Fujitsu's AI laboratory have been working on neural networks for two

John Hartley, trained as an engineer, is an industrial journalist.

Why JAL?

"I like organizations that pay attention to details. I feel at home."

Akio Morita
Chairman of the Board &
Chief Executive Officer
Sony Corporation
Tokyo, Japan

"I respect an operation when it's managed by true professionals."

Jeffrey A. Barks
Associate Dean for Master's
and Bachelor's Programs
Sloan School of Management
M.I.T.

In a world of increasingly unpredictable service standards, Japan Air Lines remains a reference of professionalism. That's one of the reasons why more people from all over the world choose JAL to fly to the Orient than any other airline. On your next trip to the Orient, fly the airline that's run the way you'd run an airline. For reservations, call your travel agent or Japan Air Lines directly at 1-800-JAL-FONE.

JAL
JAPAN AIR LINES
Comfortably ahead, worldwide.

Tokyo Osaka Seoul Hong Kong Beijing Shanghai Manila Singapore Bangkok and Beyond

years, and are concentrating on relatively simple networks that are taught and corrected when they make mistakes. The researchers are not trying to develop a network that can be placed somewhere, left to its own devices, and will then learn what to do. Practicality is the key, and the first applications are seen as robot controllers and pattern recognition with knowledge-based systems following later.

Hiromu Hayashi, manager of the AI laboratory, feels that the neural network would not be able to act as a knowledge-based system on its own. "The von Neumann computer has so much more power than the neural network, so we expect to develop a hybrid neural von Neumann machine. We don't have a method of combining enough neurons yet to build a computer, though."

The concept of a neural network is that a large number of fairly simple cells are connected together such that they can receive inputs from many other cells. The inputs are weighted to produce a suitable output.

The first practical application will probably be a real time robot-controller which will allow an assembly robot to adapt to the surroundings without being taught a sequence of events as at present.

Trench Cells, New CCD's

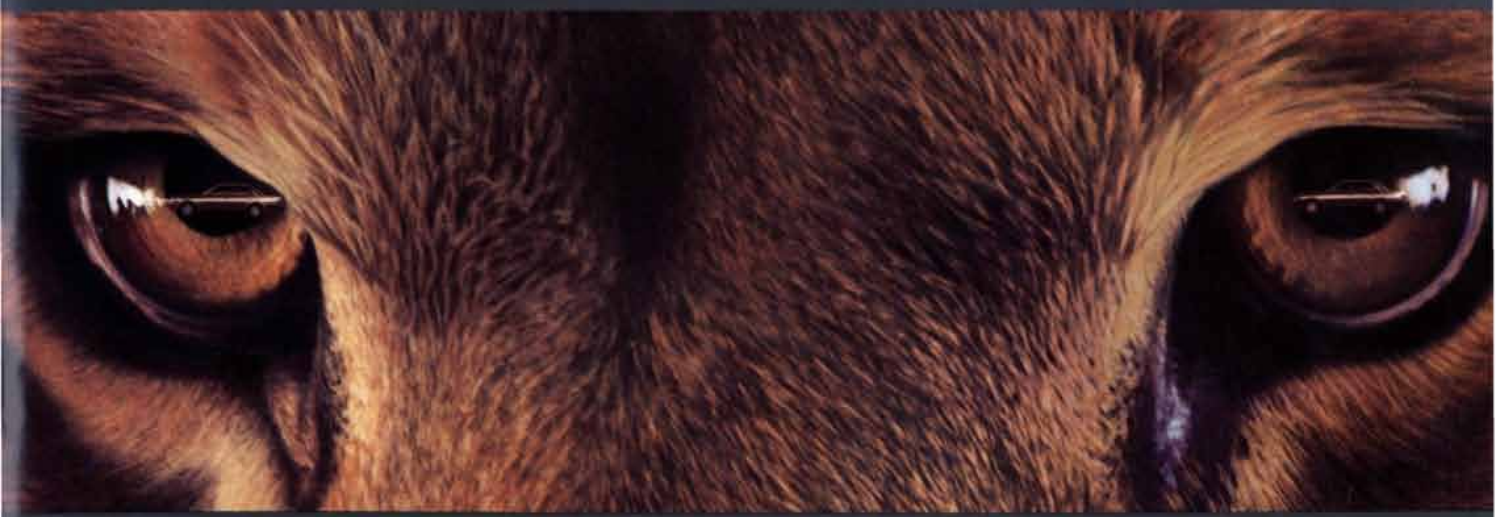
Japanese manufacturers dominate the semiconductor memory business, and are moving aggressively into logic devices, long the preserve of U.S. companies. In addition, they are moving further ahead in charge coupled devices (CCD's) essential elements in video and the forthcoming high-definition television, a Japanese development that could easily become a world standard.

Memory, however, is considered the most important benchmark of technology among Japanese semiconductor makers, not least because it accounts for the largest proportion of business. Already some Japanese companies are distributing 4M-bit DRAM samples to their main customers. These are not preproduction chips, but are intended to verify the basic functions, and give guidance on packaging and other details of production versions, which are still some time away.

Owing to the difficulties involved in designing these chips, innovation is vital. The sources of the innovations have been surprising. Last year, it was NTT, the telephone company, that announced the first 16M-bit DRAM device. This year, Matsushita came up with an innovative cell structure that allows the 4M- and 16M-bit devices to be much smaller.

Masaharu Noyori, manager of admistration and planning at Matsushita's semiconductor research laboratory, says that the company is putting equal effort into the development of memory and logic—microprocessors and controllers.

As memory devices become denser, it is difficult to maintain a small enough pattern to put the chip in a standard package. Standard packages are needed because many companies want to replace existing chips with the latest model without redesigning their complete assemblies. The solution is trench or stacked cells to reduce cell size for the larger chips.



INTRODUCING THE 1989 TOYOTA CRESSIDA... A SEDAN WITH THE HEART OF A LION.

THE PRIDE OF TOYOTA.

Toyota has uncaged an all-new Cressida for 1989. The king of luxury performance sedans has more power and grace than ever before. With a fuel-injected 6-cylinder 24-valve engine, Cressida delivers 190 lionhearted horsepower. Its redesigned chassis, featuring sure-footed MacPherson strut front and double wishbone rear suspension, combines with Cressida's new power to give cat-like handling and roaring performance. Complementing all of this power and performance is the sleek tautness of Cressida's graceful new styling. Cressida is indeed the master of the asphalt jungle.

Inside is the majestic luxury of Cressida's redesigned interior. It is an ergonomically designed environment that emphasizes comfort and control and sets new standards for standard equipment.

Cressida's optional equipment is just as impressive as its array of standard features. Leather trim interior,* compact disc player, and Anti-lock Brake System (A.B.S.) are just some of the options that make Cressida such an extraordinary combination of performance and luxury.

Finally, there is the reassurance of knowing that Toyota Quality has made Cressida the most trouble-free new car sold in America for the past two years.† Cressida is master of the asphalt jungle.

The 1989 Cressida. Make the pride of Toyota the pride of your driveway.

36-month/36,000-mile basic new vehicle limited warranty, including powertrain and electronic components with no deductible and no transfer fee. Normal maintenance and wear and tear items are not covered.

Call 1-800-GO-TOYOTA for more information and the location of your nearest dealer.

Get More From Life... Buckle Up!

*Leather interior not available on model shown in Pewter Pearl.

†Based on problems encountered in first 90 days of ownership—J. D. Power & Associates 1987, 1988 New Car Initial Quality Surveys.

TOYOTA QUALITY

WHO COULD ASK FOR ANYTHING MORE!

Dr. Hisao Oka
Senior Managing Director
Mitsubishi Electric Corporation

A revolution takes place when a social need and an element of scientific technology that satisfies the need meet in a timely manner, such as with the industrial revolution. I think that this is another age of revolution.

In this respect, there are various elements of scientific technology supporting the current wave of globalization. At the foundation of the globalization is the active exchange and utilization of information.

A number of elements of scientific technology have emerged to meet society's orientation toward information. One is electronics. The largest element of scientific technology contributing to globalization is microelectronics, or primarily semiconductors. Use in computers and communications has spread microelectronics throughout society.

Microelectronics is supported by semiconductors. Semiconductor technology has enabled such extremely efficient communication technologies as VLSI and optical communication.

Integration of semiconductors has been increased from one megabyte to two megabytes and up to 16 megabytes. The next issue that we face is making the semiconductor "intelligent." In essence, such intelligent computers as the fifth-generation computer, the recently-introduced neurocomputer, and the "fuzzy" computer, that aid man's thinking process, and not those that must be programmed by man for each task, will be required in the future.

Communications, too, must become increasingly intelligent. The wave of globlization is likely to continue at least into the first half of the 21st century. An essential factor for this process is the development of technologies for incorporating intelligence and functions.

In today's dynamic society, it is difficult to anticipate the kinds of needs that will arise. In order to promptly meet any type of need and then develop a corresponding product, a company must maintain solid fundamental technologies. Of such technologies, new material technology, in my opinion, is extremely important. The second extremely important technology is that of system software.

The third is processing technology, in which the Japanese excel. I think that Japan should take careful measures to maintain its lead in advanced processing technology.

Improvement in intellectual productivity will become extremely important in the future. While the environment and training of engineers and scientists are obvious issues in the improvement of intellectual productivity, providing adequate tools is another. In this respect, computer-aided engineering including CAD/CAM will become extremely important.

All scientific technologies are developed by studying and imitating the human mechanism. In this respect, the next step is to study and imitate human intelligence and the mechanism of man's nervous system. If this study evolves into the development of the final information system or communication system of the information-oriented age, then biotechnology research must first advance.

In the area of superconductivity, too, even if we could find a substance which is superconductive at room temperature, we do not know whether we could find suitable materials or practical applications. Nevertheless, we must explore a wide range of possibilities.

Fundamental research projects we currently have underway include biotechnology research, research on quantum-well devices or optical devices and extremely fundamental research on computer mechanisms to aid in the development of neurocomputers.

In the past, very few creative technologies were developed by the Japanese in such global projects. Japanese were criticized by Europeans and Americans who claimed the Japanese developed none of the fundamental scientific technologies. This is a problem of environment and education in Japan; I do not think that Japanese lack creativity.

We Japanese must carry out creative research in the future. To accomplish this, the social environment and education system must be changed. Although it is likely to take a considerable amount of time, education problems are probably those that most need to be solved.

Dr. Morio Onoe
Executive Vice President & General Manager
Research & Development Center
Ricoh Company, Ltd.

The mainstay products of Ricoh Co., Ltd. are in the field of cameras, watches and precision instruments. Although Ricoh is not strong in single-lens reflex cameras, its compact camera sales have been increasing and Ricoh intends to further exploit this field. We are supplying video equipment under OEM and have an interest in popular types of video devices. Another field is semiconductors. Ricoh is not engaged in general purpose memories, but in the field of ASIC's it is one of the leading makers. As a parts supplier, it is also manufacturing thermal heads and sensors.

However, the main products of Ricoh, among others, are copiers and facsimiles. Ricoh started in the field of sensitized paper, supplying paper that permitted making copies of drawings readily within an office. Therefore when electrophotographic technology appeared, it was only natural for Ricoh to enter into the field, and in regard to facsimiles where Ricoh has the top share of the market, it became engaged in this product naturally as it was a technology related to the transmission of drawings.

When we look back at the history of our company, the themes of research and development were decided as a matter of course. One such theme is "multi-media." Our self-appointed task is to develop instruments and systems that will utilize not only written language but also other various media such as drawings, images and sounds for use in the office.

The other pillar of research at our company is the "man-machine interface" or the "user-friendly interface." Office equipment is used by persons who are not specialists, and therefore the machines must be designed to be used without training and without the necessity of referring to manuals. In addition, such instruments must be usable by people in any country. The technology that Ricoh is aiming at is the development of an instrument that any ordinary individual would be able to use without difficulty.

At present, facsimile machines are a growing market. Although facsimiles were being studied both in America and Europe, it is the Japanese firms that have perfected the facsimile to its stage of practical technology.

Needless to say, Ricoh has expanded abroad and has established support systems in several countries. It is expected that the internationalization of laboratories and staff will accelerate in the future.

Toshio Nakatsubo
Senior Managing Director
Olympus Optical Company, Limited

At Olympus Company, Limited, there are approximately 10 "unit operations" of which only the microscope, camera, endoscope and medical analysis equipment operations are proving to be profitable. Although the others are surviving on the profits of these operations, one of these, the biotechnology operation, is expected to become profitable in the near future. Moreover, research at Olympus is advancing in the development of printers and high-density optical recording systems in the field of information-related equipment.

In furthering such new operations, the Research and Development Department plays a vital role. However, as it is impossible to ask a researcher who has been developing microscopes to suddenly switch to biotechnology research, Olympus has had no choice but to recruit experienced personnel, assign Olympus research and development personnel to other research institutions and universities for training and merge with other firms to quickly acquire new technologies.

In these circumstances, we are currently focusing our attention on the research environment. Although it was not unheard of previously to conduct biotechnology research in a corner of a hardware factory, such a situation would be inconceivable today. This is because we now place emphasis on providing the guidance that allows researchers to do their most creative work.

No one knows, when assigning a particular individual to a research center, if the best choice is made. Some people are better suited to sales or management than research, and vice versa. In the past, however, many researchers had to become managers after they reached a certain age. Therefore, we recently implemented a "specialist system" under which technicians are thoroughly trained to become specialists. At the age of 40, technicians who show a greater potential for competency in areas other than research are transferred. This system has two merits. The first one is in personnel, where the shortage of directors is eased. In the second, the ability of specialized technicians is improved. We want to develop specialists who lead not only the company, but the academic society and the world as well, because we recognize that the future development of leading-edge technologies can be accomplished only by such individuals.

It is said that the strength of Japan lies in its smooth handling of retraining. I feel that this belief is no longer valid today. Specifically, the traditional method of older employees training their juniors is no longer effective. Based on this idea, we gave to employees of around 40 years of age the title of "group leader" and gave them actual authority over research and development. We also place with them the responsibility of training subordinates.

I feel that the role of the individual in charge of research and development will become increasingly important. He must not be content with the status quo, as competitors may gain an advantage, and he must secure patents in new fields of research. Appropriate research subjects must be pinpointed, a task with which all Japanese firms involved with high technology are surely struggling. In this respect, Japanese firms may have approached the position of American firms of 20 years ago. We have no choice but to work effectively under such circumstances.

Hitoshi Aoike
Managing Director
JVC

We have been involved in audio-video since 1927. When we consider the current environment, we find that we are now in the so-called advanced information society, and our problem is to determine how to expand audio-video in this society. If we assume that audio is one form of information, it is incomplete in the sense of "communication" as the word is used today. Communication only occurs when information is mutually exchanged between two parties, and what we have provided up until now has been one-way traffic. We therefore believe that in the field of audio-video, the one-way traffic era we have seen thus far will become an era of two-way audio-video traffic. We consider this to be the basis of our future vision. The position of technology in such an era, by our thinking, will be to elevate current audio-video to a standard where the "communication" flows in both directions. Communication in two directions previously was the mutual exchange of data in the form of characters and numbers. The role of sound and images has been steadily gaining importance in such communication.

It is our belief that there may come a time when we will be performing the same functions as other companies and, in that case, we will be forced into a competition that is different from what we have been experiencing. We surmise that under those circumstances it will become a matter of determining how to survive. When we view the current environment, we can no longer devote ourselves to manufacturing products based on the hardware-oriented concept as we have been doing, because we will encounter competition from participants from other fields. It is no longer a period when business can be carried on by merely designing and producing quality products. We feel that it is of vital importance to create products that will benefit users and that are acceptable to users even though prices may be higher. I believe we can express our target direction for both industrial and household markets to move from hardware to software.

Internationalization up until the present time has been internationalization of production only where manufacturing is involved. We feel that second-generation internationalization should not be limited to production, but should involve the design of products as well as the research and development that are the basis for such products. The "concept" is extremely important in research and development, and the uniqueness of a concept is a necessity. Concepts originating among Japanese and from discussions among Japanese can only produce opinions and views that are similar. In this respect, the effective utilization of a foreign research engineer can produce different concepts. The second benefit is in the area of new applications, and unlike the inclination up until now of making products from the orientation of available technology, the new approach will be to view products from the perspective of application, as we find that the culture of a particular country greatly influences product application. A person raised in Japan will not be able to approach the American type of application nor the European type of application in the true sense regardless of the effort that he may make. It is therefore necessary to go to a particular country to know the market of that country. In this respect, establishing research in foreign countries will have a significant impact.

SUPER VHS FROM JVC.

ACCEPT NO LIMITATIONS.

Begin with unsurpassed picture clarity.
Add every imaginable feature and function.
Top it off with the JVC name, the company that invented VHS and Super VHS.
Now you know why many experts consider these Super VHS decks with over 400 lines of resolution the most advanced machines you can own.
Consider the HR-S8000U, JVC's finest video achievement. Here you'll find an array of digital functions including Freeze, Strobe, Solarization, Search and Picture-in-Picture.
And JVC's LCD remote puts each of these features in the palm of your hand.
And with the HR-S5000U, JVC puts the superb picture quality of Super VHS within everyone's reach. This new Super VHS VCR has a bold, sleek design complete with a full range of advanced features.
Choose Super VHS from JVC. Now there's no limit to how good home video can be.

JVC®

WE'VE SEEN THE FUTURE. AND IT'S JVC.

Dr. Yasuo Nakajima
General Manager
Central Engineering Laboratories
Nissan Motor Co., Ltd.

Automotive technology is supported by numerous industries and is a congregation of various technologies. Therefore, the automobile industry of the future is one in which high technology is systemized.

An automobile is composed of such technologies as electronics, new materials, information and human engineering. These technologies must be systemized for the automobile to improve. For this to happen, it is necessary to make automobiles more intelligent, enhance their function and improve their reliability. In the past, for example, the computer controlled only the engine, but today it controls both the engine and the transmission.

Whereas automobiles first used eight-bit microcomputers, today they use 16-bit computers and are expected to use 32-bit computers in the near future. The 16-bit computer was developed jointly with Intel over a 10-year period. It is very rare for a Japanese company to conduct a joint development for such a long time with a foreign company.

Although electronics have played a major role in providing convenience in various controls, we can't ignore that mechanical devices also play a major role. Materials, used for function or structure, will support this improvement effort. An example is the ceramic turbocharger. Our company is the only one to offer this in production automobiles.

In March of last year, we established the Semiconductor Research Center. Automotive LSI's are used in a severe environment. We have a difficult time finding an ordinary electronics manufacturer who can develop an LSI that is resistant to such adverse conditions. The research center only conducts development, production is assigned to specialized manufacturers.

One of what we call the "four pillars of technology" is continuous reinforcement of research and development. A facet of this reinforcement is an investment in fundamental research. The second pillar is the development of a support system for research and development. Since the automobile is extremely complex, we must rely on such support systems for calculations, and not rely solely on human judgments. The next pillar is the promotion of the internationalization of our research and development. Last January we founded a company called Nissan Research and Development (NRD Inc.). The company has already begun to conduct development tasks in the U.S.

In May of this year we founded a company called NETC, Ltd. (Nissan European Technology Center, Limited). This company holds similar objectives to those of NRD. We are making our research results public. There have been a lot of discussions concerning the issue of employing foreign researchers. We plan to take positive steps in this area too.

We have been increasing our research expenses by a substantial margin. Our total research and development budget is approximately 4.5 percent of sales.

We have also increased the number of personnel by a substantial margin, and we are scouting experienced personnel so that we can have a staff of mixed backgrounds.

I think that the automobile of the future will have to be made so that the driver does not get tired from driving it. I think that a car that is comfortable and fun, and accurately conveys messages from the outside will be necessary. What is most important, however, is safety.

Hiroshi Tanaka
Managing Director and Chief Executive,
Business Machine Operations
Canon Incorporated

With innovations in copiers, cameras and laser printing, Canon has established itself a deserved reputation as a leader in innovation. Now it is moving on to greater things with the introduction of a color laser copier with editing capability and bubble-jet printing. Like the laser-beam printer, both are opening up new markets.

The importance of Canon's laser-beam printers cannot be overstated. Not only did the introduction of these desktop machines bring superb quality into the office, but it opened the way to desktop publishing, a business that is growing amazingly rapidly at present. Without laser-beam printers there would be no desktop publishing, and without Canon, which has 80 percent of the market, there would hardly be laser-beam printers.

Such is the technology involved in the production of laser-beam printers that only a few manufacturers have entered the business, the bulk of the others relying on Canon and arch rival Ricoh for OEM supplies.

Having mastered that technology, Canon has now moved on to the color laser copier, the result of innovative thinking in design and in marketing. The result is stunning quality and capability that allows editors and copywriters to be creative in double quick time.

Canon was the first into the market with a color laser copier, and is still the only company that has adopted digital processing for the colors. We started work on a color copier with analog processing 12 years ago, but it was not good enough, so we abandoned that project and started on digital processing.

Digital processing is used therefore for the data that is sent to the laser. This makes it easier to get good color combinations than with analog processing, which Kodak's color laser copier uses.

As it is, the copier has a resolution of 400 dots per inch and produces excellent copies. Tanaka said that the image is now approaching that of silver halide.

Not content with that innovative work, Canon has been busy perfecting the bubble-jet printer, which should not be confused with other ink jet printers. The bubble-jet technology was invented by Canon, and it has been selected as an important technology of the future by the Society of Photographic Science Engineering in the U.S.

The bubble jet is ideal for such large copying because the head that discharges the ink can travel across a rail in front of the paper, and so irrespective of the size of the paper, the resolution remains constant, and the cost of the mechanism itself is hardly affected by the size of the paper.

In spite of the innovative concept of the bubble-jet printer, it did not do well at first. Special ink is required, and those early models required special paper; on plain paper, the edges of the blobs blurred. The early machines established a bad reputation but now you can use plain paper, and the latest machines are very good.

Showing that innovation alone is not enough, up until July, 1987, bubble-printer development had not been concentrated, but had been spread across the company. In July, 1987, we established a bubble-jet project team; before that they were not properly organized. We are now confident that the product is good, and that it will improve progressively.

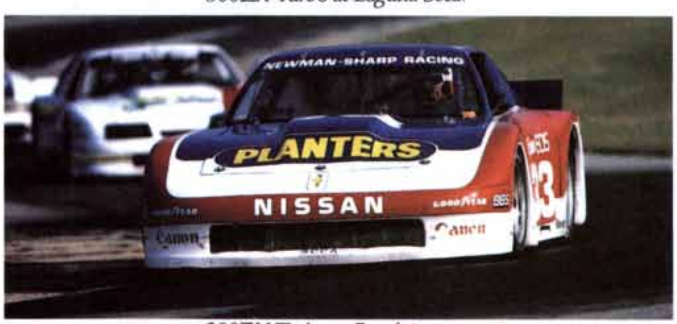
To learn how to build a better car, we took a lot of demanding courses.

GTP ZX-Turbo at Lime Rock Park.

300ZX at Road Atlanta.

300ZX-Turbo at Laguna Seca.

GTP ZX-Turbo at Mid Ohio Sports Car Course.



300ZX-Turbo at Road America.

The toughest race of all.

For years, racing and winning have been important priorities at Nissan®.

We've competed in thousands of races. Taken 56 SCCA National Championships, more than any other manufacturer. And sponsored drivers, racing teams and events all across the country. In all classes and types of racing.

But racing means more to us than just a nice trophy collection. What we find out on the racetrack helps us improve the other cars we build. Cars that have to handle even tougher situations. Like rush hour traffic. Or slippery school crossings.

And for as long as we continue learning, we'll definitely continue racing.

Because no matter how much you already know, we think you can never take enough classes.

NISSAN®

Built for the Human Race.™

Ichiro Yoshiyama
Senior Executive Director
Minolta Camera Co., Ltd.

A look at the current situation in industry shows that a change in industrial pattern is inevitable. This is a widely-held view. Man, having developed the use of tools, eventually built up industries utilizing enormous technology, and has gained the capability of supplying in abundance whatever society demands. As a result, in both quality and quantity, the industrial world has acquired awesome potential capacities.

In the past there was a limit to the reserve capacities of a manufacturer and therefore when a product was originated, it was not possible for the originator to completely dominate the industrial sector or the economy. It was not possible for one company or one country to meet and satisfy the entire demand. However, there came a time when an original product could be developed into a big success. There came a time when all of a sudden, a company was capable of supplying an entire country's demand. This completely changed the nature of industry.

On the other hand, the world has undergone other rapid changes within a short time. With these changes, it became necessary for industry to revise the goals and methods of product conception from that employed when quantitative production power was lacking. In short, the meaning of the coming change in the industrial paradigm is that if the former operational norm continued, we would be irretrievably lost.

In the past there was a period when military demands took the initiative, but now the mass demand for civilian consumer products is necessary for technical progress to proceed economically, yielding profits to support investment in research and development. In order to generate collateral funds for investment in research and development, technical advancement must continue ceaselessly. Once technical achievement subsides, the company will become a loser.

However, when advancement is too rapid, a product that has been developed with much effort will become outmoded within two or three years. After a technology brings forth a product, within two or three years an adequate volume is attained to satisfy all markets. Industry cannot survive unless it is able to produce what the consumers demand and buy. It is because new products are developed that former products become antiquated. This cycle is repeated.

In order to produce a product with revolutionary performance, the design method and production method must change completely in only half the time as before. However, as a product life of 10 years becomes five and then only three years, we have already commenced development of our next generation of products.

I feel that an important requirement for an R&D leader who is working diligently in such an environment is not to interfere with the energy that may burst forth from the young engineers.

Man by nature has an active personality. How is it possible to foster something not yet possessed? It would be something we would not have thought about and therefore we would not be able to make any assumptions or give any definitions. For an engineer engaged in development work, I feel that it is extremely important that they preciously harbor this feeling every day.

Dr. Hiroshi Watanabe
Executive Vice President
Hitachi, Limited

Probably the largest change we currently face is the problem of the appreciation of the yen against the U.S. dollar. While all Japanese companies have made efforts to maintain export competitiveness by reducing costs through implementation of new production technology, etc., there are only so many costs that can be reduced because of the rapid appreciation of the yen. We have nevertheless put our utmost efforts into this area and have managed to reduce some costs, and as a result, some of our products are still competitive in the export market.

One of the countermeasures we took against the appreciation of the yen was to transfer production overseas, and another was to develop new products with high added value which no one else could develop.

In reducing costs, we implemented an all-out cost-cutting research effort with the participation of every researcher in our firm. As a result, we managed to cut costs by 30 percent. The goals for such research, in addition to reducing product price by 1/2, are to reduce product volume by 1/3 and weight by 1/3. Although the goals have to be compromised somewhat in the end, the significance of such an activity is that every employee participates. This kind of participation, I feel, is the strength of Japan and Hitachi.

There are always 30 to 40 special research areas for technical development in the company. Of these, 1/2 are in new product development, 1/4 in cost reduction, and 1/4 in the development of fundamental technology. We have eight research centers operated directly by the company and others affiliated with divisions and factories. These research centers spent 261.9 billion yen for development last fiscal year.

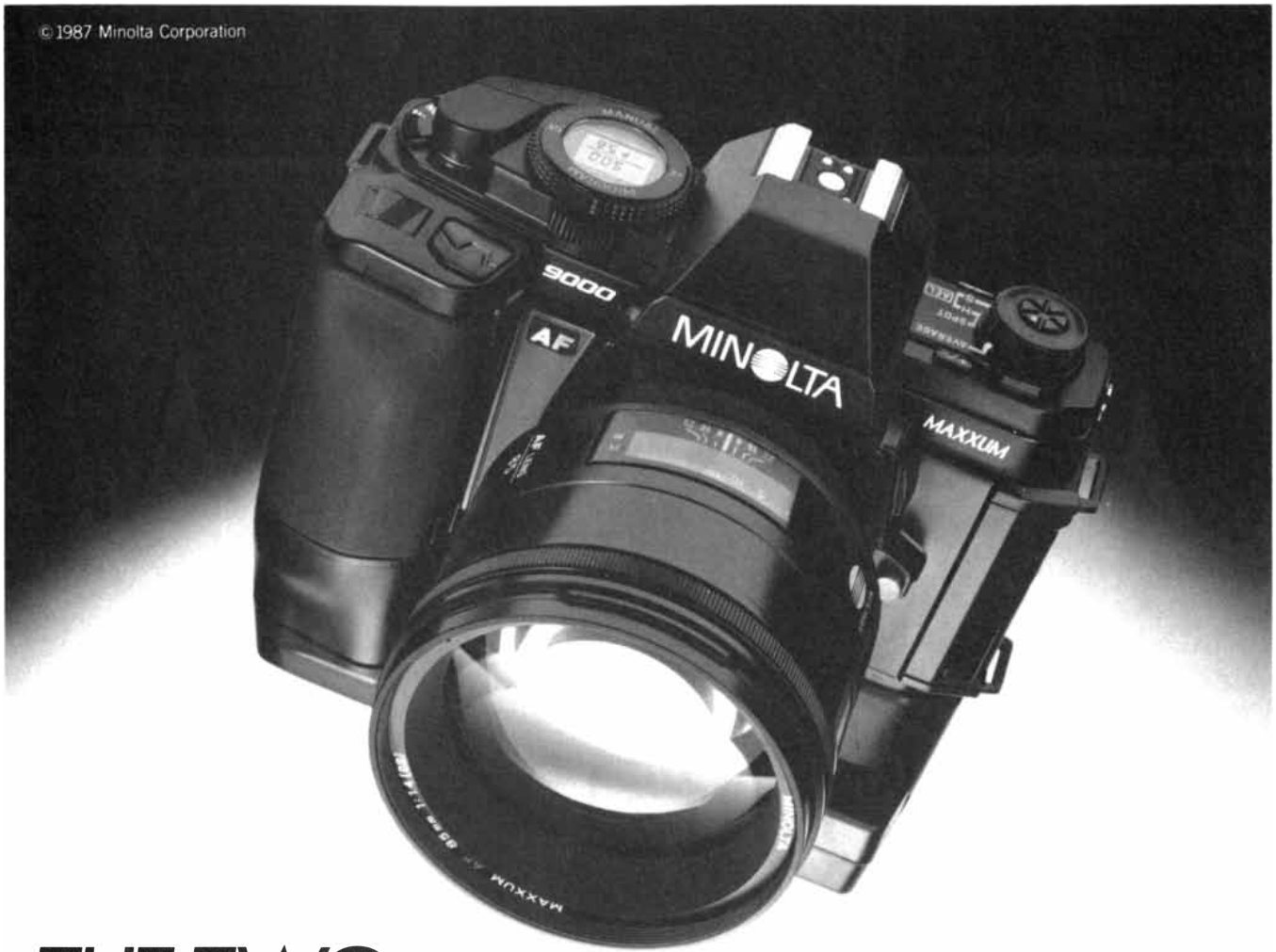
The research centers are aggressively pursuing research that shows future potential. For example, if a new field such as optical electronics is expected to be of importance in the future, a project team is formed by the relevant factory, the Communication Systems Division, and the relevant research center. Led by a manager, this project team then combines all our efforts in working toward the goal.

Research subjects are very carefully selected and divided into those relevant to Hitachi's current problems, those expected to bear fruit in two to three years, and those expected to do so in five to 10 years. Almost 12,000 researchers are grouped in various ways into project teams. In general, the long-term projects are carried out by the research centers and the short-term ones by factories and divisions.

Of our 710 new employees, 170 or 180 were assigned to research centers and the rest to factories where they are making contributions as technicians.

We feel that it is vital to establish research centers overseas, and we are studying ways to accomplish this. I believe there are two ways. One is to conduct research overseas only, to utilize superior talent and technologies. The other is to expand an overseas operation, including production, and assign supporting research and development teams. We are now using the method of assigning technical teams as required by developments in foreign operations.

In the future, advancements in electronics will be supplemented by biotechnology, new materials and software development.



THE TWO MOST DYNAMIC FORCES IN PHOTOGRAPHY

Your flair, your talent, your creativity and the technological sophistication of the Professional Maxxum 9000. An unsurpassable combination. But if the potential unleashed by this camera's ground-breaking continuous autofocus is readily apparent, consider the opportunities unlocked by its many other special features. By selecting Maxxum's center-weighted metering or shifting to spot metering with highlight or shadow biasing, you can use light and color like a fine painter to add drama and sensitivity. With Automatic Multi-Program Selection to put you in the program that matches your focal length, even while you zoom, you need never be distracted from the subject before you. And with Professional Maxxum's 1/4000 sec. shutter speed and 1/250 sec. flash sync you can freeze the subject instantaneously. Other important considerations are options like auto bracketing, 5fps motordrive and a 100 exposure filmback, and the 26 lenses, 4 dedicated flash units, and 2 data backs that comprise the most complete autofocus system anywhere. With the force of Minolta technology in support of your vision, there's virtually nothing you can't do.

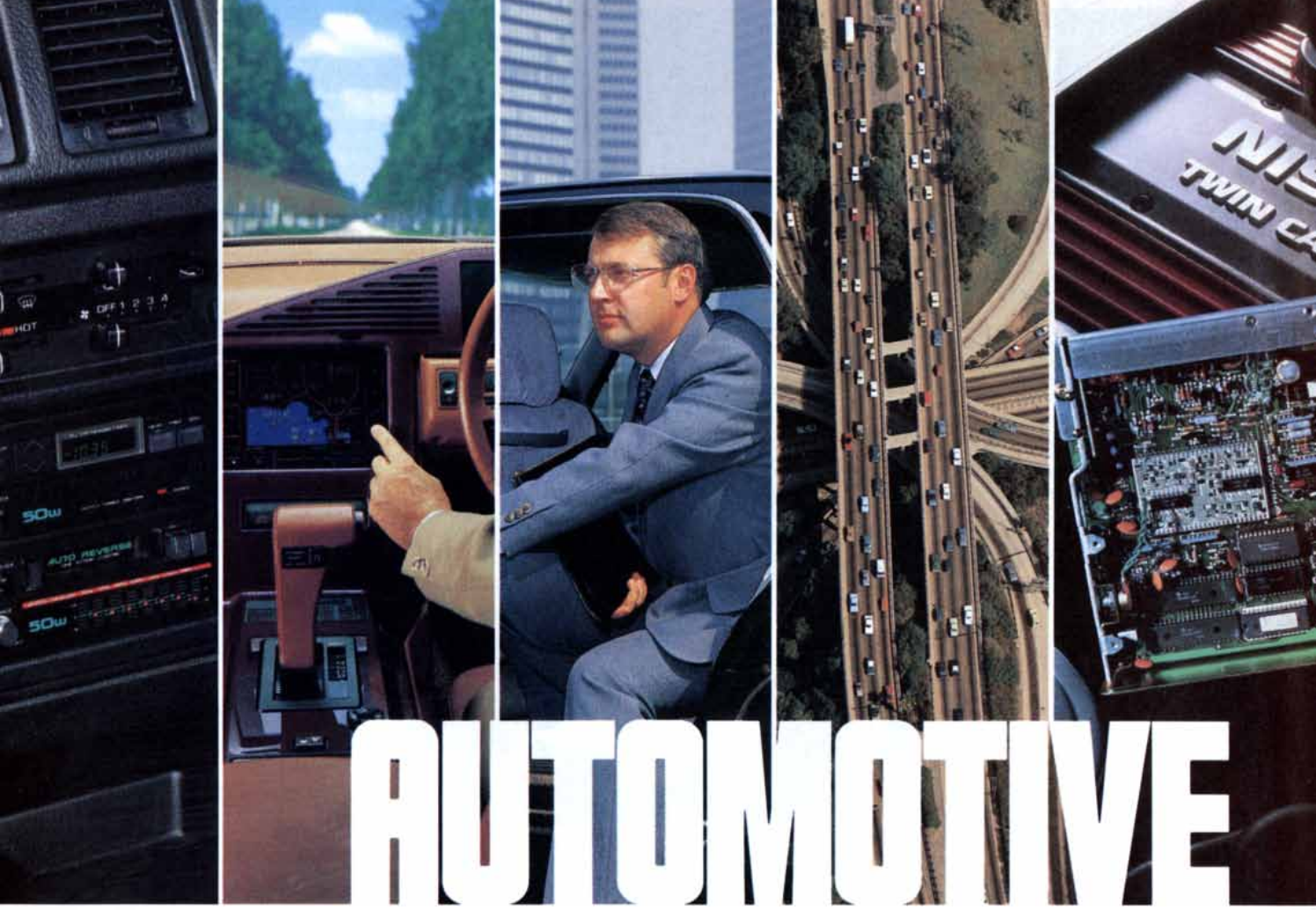
YOU AND THE MIND OF MINOLTA

Motordrive and AF85mm lens optional. Be certain the valuable Minolta USA 2-yr. camera/5 yr. lens ltd. warranty is in your package. For more information see your Minolta dealer or write Minolta Corp., 101 Williams Dr., Ramsey, N.J. 07446. In Canada: Minolta Canada, Inc. Ontario.

MAXXUM®
AUTOFOCUS SYSTEM

ONLY FROM THE MIND OF MINOLTA

MINOLTA



Hitachi's wide-ranging automotive technologies include car audio, the Satellite Drive Information System featured on Nissan's CUE-X concept car and a microcomputer engine control system.

Automobiles should be more than safe, comfortable machines. They should also be able to communicate with the world around them.

Recent advances in car electronics technology have been remarkable. They've not only improved basic functions such as engine control, they're now being seen in man-machine interfaces providing more comfort and operating ease, and even in communications with the surrounding world. Down the road there are things even more exciting.

Hitachi's scientists and engineers are at work on a Multi Information System using a color thin film transistor LCD to display operating information, road maps and a navigational system using these maps. With this system a driver could obtain a variety of driving information simply by touching the display screen. Eventually, he'll be able to issue verbal commands to, for instance, regulate the temperature within his car. Hitachi electronics and semiconductor technology can also bring free communication with the outside and determine a car's exact location through use of Global Positioning System satellites.

Hitachi has also developed a highly acclaimed hot wire air flow sensor used in engine management. It helps achieve the diametrically opposed goals of maximum power and fuel economy. And we've created many other superior products for driving control, suspension control, air conditioning and audio.

We link technology to human needs; and believe that our special knowledge will create new, highly sophisticated functions that are also easy to operate. Our goal in automotive electronics — and medicine, energy and consumer electronics as well — is to create and put into practice innovations that will improve the quality of life the world around.

HITACHI

Hitachi, Ltd. Tokyo, Japan

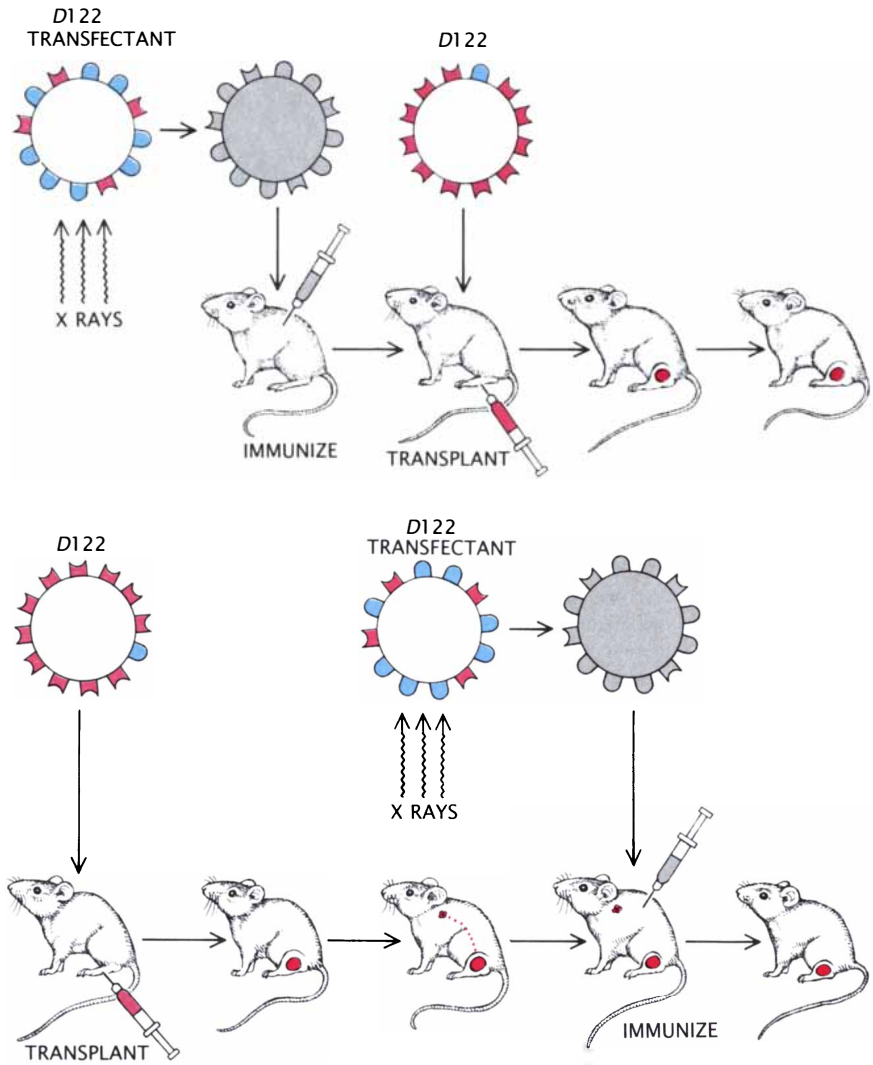
Indeed, from the correlations we observed in these experiments we can assume that it is the *fos* gene product that controls the expression of MHC genes in normal cells.

The connections we discovered between the *fos* oncogene, MHC-gene expression and metastatic competence broadened our understanding of the interactions between metastatic cells and the immune system, but we had begun screening for oncogenes with the intention of finding properties of metastatic cells that were *not* related to their immunogenicity. And when we continued the screening, that is what we found. We stumbled on an as yet unidentified gene that is preferentially expressed in metastatic clones of both the 3LL carcinoma and the T10 sarcoma. The probe we were using was supposed to detect an oncogene called *fms,* which encodes a cell-surface receptor for a growth factor that spurs blood-cell maturation. The gene we found instead gives rise to a product larger than but similar to the product of the *fms* oncogene. Could the unidentified gene also code for a growth-factor receptor?

We have examined the membranes of highly metastatic cells, and we find that they do indeed exhibit the biochemical properties typical of cell-surface growth-factor receptors. Furthermore, we discovered recently that metastatic cells from other mouse tumors that give rise to lung metastases also express a gene resembling the *fms* oncogene. A graduate student in our laboratory, John Gubbay, is currently cloning the gene and attempting initial characterizations. Our guess is that the gene codes for a receptor that recognizes some factor specific for growth in lung tissue.

The idea is appealing because it rounds out our list of comparisons between metastatic and nonmetastatic tumor cells. Our research with enzymes suggests that most tumor cells are capable of invading the intercellular matrix and penetrating blood-vessel walls. Our work with MHC molecules shows that some cells survive the journey through the bloodstream better than others because they can evade the attack of killer *T* cells. Now we may have hit on a reason some cells are better than others at forming new tumors when they reach their target site: they might carry a receptor for a growth factor that prompts their proliferation.

We are anxious to learn how much of what we have discovered can be applied to human cancers, and we are currently collaborating with investigators at the Memorial Sloan-Kettering Cancer Center in New York to try to find out. Our first step is to investigate to what extent metastatic competence in human tumors is also determined by ratios of major histocompatibility types. If the principles we have established for mouse tumors apply to human malignancy as well, we shall have helped to pave the way for eventual immunotherapy of human cancers. In the meantime we plan to continue our basic research efforts with mice, in the hunt for other determinants of tumor metastasis.



IMMUNIZATION AGAINST METASTASIS can be accomplished with transfected *D*122 cells. In one experiment the authors inoculated healthy mice with transfected cells that had been inactivated with X rays so that they could not form tumors. The authors then transplanted metastatic *D*122 cells into the inoculated mice (*top*). The animals developed primary tumors but no lung metastases. In another experiment mice in which primary and secondary tumors had already been induced were injected with the inactivated *D*122 transfectant. The metastases diminished. Apparently transfected *D*122 cells provoke an immune response powerful enough to defeat highly metastatic cells and secondary tumors. The response, however, does not seem to be strong enough to destroy the primary growth of a metastatic clone.

FURTHER READING

METASTASIS RESULTS FROM PREEXISTING VARIANT CELLS WITHIN A MALIGNANT TUMOR. Isaiah J. Fidler and Margaret L. Kripke in *Science,* Vol. 197, No. 4306, pages 893-895; August 26, 1977.

REJECTION OF TRANSPLANTABLE AKR LEUKEMIA CELLS FOLLOWING MHC DNA-MEDIATED CELL TRANSFORMATION. Kam Hui, Frank Grosveld and Hilliard Festenstein in *Nature,* Vol. 311, No. 5988, pages 750-752; October 25, 1984.

MHC GENES AND ONCOGENES CONTROLLING THE METASTATIC PHENOTYPE OF TUMOR CELLS. Lea Eisenbach, Gil Kushtai, Daniel Plaksin and Michael Feldman in *Cancer Reviews,* Vol. 5, pages 1-18; 1986.

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.

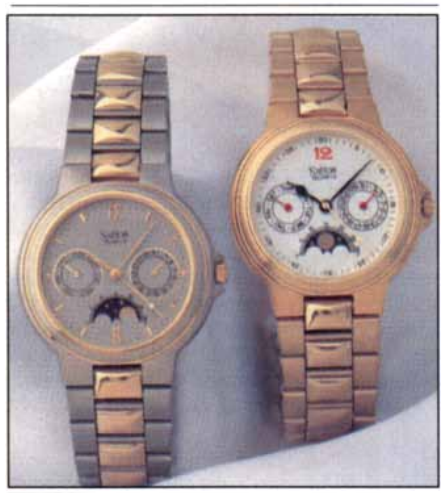
CIVILIZED AWAKENING

NEW from Nakamichi! - a civilized way to wake up. This Nakamichi Clock Radio gradually turns up the volume on your day, allowing you to skip the usual shock of your radio alarm. Created in the true Nakamichi tradition, this clock radio is a high fidelity AM/FM receiver. Paired with its companion unit, the Nakamichi creates a complete stereo system. The clock radio has a digital quartz synthesized, drift-free tuner with a built-in FM antenna. High quality 5″ speaker is encased in acoustically engineered enclosures for superior sound. Two independent, programmable alarms and a six minute snooze bar. It has separate bass/treble control, automatic loudness compensation to enrich tone quality for low listening and 8 station presets. The LCD display adjusts brightness automatically. Volume gently eases up to a preset level over 10 seconds. Plus a 60 minute sleep control and backup battery. The matching companion unit ties into the main unit through an 18′ cord. When used in combination with the main unit, the companion provides another superbly designed 5″ speaker for extraordinary stereo sound, a second clock/alarm display that can be switched off and an additional snooze button. This state-of-the-art clock radio system is a musical masterpiece. Main unit **$139.95** in black #2400 or white #2410; complete stereo system (main unit with companion) **$229.95** in black #2420 or white #2430.

SHARP AS A RAZOR

The Chef'sChoice® Knife Sharpener safely and easily hones the dullest of knives to razor sharpness in minutes. According to *Food and Wine*, the Chef'sChoice Knife Sharpener is "unquestionably the best." *Popular Science* calls it "idiot proof." Its ingenious patented design is the brainchild of former DuPont engineer, Daniel Friel. His unique sharpening method uses diamond abrasives to sharpen a knife in three stages to form a progressively sharper edge. Since magnets hold the knife at the precise angle required for each operation, the guesswork of traditional methods is eliminated. The resulting triple bevel edge stays sharper longer than conventional edges. Only the minimum amount of steel is removed during each stage, making it safe for all high quality steel or alloy knives. World class chef, Pierre Franey, hailed the Chef'sChoice as "unparalleled, in my experience, for putting an exceedingly sharp and long-lasting edge on knives of all kinds." Manufacturer's two year limited warranty, UL listed, measures 9″ x 3½″ x 4″. **$79.95** #2450.

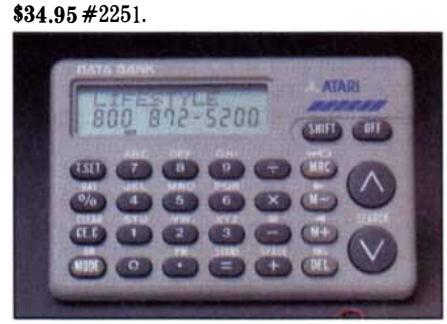
CALENDAR ON YOUR WRIST



While poets pondered the mysteries of the moon, the ancient mathematicians and astronomers were able to find the key to timekeeping in the moon's movements. Sutton's moonphase watch brings the irresistible appeal of the moon to the precise art of modern timekeeping. This unique lunar phase chronograph, with sweep second hand, is a virtual calendar-on-a-wrist. It offers precision quartz accuracy, and features twin inset registers that indicate both day and date at a glance. A starry dark blue third dial is synchronized with the moon's 29.5 day cycle to reflect its exact phase. Water resistant to 100 feet. Available in two styles, both with an adjustable wristband. Manufacturer's 1-year limited warranty. **$99.95** each. Stainless steel/gold-plated # 2360 or gold-plated #2350.

POCKET-SIZED MEMORY

This is the most innovative, easy-to-use, multi-function Data Bank in its class. No more fumbling through notes or phone books, with the touch of a key the information you need is at your fingertips. The Atari Data Bank holds up to 50 names and numbers and 50 memos in its 2K memory. It has raised rubber keys for easy data entry, a built in clock with date display, an 8 function calculator, and a secret function to safeguard confidential data, all displayed on a 2 line 24 character display. The Data Bank offers a scroll feature for access to information, all in alphabetical order, or direct search for instant retrieval. This Data Bank is so smart it will even turn itself off if you forget. Atari's multi-function Data Bank is the perfect solution for complete business and/or personal organization. **$34.95** #2251.



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.

THE LIFESTYLE RESOURCE™

ALLERGY AND COLD BUSTER

Now a major scientific breakthrough — The Viralizer® System — gives you relief from cold, sinus and allergy symptoms. It's the newest development of a concept pioneered at the Pasteur Institute in Paris. The cause of the common cold is the Rhinovirus family which lives and multiplies in the nose and throat, but cannot thrive in temperatures over 110°F. The Viralizer is designed to deliver a gentle, controlled heat which penetrates the nose and throat. After a pleasant heat treatment, the Viralizer dispenses either of two mild, over-the-counter, medicated sprays. Vira-Spray I is an analgesic, anti-bacterial spray. Vira-Spray II is a decongestant that provides temporary relief of nasal congestion due to colds, hayfever, sinusitis or allergies. These therapeutic sprays further discourage the stubborn cold and sinusitis germs so you're less likely to be re-infected or spread your cold to others. The Viralizer can produce effective relief by using it for only 3 or 4 minutes, several times a day. Proven in clinical tests 90% effective in treating the symptoms of upper respiratory infection in 24 hours or less, the Viralizer works without pills. Viralizer, is safe for children and adults, has been tested and recommended by doctors. The complete Viralizer ® System includes 1 electric Viralizer with Vira-Spray I and Vira-Spray II plus a 3-pak refill of medicated sprays. **$39.95** #1690.

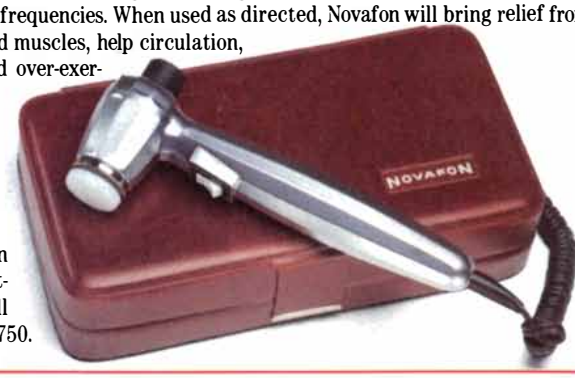
WHEN YOU'RE IN A HURRY

When you're running short of time, ask one of our phone operators about UPS Blue Label delivery. In most cases, we can get your shipment to you in three or four working days.

CALL TOLL-FREE 800-872-5200

SOOTHING SOUND MASSAGE

Tested more than 30 years, the Novafon sound massager is used in clinics and health centers around the world. 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 plush case, perfect for carry along. 1 year warranty. A precision made instrument with no interacting parts to wear out, the unit will give years of service. **$169.95** #1750.



THE SIX INCH, FIVE POUND FURNACE

Ceramic heaters are unique because of their size, exceptional safety, spartan use of electricity and tremendous heating ability. Two North American companies, GTE and Micromar, have come together to create what *The Lifestyle Resource* considers to be the best and safest 1500 watt ceramic heater on the market today — and we looked at 12 competitive heaters. Of all those heaters, *we chose Micromar's Heat Machine Plus to offer to you.* GTE spent 18 months of intensive research and development creating, to our knowledge, the *only* ceramic element specifically designed for a portable ceramic heater. This U.S. patented element is manufactured exclusively for Micromar who was the first company to market portable ceramic heaters in North America. They've built on their experience to design this third generation ceramic heater using the finest components they could find. Let's look at some of the unique features of this remarkable product. SAFETY: This ceramic element operates at a temperature below the ignition point of tissue paper. No fuels, odors, flames or red hot coils — safety grille gives added protection. A "tip-over" switch automatically cycles the heater down if it is turned over. It is designed to run around the clock, is exceptionally safe and great for a child's room. AIRFLOW: The high quality fan is manufactured in West Germany by Papst, who also supplies fans to other top notch companies. The ceramic elements are manufactured using a special pressing process which allows larger holes than any other solid ceramic element — creating an airflow almost twice as much as some others. That means quicker, more even heat distribution. EFFICIENCY: At full capacity of 1500 watts, The Heat Machine Plus puts out 5120 BTU's per hour. As it reaches the desired temperature you set, electrical consumption drops, thereby reducing your electrical costs. Temperature is controlled by a special thermistor circuit which varies fan speed and eliminates the on-off, hot-cold cycling common to many heaters. VERSATILITY: The Heat Machine Plus goes anywhere, gives instant heat and allows you to lower the heat in other areas of the house and selectively heat those areas you occupy. DURABILITY: The Micromar is built like a tank, with a heavy-gauge metal case, industrial grade components and a meticulous attention to detail. It was laboratory tested for over 100,000 on/off cycles and carries a manufacturer's 5-year limited warranty. While not the least expensive of all the brands we looked at, we believe Micromar's Heat Machine Plus to be the best value among this year's ceramic heaters. Built-in handle and washable air filter. Standard 110 volts AC. UL listed. **$179.95** #2280.

THE LIFESTYLE RESOURCE™

CREDIT CARD ORDERS CALL TOLL-FREE 24 HOURS **800-872-5200**

DEPT. SFAL08; 921 EASTWIND DR. SUITE 114: WESTERVILLE, OH 43081

SEND TO (PLEASE PRINT)

ADDRESS

CITY STATE ZIP

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

ACCT. # EXP. DATE

ITEM NO.	QTY.	DESCRIPTION	ITEM PRICE	TOTAL PRICE

Shipping Charge covers UPS. handling and insurance for guaranteed delivery.

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

UPS Second Day available for an additional $7.50 per order.

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

©1988 C80 ADVERTISEMENT

The Electrification of Thunderstorms

Although it has been known for two centuries that lightning is a form of electricity, the exact microphysical processes responsible for the charging of storm clouds remain in dispute

by Earle R. Williams

Lightning is one of the commonest and most spectacular of natural phenomena, and in the two centuries since Benjamin Franklin demonstrated that a lightning bolt is a giant electrical discharge, lightning and thunderstorms have been the subject of numerous scientific investigations. Yet, in spite of a barrage of new equipment and investigative techniques, lightning's exact origins and the mechanism by which rain clouds are electrified remain elusive.

The intractability of the problem stems from the fact that the physics of lightning and thunderstorms spans 15 orders of magnitude in scale. At the one end are the atomic phenomena that initiate the electrification of the storm cloud and that take place on scales of 10^{-13} kilometer; at the other end is the air motion of the full thundercloud, which completes the charging process and may take place over scales of tens or hundreds of kilometers. At each scale significant physics is not understood.

Franklin himself, perhaps unknowingly, identified one of the basic difficulties. In 1752 he observed that "the clouds of a thunder-gust are most commonly in a negative state of electricity, but sometimes in a positive state." Whether this ambiguity is the result of faulty observations or inherent in nature has been clarified only recently. Nevertheless, since Franklin's words were written, it has been accepted that lightning is the transfer of either positive or negative electric charge from one region of a cloud to another or between the cloud and the earth. For this charge transfer to take place the cloud must be electrified, that is, the positive and negative charges must be separated. How does charge separation come about?

As will become apparent, only a partial answer to this question can be given. In objects one ordinarily encounters, such as coffee cups or telephones, there are equal numbers of positive and negative charges; moreover, these charges are spread uniformly over the object, which is then said to be electrically neutral, or uncharged. Many microphysical processes, however, might cause the charges to separate, with the result that, although the object as a whole remains neutral, one region has more positive or negative charges than another. The object is then said to be charged or electrified. Charge separation is measured in volts; the greater the separation, the greater the voltage. When you walk across a room, the entire room remains neutral but the action of your shoes on the rug may charge the rug with one polarity and your shoes and body with the opposite polarity. This can lead to a potential difference of 100,000 volts over a distance of centimeters, a charge that is

EARLE R. WILLIAMS is assistant professor of meteorology in the Department of Earth, Atmospheric and Planetary Sciences at the Massachusetts Institute of Technology. He received a B.A. in physics at Swarthmore College in 1974 and a Ph.D. in geophysics from M.I.T. in 1981. His interest in lightning stems from a summer at the Brookhaven National Laboratory, where he observed large-scale electrical breakdowns on high-voltage terminals, a great nuisance to nuclear physicists but sufficiently spectacular to initiate a career in the physics of thunderstorms. His current research interests center on radar meteorology.

LIGHTNING STRIKES SEATTLE in an unusually large storm on July 31, 1984. A typical lightning bolt bridges a potential difference of several hundred million volts; it transfers the charge of approximately 10^{20} electrons in a fraction of a second, for a peak current of up to 10 kiloamperes. A moderate thunderstorm generates several hundred megawatts of electrical power, equivalent to the output of a small nuclear power plant.

evident when you grasp the doorknob.

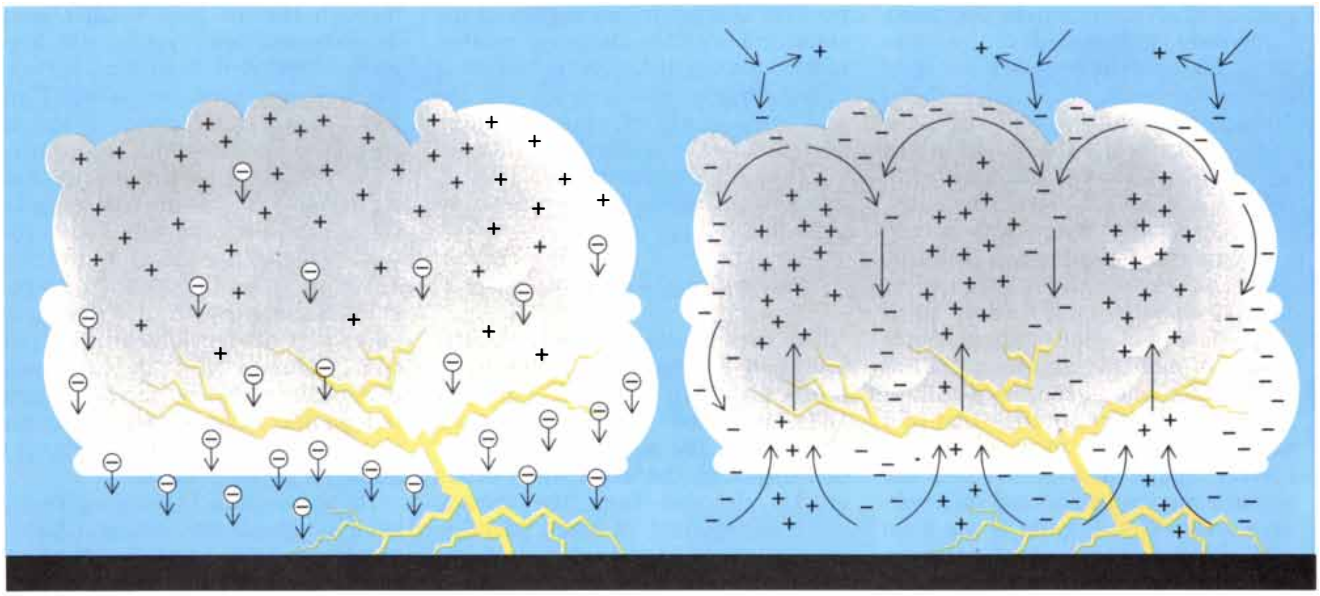
A typical lightning bolt represents a potential difference of several hundred million volts, and it may transfer 10 or more coulombs of charge to the ground; this is the charge carried by about 10^{20} electrons. The transfer of one coulomb of charge in one second is by definition an electrical current of one ampere. A lightning bolt therefore represents a current of much more than 10 amperes since its duration is much less than one second. Storm clouds of modest size produce a few flashes per minute and a power of a few hundred megawatts—that of a small nuclear power plant. To find the correct charge distribution and the physical mechanism behind such voltages and power outputs is the main task of thunderstorm physics. Historically investigations have centered on the electrical structure of clouds.

After Franklin's observation it was natural to assume that the charge distribution in a rain cloud conformed to the simplest pattern imaginable: positive charges in one region of the cloud and negative charges in another region. Such a structure is termed a dipole. In attempting to explain the presumed dipole structure of thunderclouds, investigators have invoked two very different models: the precipitation hypothesis and the convection hypothesis.

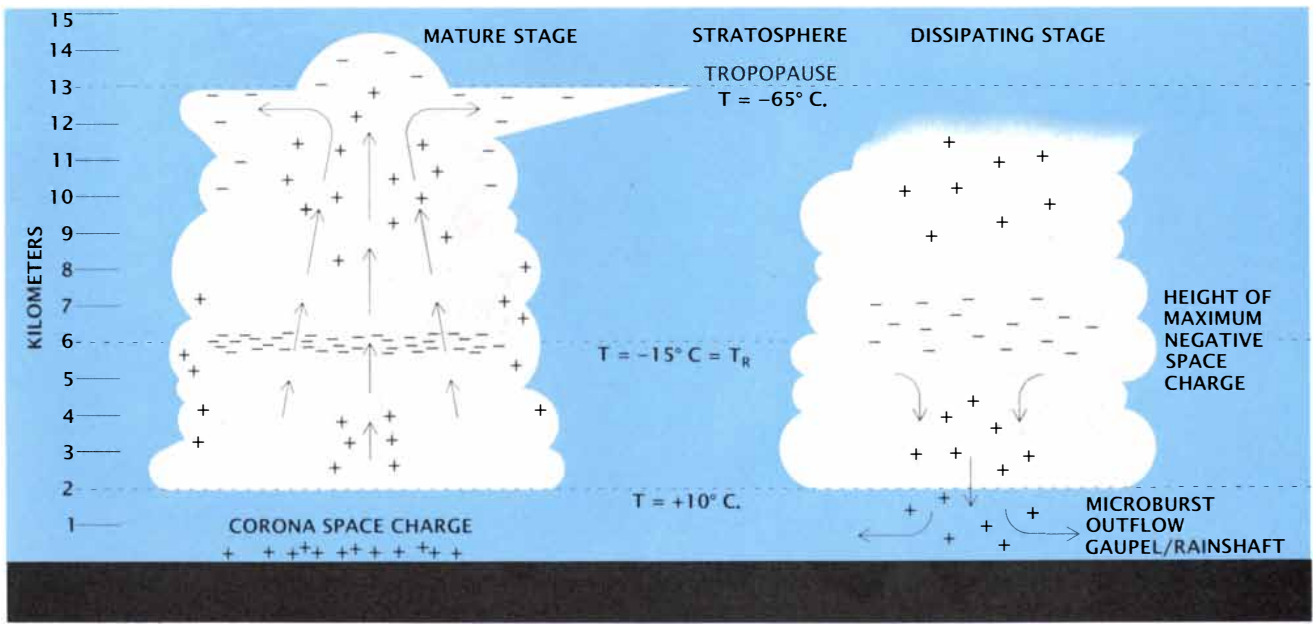
Precipitation vs. Convection

The precipitation hypothesis, first proposed by the German physicists Julius Elster and Hans F. Geitel in 1885, is based on a phenomenon observed in the working of any garden sprinkler: the larger water drops quickly descend from the stream, whereas the mist of small particles remains suspended in the air to be blown away by the wind. In the same way, the precipitation hypothesis assumes that raindrops, hailstones and graupel particles (millimeter-to-centimeter-size ice pellets) in a thundercloud are pulled by gravity downward through the air past smaller water droplets and ice crystals, which remain suspended. Collisions between the large precipitation particles and the mist of water droplets and ice crystals are conjectured to transfer negative charge to the precipitation particles (as charges are transferred to shoes from a rug) and, by charge conservation, positive charge to the mist. It follows that if the precipitation particles become negatively charged, the lower part of the cloud will accumulate negative charge and the upper part positive charge [*see top illustration on next page*]. A charge structure with the positive region uppermost is termed a positive dipole.

The convection hypothesis, formulated independently by Gaston Grenet of the University of Paris in 1947 and by Bernard Vonnegut of the State University of New York at Albany in 1953, is somewhat more complicated. The analogue here is the familiar Van de Graaff generator. In such a device a positive or negative electric charge is



TWO MODELS attempt to explain the electrical structure of thunderclouds. The precipitation model (*left*) suggests that gravity pulls heavy raindrops, hailstones and millimeter-size ice particles called graupel past smaller water droplets and ice crystals, which remain suspended. Collisions between the falling particles and the suspended mist are conjectured to transfer positive charge to the mist and negative charge to the heavier particles. As these heavier particles fall, the lower part of the cloud becomes negatively charged and the upper part becomes positively charged—a structure known as a positive dipole. The convection hypothesis (*right*) proposes that warm air currents carry positive charges released from the earth's surface to the top of the cloud. Negative charges, produced by cosmic rays above the cloud, are attracted to the cloud's surface by the positive charges within it. The negative charges attach themselves to cloud particles to form a negative "screening layer." Downdrafts are assumed to carry the negative charges downward; this process again results in a positive dipole. Note that the convection model invokes no precipitation and the precipitation model no convection.



ACTUAL STRUCTURE of a thundercloud is not dipolar but tripolar, with a main negatively charged region sandwiched between two positively charged regions. In a mature thundercloud (*left*) the main region of negative charge is at a height of about six kilometers and a temperature of about −15 degrees Celsius. Its thickness is only a few hundred meters, giving it a pancake shape. The upper positive region often extends to the tropopause, at a height of about 13 kilometers. At the very top of the cloud there is a thin layer of negative charges called the screening layer; its origins may be due to cosmic rays, which ionize air molecules. At the bottom of the cloud there is a second region of positive charge, smaller than the first. In a mature thundercloud updrafts dominate (*arrows*), but in a dissipating thundercloud (*right*) the lower region of positive charge precipitates out with strong downdrafts. The naive precipitation model does not account for the tripolar structure of thunderclouds. The convection hypothesis does this by assuming that the lower positive region is produced by so-called corona discharge given off by sharp objects on the earth. Recent evidence, however, indicates that the correct explanation for the tripolar structure lies in the microphysics of charge transfer between graupel particles and ice crystals.

sprayed onto a moving rubber belt, which then transports the charges, or ions, to a high-voltage terminal. The convection model assumes that the electric charges in the cloud are supplied initially by two external sources. The first source is cosmic rays, which impinge on air molecules above the cloud and ionize them (separate the positive and negative charges). The second source is the strong electric field around sharp objects at the earth's surface, which produces a "corona discharge" of positive ions. These positive ions are carried upward by warm air, which, rising by convection, acts like the belt on the Van de Graaff generator. After reaching the upper regions of the cloud, these positive ions attract the negative ions that were formed by cosmic rays above the cloud. The negative ions enter the cloud and quickly attach themselves to water droplets and ice crystals, thereby forming a negatively charged "screening layer." By hypothesis the downdrafts at the cloud's periphery then carry the negatively charged particles of the screening layer downward; this again results in a positive dipole structure.

Although precipitation and convection are observed in all clouds that produce lightning (and these phenomena are indeed inseparable in large clouds), one sees that the elementary precipitation hypothesis invokes no convection and the convection hypothesis invokes no precipitation. The marked distinction between the two models has played an important role in guiding investigators toward an understanding of the respective roles of precipitation and convection in the electrification of clouds.

These models were developed to explain the dipole structure of thunderclouds. But, as already mentioned, Franklin's first observation in 1752 hints at an ambiguity: Is the positive or negative charge uppermost? This question led to an early controversy between C. T. R. Wilson and George C. Simpson over the charge structure of thunderclouds. The debate is instructive, for it shows some of the difficulties in collecting meaningful thunderstorm data.

Positive Dipole or Negative?

In the 1920's Wilson, who earlier had invented the cloud chamber, made observations of a number of thunderstorms from a distance and concluded that the basic structure of a thundercloud was that of a positive dipole. At about the same time Simpson, measuring the charge on rain falling from thunderclouds, concluded the opposite: that the lower region of a thundercloud was positively charged and the upper region was negatively charged—a negative dipole.

It is only within the past 20 years that investigators have been able to explain these seemingly incompatible results. With the benefit of hindsight one can say the most important reason for the persistence of the discrepancy is that one rarely measures the charge in a thundercloud; it is inferred from a measurement of the cloud's electric field. An electric field surrounding a charged body is analogous in almost every respect to the gravitational field surrounding a massive body. Both fields cause other objects within them to move; gravity attracts massive objects, whereas an electric field attracts or repels charged objects. The gravitational or electrical force acting on such "test particles" diminishes with the square of the distance between the test particle and the central body. Both fields are therefore characterized by a strength (determined by the distance to the attractive or repulsive body) and a direction (attractive or repulsive). Fields characterized by a strength and a direction are called vector fields.

When more than one charged body is present, the electric field can become very complicated. Moreover, any number of charge configurations can produce the same field strength and direction at a given point. As a result a single measurement of the electric field cannot uniquely determine the charge distribution. Many measurements are required; in principle one must actually measure the electric field everywhere to deduce the true charge distribution. Wilson and Simpson each made measurements from a single position, which is not enough to infer the charge structure correctly.

Since the Wilson-Simpson controversy, 50 more years of observation have established that the basic structure of thunderclouds is not dipolar but tripolar: there is a main region of negative charge in the center with one region of positive charge above it and a second, smaller region of positive charge below it [*see bottom illustration on opposite page*]. The most notable feature of the main, negatively-charged layer is its pancake shape: its vertical thickness is less than a kilometer, but it may extend horizontally several kilometers or more. It is at an altitude of approximately six kilometers, where the temperature is roughly −15 degrees Celsius. Under conditions prevailing there all three phases of water—ice, liquid and vapor—can coexist. The largest electric fields in the thundercloud are found at the upper and lower boundaries of the main negatively charged layer.

The upper region of positive charge is more diffuse than the negative layer and may extend vertically several kilometers—as high as the cloud itself. The lower region of positive charge, on the other hand, is so small that the electric field at the surface of the earth is frequently dominated by the main negative charge. One other feature is observed in many clouds: a layer of negative charge, about 100 meters thick, above the upper positive region. This layer may result from negative ions produced above and outside the cloud, which are then captured by cloud droplets or ice particles; it is the screening layer predicted by the convection hypothesis. Regardless of its origins, however, the screening layer appears to be a secondary feature that does not significantly alter the basic tripolar structure of the cloud.

The tripolar structure allows one to understand the Wilson and Simpson results. Wilson made his observations from a considerable distance; the electrical effect of the small positive region at the base of the cloud was overwhelmed by the main negative region. He therefore saw only the uppermost positive charge and a negative charge below it: a positive dipole. On the other hand, Simpson's observations were carried out right under the cloud. His instruments detected the lower positive region directly above him. Since the higher negative central region screened the top positive region, Simpson concluded that the negative charge was uppermost and hence that the structure of the cloud was a negative dipole.

Microphysics

The tripolar structure of thunderclouds requires some modification of the naive precipitation model, which can account only for a simple dipole, quite apart from the fact that the microphysics of charge transfer was left virtually unexplained. On the other hand, it might seem that the convection model leads more naturally to a tripole structure because it assumes that corona discharge from sharp objects on the surface of the earth produces a flux of positive charges toward the base of the cloud. It has been widely believed this flux might account for the lower, positively charged region of the tripole. Recent measure-

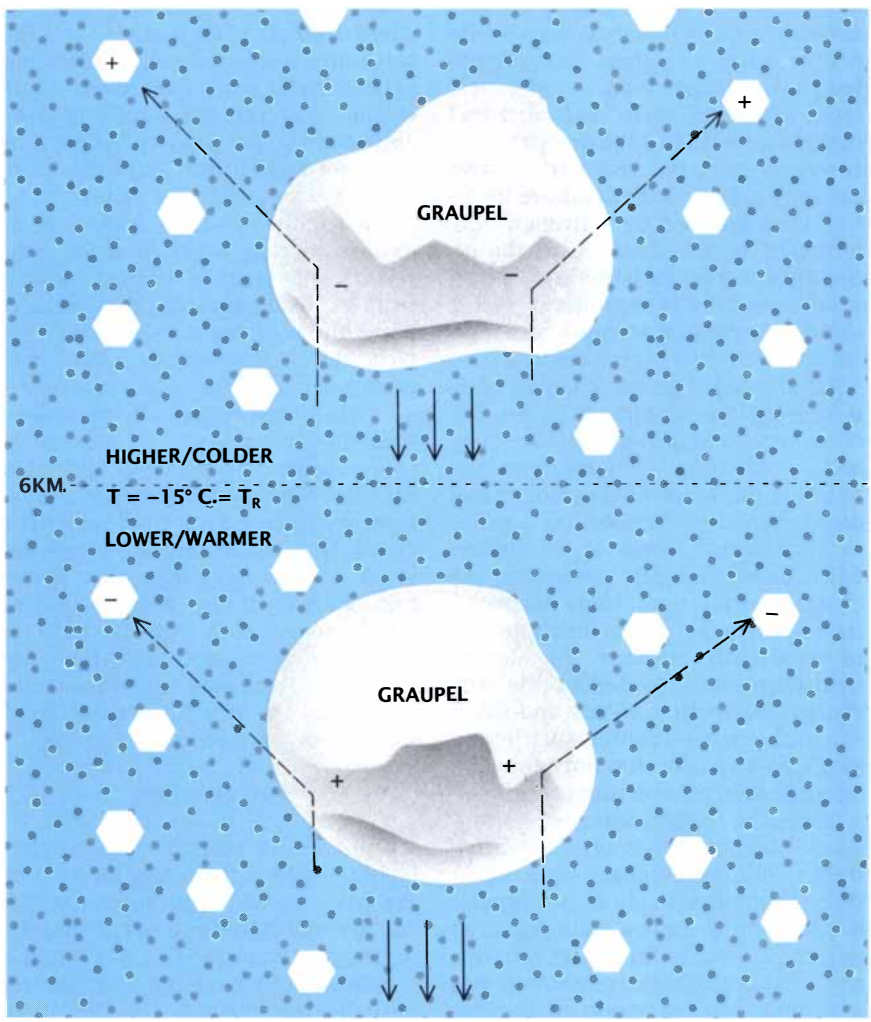
ments of the size of the flux, however, suggest that it may be an order of magnitude too small to account for the observed rate of cloud charging. Partly for this reason the convection model has fallen into disfavor. One therefore attempts to modify the precipitation model.

Several modifications of the precipitation model have been proposed to account for the lower positive charge region, as well as for the fact that rain ordinarily carries a positive charge. Simpson made the first attempt at explaining these observations. Empirically it is known from studies of waterfalls that the larger droplets selectively acquire a positive charge on breakup. (The amateur can perform this experiment by taking a microammeter on board the *Maid of the Mist* at Niagara Falls.) Simpson proposed that precipitating water droplets in a thundercloud also fragmented near the base of the cloud and thus accounted for the lower positive region. Measurements made under the main negative charge region in thunderclouds, however, show that precipitation particles carry charges that are substantially greater than those produced in the waterfall fragmentation process; this raises serious questions about whether fragmentation can account for the tripole's lower positive region. Furthermore, it is now recognized that most of the positively charged particles that fall below the main negative charge region are not water droplets but ice.

Ice plays a role in the other proposals to explain the thundercloud's tripole structure. Laboratory studies in the 1940's showed that ice particles pick up a strong positive charge in the course of melting. This observation is still often invoked to explain the lower positive region of the cloud. Although melting may conceivably explain the existence of positively charged particles at altitudes below 4,000 meters, where ice begins to melt in mid-latitude thunderstorms, it cannot, however, explain their existence at higher altitudes where positively charged particles are also observed.

Melting of ice does not appear to account for the observed tripole structure, but there is now considerable evidence that collisions between ice crystals and graupel particles play a fundamental role. Over the past 20 years laboratory studies by many investigators, notably Stephen E. Reynolds, Marx Brook and their collaborators at the New Mexico Institute of Mining and Technology, Tsutomu Takahashi of the University of Hawaii at Manoa and Clive P. R. Saunders, John Latham and Anthony J. Illingworth at the Victoria University of Manchester, have shown that when graupel particles collide with ice crystals, the polarity of the charge transferred to the particles is strongly dependent on temperature. Below a critical temperature, called the charge-reversal temperature, negative charge is transferred; at higher temperatures (corresponding to lower altitudes in thunderclouds) positive charge is transferred [*see illustration at left*]. The exact value of the charge-reversal temperature is still a matter of dispute, but most laboratory investigators agree that its value is between −20 and −10 degrees C.

Thundercloud observations by a variety of methods show that the main negatively charged layer is at an altitude where the temperature is about −15 degrees. The charge-reversal hypothesis then explains why negative charges are found less frequently below this altitude: the graupel particles become positively charged as they fall past and collide with suspended ice crystals. These falling positive charges form the lower positive region of the tripole. Moreover, the amount of charge transfer per collision in the laboratory is of sufficient magnitude to account for the charge transferred by lightning in clouds of modest electrical activity. The establishment of a charge-reversal temperature consistent with both laboratory experiments



MICROPHYSICS OF CHARGE TRANSFER involves collisions between graupel particles and ice crystals. The heavy graupel particles fall through a suspension of smaller ice crystals (*hexagons*) and supercooled water droplets (*dots*). Laboratory experiments show that when the temperature is below a critical value called the charge-reversal temperature, T_R, the falling graupel particles acquire a negative charge in collisions with the ice crystals. At temperatures above T_R they acquire a positive charge. T_R is thought to be about −15 degrees C., the temperature of the main negative region found in thunderclouds; thus graupel picks up a positive charge when it falls below this altitude to higher temperatures. There is now evidence that these positively charged graupel particles form the lower positive region of the thundercloud tripole.

THE X-30 IS KID STUFF.

The power to break the bonds of Earth. To leap up and away. To fly 18,000 miles an hour, 200,000 feet high. For a kid, that's the stuff dreams are made of.

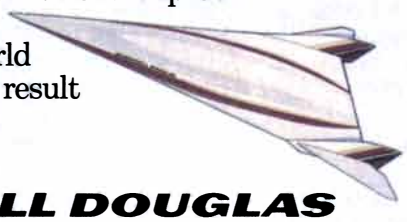
Kid stuff, because it's your kids, and ours, who will gain most from the advances in propulsion technology required for the X-30 National AeroSpace Plane (NASP).

We'll need new, advanced powerplants, including those which move air through the engine at supersonic speeds. Fuel? Liquid methane, or hydrogen which would serve as both fuel and coolant for the plane's superhot skin surfaces during flight.

The X-30 is kid stuff, because it is our children who will move from Earth to space and back at far less cost than we do today. And one day fly from Los Angeles to Tokyo in two hours aboard an Orient Express.

And it's they who will have—or not have—world leadership in aerospace as a result of America's work today.

Let's go full throttle.



MCDONNELL DOUGLAS

and thundercloud observations must be considered the main advance in thunderstorm electricity in the past two decades.

At the same time, the exact microphysical processes that would explain the systematic transfer of charge of one polarity to the graupel particles, as well as the reversal temperature, remain almost entirely unknown. The underlying physical mechanism may well be related to whatever causes the shoes to charge when one walks on a rug or a glass rod to charge when it is rubbed with a piece of wool. Although these phenomena were known to the ancients, however, the basic microphysics behind them remains to this day a neglected and unsolved problem. The lack of a microphysical description of static electrification is the most serious gap in the understanding of thundercloud electricity.

Convection

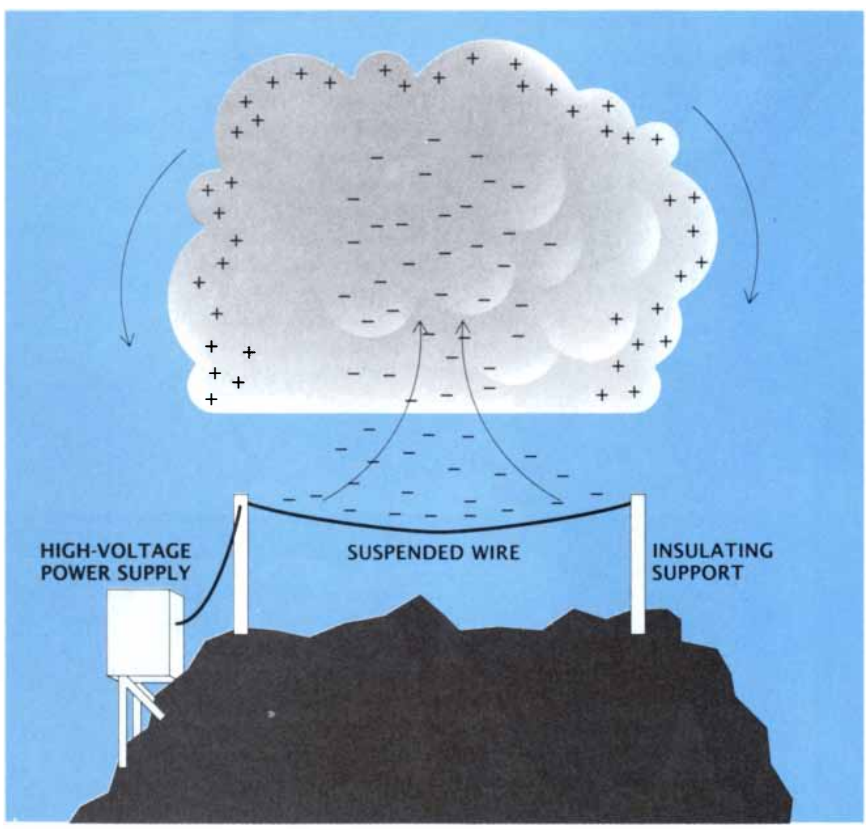
Although the convection model may be inadequate in accounting for the magnitude of the lower positive region, there is substantial evidence that thunderstorms are regions of vigorous updrafts and downdrafts; convection is indisputably present. It has also been observed that maximum lightning-flash rates are associated with the upward motion of graupel and hail above the main region of negative charge. This picture contradicts the naive precipitation hypothesis, in which only downward-moving graupel particles cause the electrification. It is probably the relative motion between the ice crystals and the graupel particles, however, that causes large-scale charge separation. The important requirement is that the ice crystals rise relative to the earth more rapidly than the graupel particles; this is equivalent to falling graupel. Moreover, vigorous updrafts are not only consistent with electrification but also essential for it: the updraft maintains the supply of supercooled water droplets above the charge-reversal altitude. These droplets provide for the growth of graupel particles required for electrification and, according to the laboratory experiments described above, negligible charge transfer takes place between graupel particles and ice crystals if the droplets are absent.

During the past decade thunderstorm downdrafts have been a growing concern from the point of view of aviation safety. Unusually strong downdrafts, termed microbursts by Tetsuya T. Fujita of the University of Chicago, are believed to have caused major commercial-airliner crashes. Recent studies of stationary thunderstorms have shown that these downdrafts follow by five to 10 minutes the time of peak updraft and maximum intracloud lightning activity; the downdrafts are also associated with the intense precipitation that results when the updraft collapses.

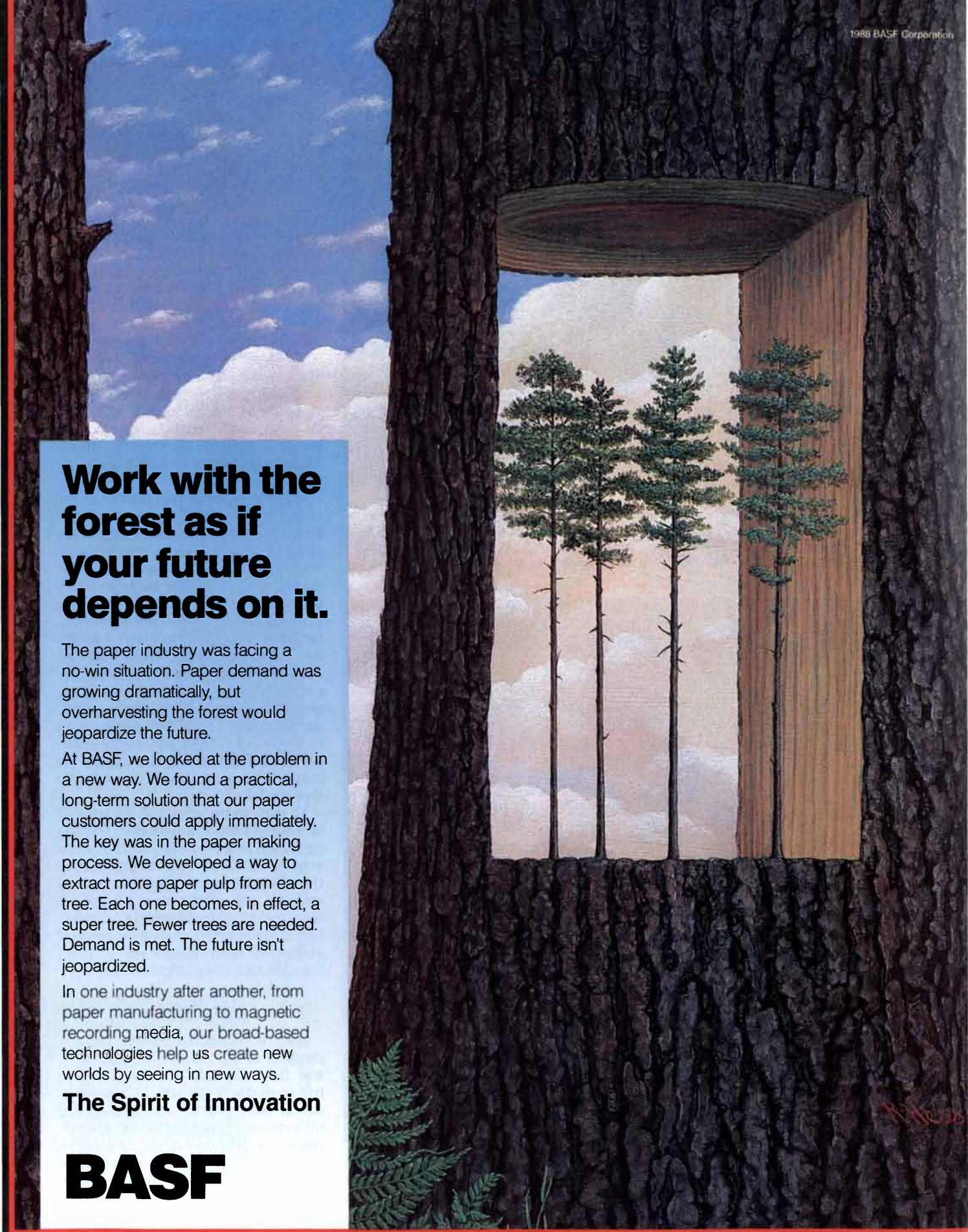
Measurements also show that the electric field at the ground reverses at this time from upward-directed to downward-directed. Positive charge is found on the precipitation, which suggests that the lower positive-charge region of the cloud is carried to the ground during the microburst phase of the downdraft. Both the intracloud lightning rate and reversals of the electric field might serve as valuable precursors to warn air traffic controllers about hazardous surface conditions.

Given that such strong convection currents are characteristic features of thunderstorms, one might think that the convection model holds promise in explaining some aspects of cloud electrification. As already mentioned, it does predict the screening layer. It is for these reasons that researchers, notably Charles B. Moore of the New Mexico Institute of Mining and Technology and Vonnegut, have continued to test the model. In their experiments air under a fair-weather cumulus cloud was charged by a wire connected to a high-voltage terminal [*see illustration at left*]. Electrical observations from an airplane showed that the charge released by corona discharge from the wire was carried up through the cloud by convective air motions. What is more, when positive charge was released, the top of the cloud became positively charged and the bottom negatively charged—a positive dipole. But when the polarity of the charges released from the power supply was changed to negative, the cloud took on a negative-dipole structure. These results indicated that convection was carrying charge to the top of the clouds.

The charge within the clouds produced by these experiments, how-



EXPERIMENTS by Charles B. Moore and Bernard Vonnegut test the convection hypothesis. Air under a cumulus cloud is charged positively by a wire connected to a high-voltage terminal. Measurements from an airplane show that the charges are carried upward through the cloud by convection. Reversing the polarity of the power supply reverses the cloud's polarity. The experiments indicate that convection is carrying the charges upward; the field produced, however, is about 1,000 times smaller than what is necessary to produce lightning in electrically active clouds, and so the relevance of these results to thunderclouds is still open to question.



Work with the forest as if your future depends on it.

The paper industry was facing a no-win situation. Paper demand was growing dramatically, but overharvesting the forest would jeopardize the future.

At BASF, we looked at the problem in a new way. We found a practical, long-term solution that our paper customers could apply immediately. The key was in the paper making process. We developed a way to extract more paper pulp from each tree. Each one becomes, in effect, a super tree. Fewer trees are needed. Demand is met. The future isn't jeopardized.

In one industry after another, from paper manufacturing to magnetic recording media, our broad-based technologies help us create new worlds by seeing in new ways.

The Spirit of Innovation

BASF

INTRACLOUD lightning bolts are actually far more frequent than cloud-to-ground flashes but are seen less often because clouds strongly scatter light in the visible spectrum. Radar, radio-frequency direction finders and microphones, which "see" through clouds, are now helping investigators to study intracloud lightning.

LOOP-THE-LOOP shows that lightning paths do not follow obvious directions. Conflicting claims have appeared in the literature that the paths are random, are determined by the electric field configuration or are determined by the distribution of electric charge in space. The last explanation may account for many observations.

ever, was about 100 times less than is found in active thunderstorms and the electric field was more than 1,000 times less than that believed necessary to initiate lightning. Therefore the experiments did not directly test the role of convection in precipitating, electrically active thunderstorms and so the results cannot be said to strongly support the convection hypothesis.

Similar experiments have recently been performed under larger, precipitating clouds. Artificially released negative charge has resulted in a few cases with negative charge dominating above and positive charge below. This is consistent with the previous results and the convection hypothesis. The interpretation of the results is not unambiguous, however: because precipitation is now taking place, the lower positive charge may also be the result of the charge-transfer microphysics between ice and graupel already described.

Another observation relevant to the convection hypothesis is less ambivalent: the fact that the main negatively charged region is at a roughly constant altitude and temperature. In the convection model, air currents carry the negatively charged particles of the screening layer downward across distances of several kilometers. It is difficult to understand, then, why the negative charge should be mainly concentrated in a pancake-shaped region only a few hundred meters thick. As discussed above, this observation is better explained by charge-reversal microphysics and is perhaps the major argument against the convection model.

To sum up the status of the precipitation and convection models, the precipitation model can account for more aspects of cloud electrification than the convection model, but it does so by ignoring one of the most prominent features of thunderstorms: convection. Presumably in the future the best aspects of both models will be combined in one comprehensive theory.

Lightning

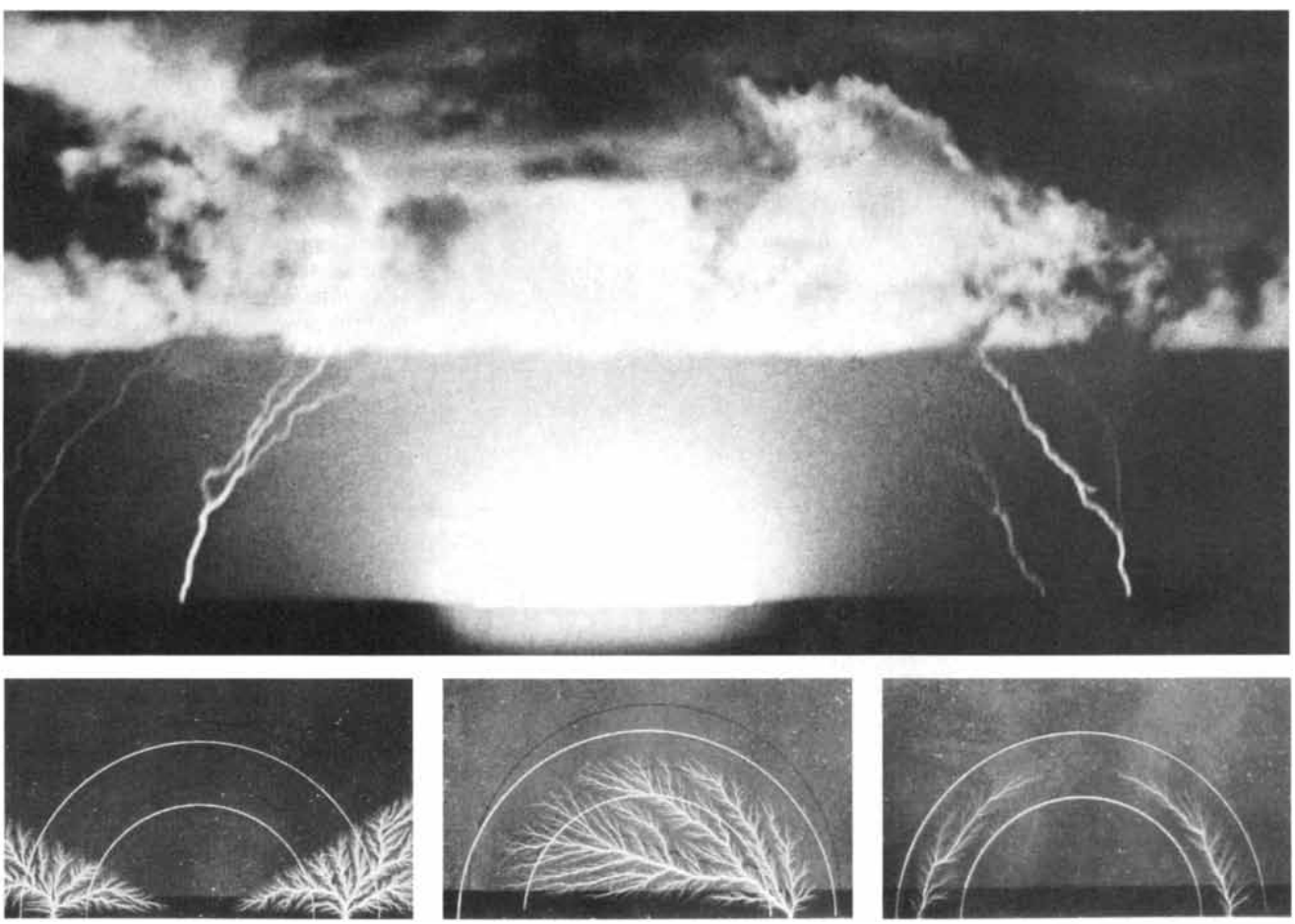
Once a thundercloud has become charged to the point where the electric field exceeds the local dielectric strength of the atmosphere—that is, the strength of the atmosphere to support a separation of electric charge—a lightning flash results. The electric field at this instant is on the order of one million volts per meter, and in less than a second the lightning bolt will transfer 10^{20} electrons' worth

of charge and provide the electric power equivalent to about 100 million ordinary light bulbs. During that fraction of a second the electrostatic energy of accumulated charge is transformed into electromagnetic energy (the visible flash as well as radio interference), acoustical energy (thunder) and ultimately heat.

Almost all natural lightning is initiated within the cloud and evolves as a double-ended "tree," with one end invading negative-charge regions and the other end invading positive-charge regions. In the case of a cloud-to-ground discharge the negative end of the tree becomes a "stepped leader," which carries a negative current of a few hundred amperes downward. When the stepped leader is within roughly 100 meters of the ground, a return stroke is initiated, which transfers a 10-kiloampere current, or 10,000 coulombs of positive charge per second, upward. It is the luminous return stroke that one actually sees with the eye and so, when one speaks of cloud-to-ground lightning, one must bear in mind that the lightning travels in both directions—sometimes several dozen times [see "Thunder," by Arthur A. Few; SCIENTIFIC AMERICAN, July, 1975].

Early studies of lightning concentrated on the cloud-to-ground flashes because those were most accessible to visual and photographic observation. It turns out, however, that lightning is far more frequent and extensive within the cloud itself, where it is hidden from view because of the opacity of the cloud. More recent studies have attempted to investigate intracloud lightning with the help of radar, radio-frequency direction finders and microphones. Interest centers on the paths taken by lightning flashes and their relation to cloud structure.

Lightning is found in regions of precipitation and in regions without precipitation, within clouds and without, and its paths often appear to be very chaotic; many claims have been made that these paths are indeed random. Theoretical models for lightning paths have often focused on the role of the electric field. In other words, it was thought the strength and direction of the local electric field determined the lightning path. Until recently little attention has been paid to the role of the electric charge itself. Although the charge gives rise to the field, so that one might think knowing one is as good as knowing the other, one must also remember that an electric field can be produced locally by any number of different charge distributions. Therefore the location of the electric



IVY-MIKE TEST of a 10-megaton hydrogen bomb in 1952 generated lightning within 10 milliseconds after detonation. The intense gamma-ray burst from the explosion strips electrons from air molecules in a process called Compton scattering; the lighter electrons are rapidly moved away from the now positive air molecules, resulting in charge separation. The hemispherical symmetry of the explosion allows the charge distribution to be simulated in the laboratory. The three experiments shown at the bottom were done at the Massachusetts Institute of Technology's High Voltage Research Laboratory, where electric charge was injected into specific regions of an insulating plastic block. The models have a similar electric field near the ground but differing charge distributions. Only model 3 reproduces the IVY-MIKE lightning pattern, indicating it is the charge distribution, not the field, that predominantly determines the path of lightning. The lightning is triggered at the point on the ground where the electric field is large, and it travels upward through the region of greatest negative charge.

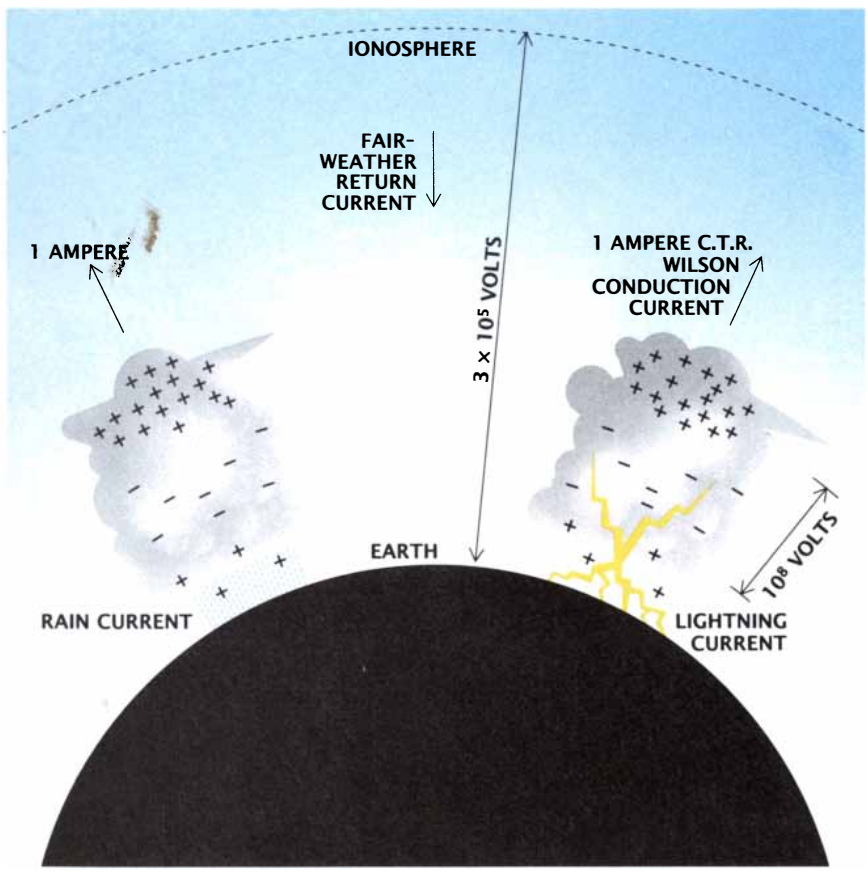
charge is a different piece of information from knowing the local field configuration. An additional complication is that the charge distribution and field are not static but dynamic; as the lightning forms and grows it will change the field dramatically, making modeling much more difficult.

Today the evidence indicates that the two ends of the lightning "tree" tend to follow paths of greatest charge concentration. For example, many observations reveal that lightning is found predominantly within the main negative-charge region.

The clearest evidence that lightning paths are governed by the distribution of charge in space, however, probably comes from studies of the behavior of lightning produced by nuclear-weapons explosions. Photographs of H-bomb tests carried out in the 1950's show that the fireball is often surrounded by lightning flashes [*see illustration on preceding page*]. In contrast to the theoretical picture of thunderclouds, here the basic charge-separation mechanism is well understood. The radial flux of high-energy photons given off by the fireball strips the surrounding air molecules of their electrons in a process called Compton scattering. The negatively charged electrons are thus concentrated in a hemispherical shell around ground zero, leaving a positively charged region in the fireball.

The hemispherical symmetry of the explosion makes it possible to construct simple theoretical and laboratory models of the charge distribution and investigate its effect on lightning paths. Chathan M. Cooke, Kenneth A. Wright and I have performed such simulations at the Massachusetts Institute of Technology's High Voltage Research Laboratory. Charge is injected in an annular pattern into blocks of highly insulating plastic that trap the charge in a manner consistent with the theoretical model. The resulting electric field is strong enough to trigger lightninglike discharges.

We find that lightning is triggered at the point near simulated ground zero where the electric field is strongest; the flash then travels upward through the region of greatest negative charge. The lightning pattern is remarkably similar to the patterns seen in the photographs of nuclear tests. Furthermore, one can construct experiments with charge distributions predicted by other theoretical models. Although some of these other models produce the same electric-field configuration at the ground, the charge distribution can be quite different. The lightning patterns do not resemble those of the nuclear tests, thereby demonstrating that it is predominantly the charge that determines the lightning path.



GLOBAL CIRCUIT is charged by thunderstorm batteries. Between the negatively charged earth and the upper atmosphere is a nearly constant potential difference of 300,000 volts. On the scale of clouds, currents of about one ampere that flow from the top of thunderstorms help to maintain the potential difference; this requires a similar current to flow from the ground to the clouds' lower regions. On the large scale, a fair-weather leakage current of about 2,000 amperes, which transfers positive charge from the upper atmosphere to the earth, would eliminate the 300-kilovolt potential difference if thunderstorms did not recharge the circuit. It is thought that thunderstorms in the Tropics, which transfer large amounts of negative charge to the ground, are the dominant agent in balancing the fair-weather current.

Energy and the Global Circuit

Most of the electrical energy of a thunderstorm is thought to be released in the form of lightning. As I mentioned above, a modest thunderstorm produces a few flashes per minute and has a power output roughly equivalent to that of a nuclear plant. Simple scaling laws based on the equations of electromagnetism show that the power output goes up approximately as the fifth power of the cloud size; doubling the dimensions of the cloud increases the power output by about thirtyfold. Large thunderstorms may produce lightning at rates in excess of 100 flashes per minute.

A well-known rule in physics is that there are no free lunches. The electrical energy released by lightning must come from somewhere. Ultimately it derives from the heat that causes water vapor to expand, to become less dense than the surrounding air and consequently to rise. As the water vapor rises it condenses or freezes; the latent heat is released and the liquid water or ice then begins to fall. The gravitational potential energy released by the falling precipitation is, according to the precipitation model, the energy available for cloud electrification; it is computed as the product of the gravitational force acting on the precipitation and the fall distance of the precipitation.

Radar measurements of falling rain and graupel particles show that in modest storms the gravitational energy is in fact much larger than the electrical energy released by lightning. For exceedingly active storms, where

the electrical energy may be orders of magnitude higher, the gravitational energy and the electrical energy are estimated to be about the same. One would then expect, by conservation of energy, that at the moment of a lightning discharge, when the electrical forces suddenly decrease, the fall velocity of the precipitation should noticeably increase. Attempts have been made to measure the phenomenon by Doppler radar, which measures the velocity of a moving object, but so far the attempts have been unsuccessful. The general absence of abrupt shifts in velocity has not yet been satisfactorily explained, but small velocity changes may be masked by the turbulent motions of thunderstorms.

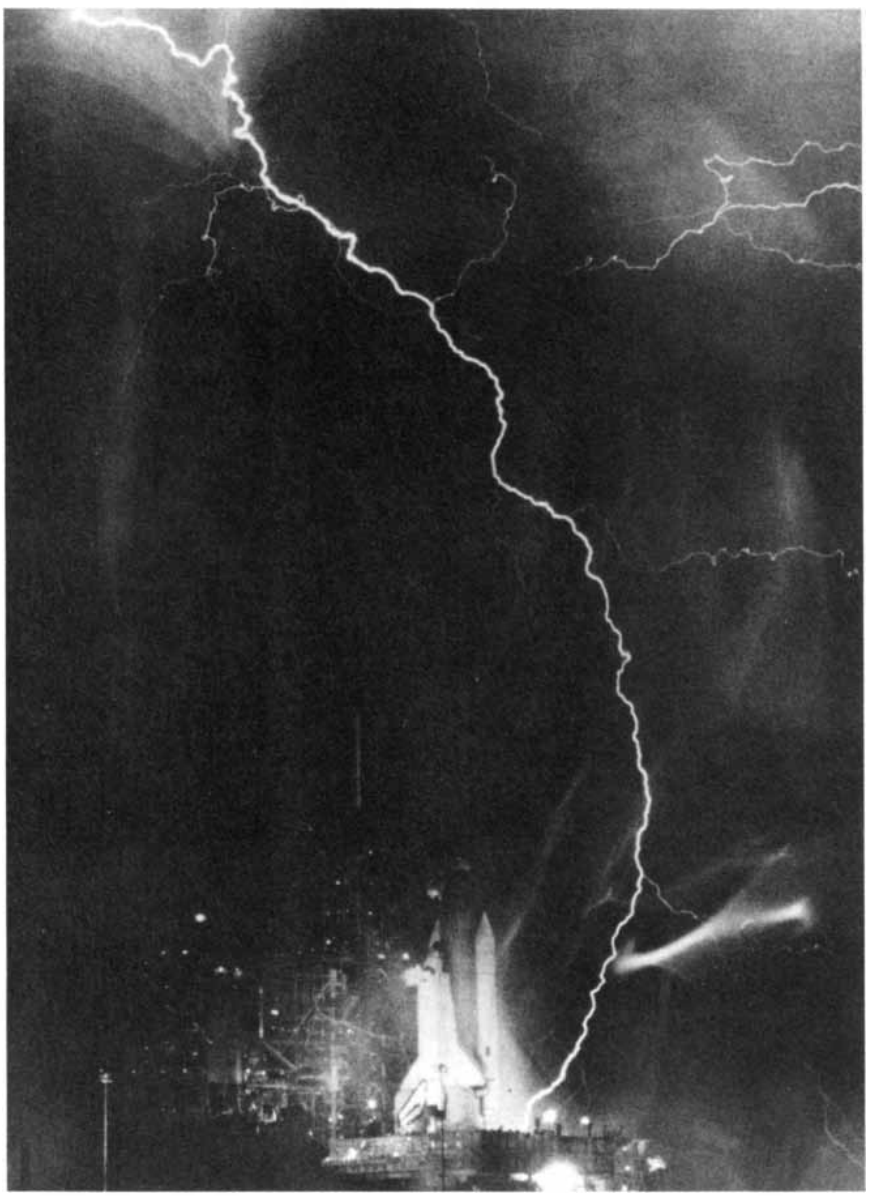
There is still another energy balance that must be maintained: that of the global electrical circuit. The earth's atmosphere is an extremely good insulator that is sandwiched between two good conductors: the earth's surface below and the upper atmosphere and ionosphere above [*see illustration on opposite page*]. These layers are the passive components of the global electrical circuit.

Between the negatively charged surface of the earth and the positively charged atmosphere is a steady potential difference of about 300,000 volts. Following the proposal originally made more than 70 years ago by Wilson, it is now generally believed this 300-kilovolt "ionospheric potential" is the result of charging by thunderstorms, which form the "batteries" of the global circuit. Electric currents of about one ampere per storm flow upward from the positive tops of thunderclouds and return to the earth in the fair-weather regions of the atmosphere.

In order for charge not to build up indefinitely in the clouds, a one-ampere current has to flow from the earth's surface to the cloud bottom. Rain currents, corona discharge and lightning all contribute to this charge transfer, but in the mid-latitudes it is not enough to balance the fair-weather return current. Where is the deficit made up? The missing batteries are found in the Tropics, where thunderstorms that are orders of magnitude larger than mid-latitude storms have flash rates large enough to charge the global circuit.

One might well ask why the earth is negatively charged in the first place. The best guess today is that the negative charge of the earth is the result of the earth's proximity to the negative end of the thunderstorm battery. The question thus reverts to why the lower part of a thundercloud is predominantly negative—and the answer to this question, once again, appears to depend on the poorly understood microphysics of ice.

In spite of the many unanswered questions, a unified picture of cloud electrification is beginning to emerge; the picture links charge separation taking place on the scale of atoms to lightning flashes that travel across distances of kilometers to an electrical circuit that spans the entire earth.



PRELAUNCH LIGHTNING at the John F. Kennedy Space Center in Florida on August 30, 1983, almost strikes the space shuttle. The storm passed and the shuttle was launched on schedule at 2:32 A.M. Approximately 44,000 thunderstorms and eight million lightning flashes take place daily around the world. In the U.S. alone lightning annually causes about 150 deaths and $20-million worth of property damage and sets 10,000 forest fires, which destroy $30-million worth of marketable timber.

FURTHER READING

ATMOSPHERIC ELECTRICITY. John A. Chalmers. Pergamon Press, 1967.

THE THUNDERCLOUD. C. B. Moore and B. Vonnegut in *Lightning, Vol. 1: Physics of Lightning,* edited by R. H. Golde. Academic Press, 1977.

THE LIGHTNING DISCHARGE. Martin A. Uman. Academic Press, 1987.

THE ROLE OF ELECTRIC SPACE CHARGE IN NUCLEAR LIGHTNING. Earle R. Williams, Chathan M. Cooke and Kenneth A. Wright in *Journal of Geophysical Research,* Vol. 93, No. D2, pages 1679–1688; February 20, 1988.

THE ELECTRIFICATION OF THUNDERSTORMS. Earle R. Williams in *Journal of Geophysical Research,* in press.

Superfluid Turbulence

Liquid helium cooled to within 2.172 degrees of absolute zero can flow without viscosity or friction, but seldom without turbulence. The odd form of turbulence that arises is quantum-mechanical in nature

by Russell J. Donnelly

Sir Horace Lamb, a British mathematician and physicist, is said to have remarked in 1932: "I am an old man now, and when I die and go to Heaven there are two matters on which I hope for enlightenment. One is quantum electrodynamics, and the other is the turbulent motion of fluids. And about the former I am really rather optimistic." As Sydney Goldstein of Harvard University has written, "Lamb was correct on two scores. All who knew him agreed that it was Heaven he would go to, and he was right to be more optimistic about quantum electrodynamics than turbulence."

Most flows of fluids, in nature and in technology, are turbulent. Since much of the energy expended by machines and devices that involve fluid flows is spent in overcoming the drag caused by turbulence, there is a strong practical motivation to understand the phenomenon. The study of turbulent flows, however, is one of the most formidably difficult subjects in physics and engineering. At present there is no substantial aspect of turbulent flow that can be understood fully from first principles.

Surprisingly, the peculiar, quantum-mechanical form of turbulence that can form in superfluid helium (a form of liquid helium that flows without viscosity or friction) may turn out to be much simpler to understand than the classical turbulence that forms in normal fluids, such as quickly flowing streams or boiling water. The surprise is even greater when one realizes that until a few years ago this field of investigation was considered a backwater in condensed-matter physics.

It now seems that the study of superfluid turbulence may provide simplified model systems for studying some forms of classical turbulence. There are also practical motivations for studying superfluid turbulence. For example, superfluid helium is often used as a coolant in superconducting machinery. Superfluid turbulence is the primary impediment to the transfer of heat by superfluid helium; an understanding of the phenomenon may make it possible to design more efficient methods of refrigeration for superconducting devices.

The medium in which superfluid turbulence develops—superfluid helium—is an unusual one. At atmospheric pressure, helium gas liquefies when it is cooled to a temperature of 4.2 degrees Kelvin. (Zero degrees Kelvin—absolute zero—is about 273 degrees below zero Celsius.) The liquid can be cooled still further by applying a vacuum pump to the vessel it is in, which decreases the pressure in the vessel and hence lowers the boiling point of the liquid; pockets of liquid that are warmer than the new boiling point boil up as gas and are sucked out by the vacuum pump. As the liquid is cooled from 4.2 degrees to 2.172 degrees it bubbles vigorously. In this temperature range liquid helium is known as helium I; it behaves like an ordinary fluid, and it conducts heat relatively poorly.

At 2.172 degrees, however, the bubbling suddenly stops, and the liquid acquires a number of unusual properties; for example, it can now flow without viscosity or friction, and it seeps through cracks that are too narrow to be penetrated even by gases. This liquid, which maintains these properties as it is cooled to absolute zero, is called helium II, or superfluid helium [see "Superfluidity," by Eugene M. Lifshitz; SCIENTIFIC AMERICAN, June, 1958]. Liquid helium stops bubbling when it is cooled below 2.172 degrees because of another remarkable property of helium II: it has an extraordinarily high heat conductivity (about 10 million times the heat conductivity of helium I), and so it is impossible for isolated pockets of heat to build up and form bubbles. The discoverers of this fact, Willem H. Keesom and his daughter Annie Keesom of Leiden University, called helium II supra-heat-conducting, and with good reason.

What underlies the unusual properties of helium II? The answer is too complex to give in detail here, but the essential fact is that helium II is a truly quantum-mechanical fluid: at such low temperatures the laws of classical fluid mechanics are supplanted by the laws of quantum mechanics.

According to quantum mechanics, an atom must occupy one of a specific set of energy levels, or states. The amount of energy inherent in a given state depends on the energies of the atom's electrons and on properties of the atom as a whole, including its vibration and state of motion. As an atom moves about it can acquire energy (from collisions, for example), but it must acquire it in set amounts: in any one event it must gain at least enough energy to drive it up into the next-higher energy state.

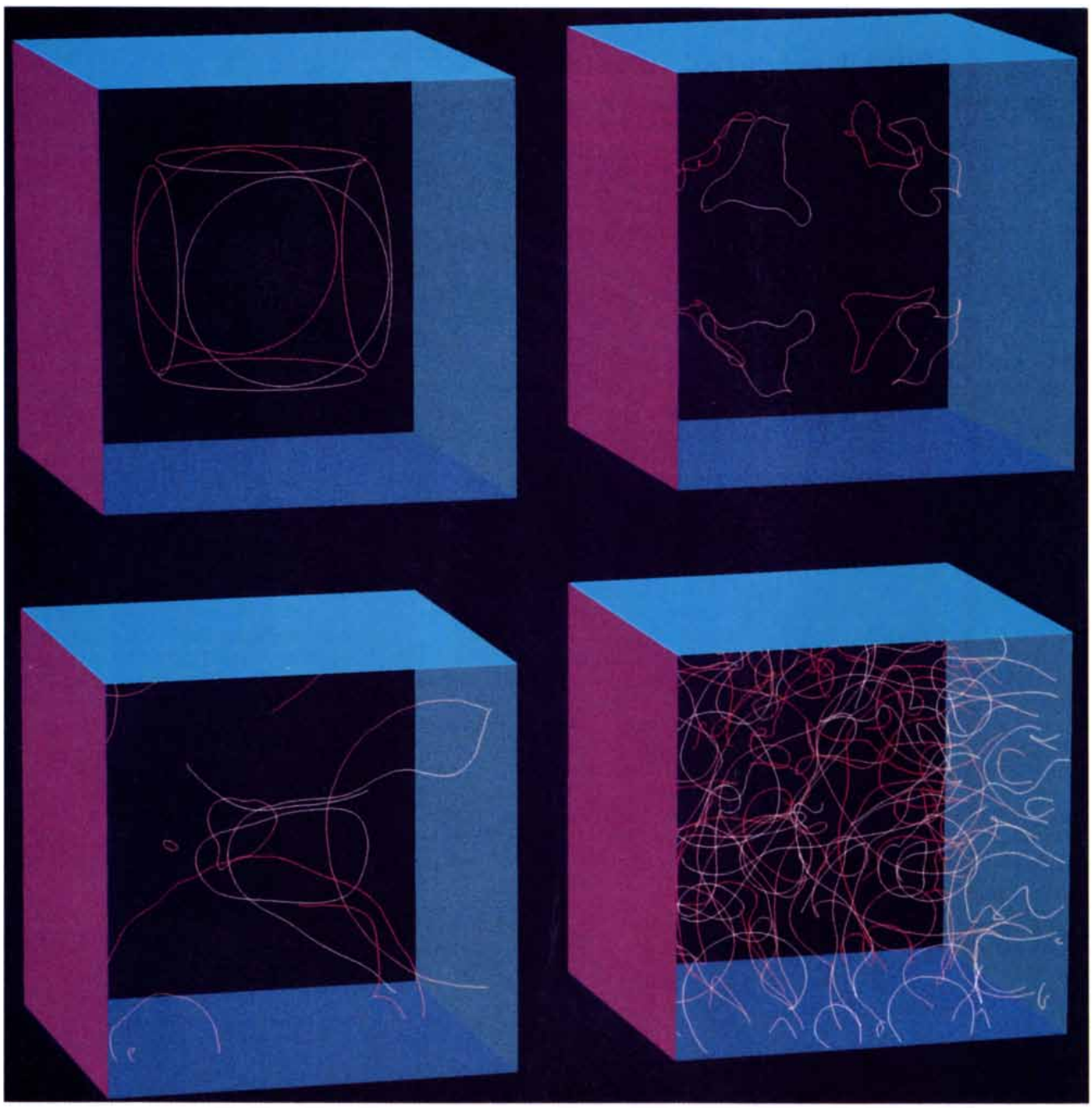
Because of certain properties of the helium nucleus, at absolute zero many atoms in a sample of liquid helium will fall into the state of lowest possible energy. The next-higher energy state has somewhat more energy than the lowest state, and so only relatively high-energy disturbances can drive the fluid up into that state. Low-energy disturbances, such as those generally caused by friction, cannot act on the fluid, and so it is left entirely undisturbed by friction. Hence a sample of

RUSSELL J. DONNELLY is professor of physics at the University of Oregon. He was born in Hamilton, Ontario, and was educated at McMaster University and Yale University, where he got his Ph.D. in 1956. He taught at the University of Chicago from 1956 until 1966, when he moved to Oregon, where he has been ever since. In 1972 he held a visiting professorship at the Niels Bohr Institute in Copenhagen. Donnelly is currently associate editor of *Physical Review A*, and he has served for many years on the board of the Oregon Bach Festival.

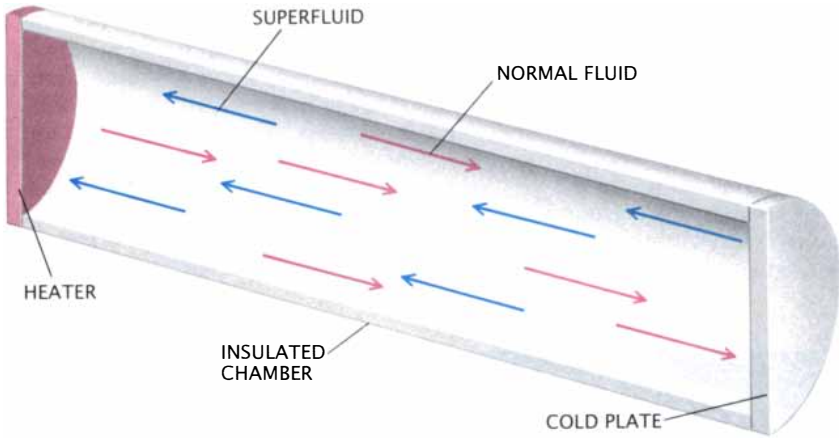
superfluid helium as a whole flows without friction.

This picture does not quite tell the whole story. At any temperature above absolute zero the liquid will be exposed to random excitations (such as vibrations) due to heat. Some of the excitations will be energetic enough to bump helium atoms out of the lowest state. It turns out that such excitations can "flow" through the superfluid; for example, a vibration can pass through the liquid from atom to atom. Indeed, the excitations behave in many ways like atoms of a normal fluid: they can "collide" with the walls of the vessel and with one another.

The quantum-mechanical description of the resulting situation is complex, but it turns out that one way to describe the transport of these excitations is to imagine liquid helium at a temperature above absolute zero as actually being a mixture of two interpenetrating fluids: a superfluid and a normal fluid. In a certain sense the superfluid component is effectively at absolute zero, and any heat in the mixture is carried by the normal fluid. The relative concentrations of super-



ONSET OF TURBULENCE in superfluid helium (a form of liquid helium that flows without viscosity or friction) is modeled in this computer-aided simulation created by Klaus W. Schwarz of the IBM Corporation's Thomas J. Watson Research Center in Yorktown Heights, N.Y. The thin red lines represent the cores of quantized vortices: vortices around which superfluid helium circulates at precisely determined speeds. At the beginning of the simulation (*top left*) the vortex cores are arranged in an orderly manner; each core is bent into a simple ring and the rings are positioned symmetrically. As the simulation evolves (*top right and bottom left*), the vortex cores twist and bend until finally they are snarled in a complex tangle (*bottom right*).



TWO-FLUID MODEL explains many properties of superfluid helium. According to this model, a sample of superfluid helium is made up of two interpenetrating fluids: a superfluid (*blue*), which flows without friction and, in one sense, has a temperature of absolute zero, and a normal fluid (*red*), which flows with normal friction and carries all the heat in the sample. A heater at one end of a channel of superfluid helium causes a counterflow: normal fluid "created" at the heater flows toward the other end of the channel while superfluid flows back in the opposite direction.

fluid and normal fluid depend on the temperature; at absolute zero the liquid is entirely superfluid, and just above 2.172 degrees it is entirely normal fluid. The two fluids can flow past each other with no friction because of the special nature of the superfluid.

The two-fluid model makes it possible to understand why superfluids conduct heat so well. Suppose an electric heater is placed at one end of a channel full of helium II and a cold plate is placed at the other end. Normal fluid is then "created" at the heater, and it flows toward the cold plate. At the same time superfluid flows in the opposite direction, toward the heater, exactly balancing the flow of normal fluid. The velocities of the two flows depend on the difference in temperature between the heater and the cold plate as well as on the relative concentrations of normal fluid and superfluid.

The normal fluid exerts drag on the walls of the channel as it carries heat away from the heater, but the superfluid (which carries no heat at all) slips freely along the channel walls; the two fluids flow past each other without friction. Thus the enormous heat conductivity of helium II has its origin not in the usual mechanisms of atomic motion, in which heat is transferred by collisions among atoms, but in a vigorous counterflow of normal fluid in one direction and superfluid in the other. The heat flows directly and swiftly from one end of the channel to the other, rather than traveling by the slower process of diffusion.

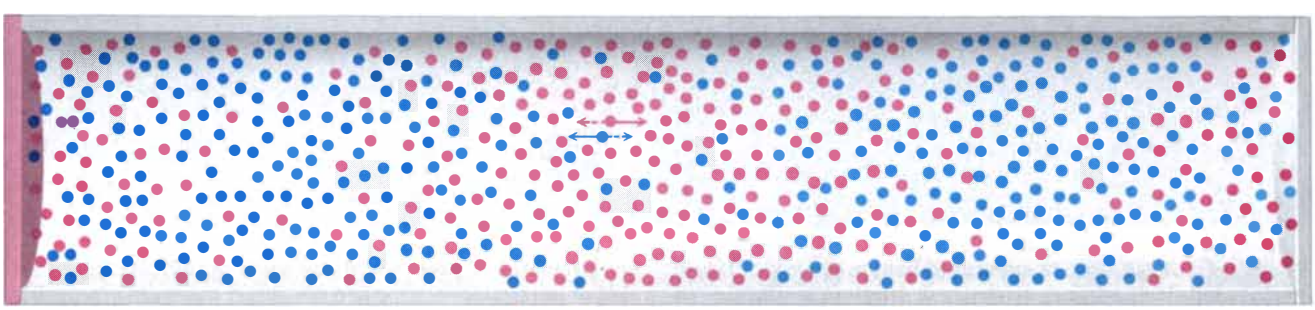
The two-fluid model raises a key question: How does this complex fluid behave under other conditions, such as rotation? One might think that if helium II were poured into a spinning bucket, the normal fluid would spin with the walls of the bucket because of friction while the superfluid remained at rest. That is indeed true for *very* slow rates of rotation—for a bucket having a radius of one centimeter, the rotation must be slower than .03 revolution per minute, or about 40 revolutions per day. At higher rates of rotation, however, the two fluids appear to rotate together.

In the early 1950's Lars Onsager of Yale University and Richard P. Feynman of the California Institute of Technology proposed, on largely theoretical grounds, that the superfluid component rotates in a much more complicated way than normal fluids do. They suggested that a number of "quantized vortices" appear in the spinning bucket.

Quantized vortices are the elements of superfluid turbulence, just as whirlpools and eddies are the elements of classical turbulence. They are remarkable objects. The core of each vortex is about one angstrom unit (10^{-10} meter) in diameter, or roughly a fourth of the average distance between atoms of the fluid, and it contains no atoms of superfluid (just as the core of the vortex that forms at a bathtub drain contains no water). In the terms of quantum mechanics, the core is a node in the "wave function" that describes the superfluid: the probability of finding a helium atom there is zero.

The way atoms circulate about the core of a vortex is also determined by quantum mechanics. According to quantum mechanics, every superfluid atom can be viewed in some sense as a wave. The atom's wavelength depends on the rate of its circular motion within the vortex, and a whole number of wavelengths must fit into the path the atom travels around the vortex core. As a result the rate of the atom's motion is quantized: an atom traveling at a certain radius from the center of the core can move only at one of a precise set of speeds.

In general the helium atoms tend to travel around the vortex core at the lowest possible speed. Then each atom's tangential velocity is equal



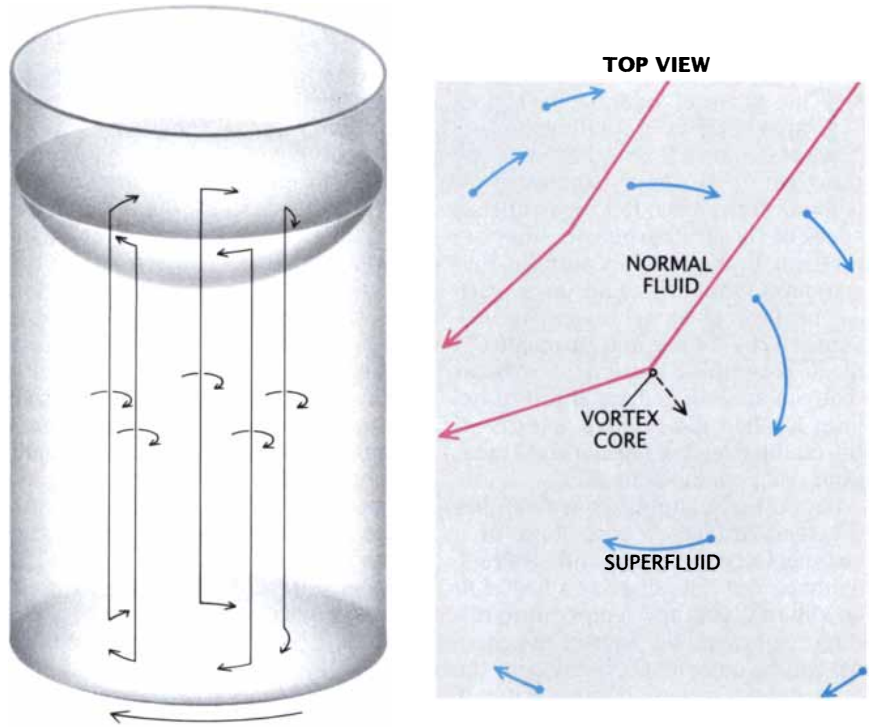
"SECOND SOUND," a phenomenon analogous in certain ways to ordinary sound, arises in a channel of superfluid helium if a heater at one end is switched on and off rapidly. In a "wave" of second sound, regions having a high concentration of normal fluid alternate with regions having a high concentration of superfluid. The fluids oscillate rapidly in opposite directions.

to Planck's constant (the fundamental constant of quantum mechanics) divided by the radius of the atom's circular motion, divided by the mass of the atom, divided by 2π. The formula by which this velocity was derived is nearly identical with the formula used by Niels Bohr to determine the characteristics of an electron's orbit about an atomic nucleus.

Vortex cores run from the upper surface of the liquid to the bottom of the bucket. If a number of vortex cores form in a single spinning bucket, the entire array of vortex cores rotates at precisely the same angular velocity as the sides of the bucket.

The normal fluid does not circulate around vortex cores, but it is affected by them. If a flow of normal fluid passes near a vortex core, the flow can be scattered or redirected. Consequently vortex cores interfere with the transmission of heat through the superfluid. In turn the flow of normal fluid can have an effect on the vortex cores: when a vortex core redirects the flow of normal fluid, the core is nudged aside and bent slightly.

It took many years to establish experimentally the correctness of Onsager and Feynman's ideas. That is not too surprising when one considers that the total volume occupied by vortex cores in a bucket spinning at 10 r.p.m. is about a trillionth of the volume of the bucket.



QUANTIZED VORTICES form when a vessel containing superfluid helium is put on a rotating turntable (*left*). Superfluid circulates around the vortices, and the vortices themselves move with the spinning vessel. The rate at which a superfluid particle circulates around a vortex core (*right*) is proportional to the reciprocal of the distance between the particle and the core: particles closer to the core travel faster. The core contains no superfluid. When normal fluid (*red*) flows near a core, the flow is deflected and the core is nudged aside by a process called mutual friction. Quantized vortices thus disrupt the flow of heat (and second sound) through the vessel.

Indeed, how can such tiny objects be detected at all? One method involves a wavelike phenomenon called second sound. Second sound is best understood by an analogy to ordinary sound.

An ordinary loudspeaker essentially consists of a plate that can be made to vibrate. As the plate vibrates, it pushes air back and forth, creating alternating regions of dense and rarefied air—a density wave. Such a density wave is a sound wave.

Now suppose a plate pierced with very fine holes is placed at one end of a channel containing helium II. If the holes are fine enough, the normal-fluid component of helium II will not be able to pass through them. The superfluid component will be able to pass through, however, because of its lack of viscosity. If the plate is vibrated, it will push the normal fluid back and forth while exerting no direct force on the superfluid. That will create regions in which there is a high density of normal fluid, alternating with regions in which there is a low density of normal fluid: a second-sound wave.

When a second-sound wave is transmitted through a channel, the superfluid component is induced to flow back and forth in a way that exactly counters the back-and-forth flow of the normal fluid; the two components flow "through" each other in opposite directions. Hence the overall density of the liquid in the channel does not vary (in contrast to the medium of an ordinary sound wave). Thus a second-sound wave is actually a wave of relative concentration. Regions in which the concentration of normal fluid is high and the concentration of superfluid is low alternate with regions in which the concentration of superfluid is high and the concentration of normal fluid is low.

Another way of looking at it is to say that a second-sound wave is a temperature wave, in which regions of relatively high temperature (regions containing a high concentration of normal fluid) alternate with regions of low temperature (regions containing a high concentration of superfluid). As a matter of fact, another method of generating second sound is based on just such a viewpoint: second sound is generated by turning a heater on and off rapidly. The heater periodically starts and stops a heat-driven counterflow of normal and superfluid helium. At high frequencies the oscillating counterflow becomes a wave of second sound.

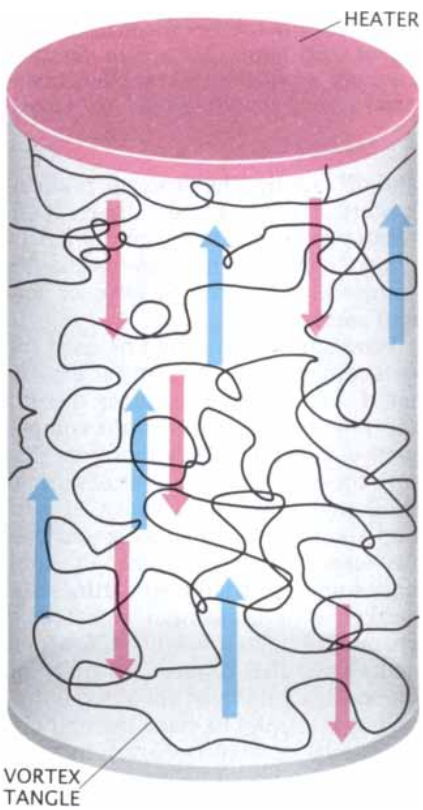
How can second sound be used for detecting quantized vortices? Bear in mind that quantized vortices disrupt the flow of the normal-fluid component of helium II. They therefore disturb the oscillatory motions of normal fluid that make up a wave of second sound. Hence, just as sound baffles can attenuate a wave of ordinary sound, so quantized vortices can weaken a second-sound signal in a channel of helium II. William F. Vinen and Henry Hall of the University of Cambridge employed the attenuation of second sound to make the first observations of a vortex array in a rotating container of helium II in 1956.

The sensitivity of second-sound-attenuation observations is extraordinary. In our laboratory at the University of Oregon we can detect changes as small as 20 centimeters in the length of vortex lines in a cubic centimeter of helium II. In such an experiment the total volume of the additional vortex

line is less than one part in 10^{14} of the total volume of liquid helium.

The array of quantized vortices that appears in rotating buckets of helium II is not what I mean by the term "superfluid turbulence"—it is too orderly. When fluid-dynamicists speak of classical turbulence, they are referring to the complex patterns that can develop in swift fluid flows, such as the flow of air in a wind tunnel. Similarly, by "superfluid turbulence" I mean a complex pattern of vortices that can arise in a channel full of helium II when the relative velocity of the counterflowing normal and superfluid components is high.

For example, suppose one monitors the temperature at both ends of a channel full of helium II while gradually increasing the power of a heater at one end. At first the temperature difference between the ends of the channel will be quite small, because of the high thermal conductivity of helium II. At some critical heater power, however, the temperature difference will climb rapidly. In other words, when a certain amount of heat is being transported down the channel—when the counterflow velocity has reached a certain critical point—the fluid's thermal conductivity decreases.



SUPERFLUID TURBULENCE, a tangle of quantized vortices, is generated when the heater at one end of a channel forces more than a certain critical amount of heat through the channel. The counterflow of normal fluid (*red*) and superfluid (*blue*) bends and twists the vortices.

Measurements of the attenuation of second sound have shown that this sharp decrease in thermal conductivity is due to the formation of quantized vortices, which interfere with the counterflow process. The absorption of second sound is almost the same in all directions, indicating that these vortices, unlike those that arise in a spinning bucket, are not all aligned in a single direction. Instead it seems that the vortex cores are twisted, bent and bunched together into a complex tangle; some of them form complete loops and others run in long, twisted paths from one wall of the channel to another. Thus, just as the rapid flow of water in a brook can lead to the formation of irregular whirlpools and eddies, so the rapid counterflow of normal and superfluid components can cause the formation of an irregular cluster of quantized vortex cores. Feynman had actually anticipated this effect on theoretical grounds before Vinen had published his first observations of quantized vortices.

The complex tangle of vortices that is generated by rapid counterflow is what I call superfluid turbulence; it is also referred to as quantum turbulence. Superfluid turbulence is one of the most surprising phenomena associated with superfluidity. After all, on the most fundamental level a vortex core represents a place where there is no superfluid. Yet the formation and motion of this entity, which is nearly without mass and occupies a tiny fraction of the volume in the channel, have a profound effect on the conduction of heat through helium II.

Superfluid turbulence is studied by observing temperature differences and the attenuation of second sound, and also by means of a technique known as ion trapping. In ion trapping a stream of helium ions (electrically charged atoms) is passed through a channel containing helium II. Some of the ions are sucked into the vortex cores and trapped there, because the pressure is lowest at the center of a vortex. It is then possible to observe the positions of trapped ions, follow their movements and even manipulate them (and hence the vortex cores in which they are trapped) by applying external electric fields.

None of those methods of observation is entirely satisfactory. It is not possible today to observe the location of trapped ions in turbulent flows, and measurements of temperature differences and the attenuation of ion beams and second sound are made on scales that are large compared with the average distance between vortices. The measurements give average properties of the vortex tangle, such as its density (the length of vortex core per unit of volume). In classical turbulence, in contrast, local velocities and pressure fluctuations can often be measured in a single eddy.

How, then, do we know that superfluid turbulence really is turbulence? We can answer that question by comparing some of the characteristics of classical and quantum turbulence.

Most obviously, classical turbulent flows are irregular and so are the vortex tangles of quantum turbulence. In classical turbulence local eddies and whirlpools develop, move about and disappear in an apparently random manner. Similarly, in superfluid turbulence it seems that the vortex cores twist, bend and turn in unpredictable ways. Classical turbulence is highly dissipative of energy. In superfluid turbulence the superfluid component itself, which is the component that circulates about a quantized vortex, cannot dissipate energy because it has no viscosity, but friction between the vortex cores and the normal-fluid component is indeed highly dissipative: a turbulent counterflow generates heat by viscous dissipation as the orderly motion of the normal fluid is disrupted.

There are also great differences between classical turbulence and superfluid turbulence. Superfluid turbulence seems to be produced locally everywhere in the flow, and it is truly homogeneous. Classical turbulence, on the other hand, is often produced when fluid flows past an obstacle, and the turbulent motion decays downstream of the obstacle. In classical turbulence it often happens that a vortex will elongate, or stretch, while the diameter of its core narrows. Individual quantized vortices can stretch, but they cannot reduce the diameter of their cores, because the core size is already at a quantum-mechanical minimum. On the other hand, bundles of vortices can "narrow" by crowding together as they elongate. In classical turbulence the sizes of vortices can range continuously from the size of the channel to dimensions so small that the vortices are eliminated by internal friction. Quantized vortices can form structures—tangles—that are as large as the channel itself or as

small as atomic dimensions, but on intermediate scales the strictly quantized rate of circulation about each vortex core places certain limits on the size and structure of tangles.

Superfluid turbulence is surprisingly easy to study. It is easy to produce and compact in scale, and it is produced in a fluid whose properties are essentially simple and accurately known. It is possible to produce and study it in channels ranging in diameter from one micron (a millionth of a meter) to a few centimeters. Other key properties can be varied over a broad range as well. The length of vortex per unit of volume in the channel can range over five orders of magnitude (that is, it can vary by a factor of as much as 100,000), and the ratio of the concentration of normal fluid to the concentration of superfluid can vary over more than two orders of magnitude (a factor of 100).

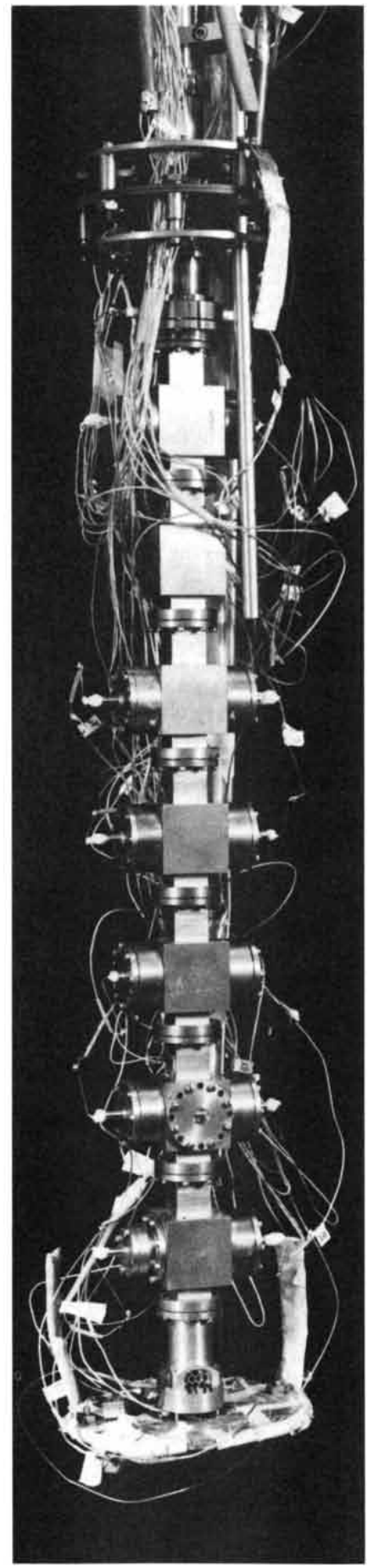
The normal fluid in most quantized-turbulence experiments seems to flow either without turbulence or with only very weak turbulence, and so it does not complicate the experimental picture. Also, the laws governing the motions of vortex cores are well understood on a fundamental level [*see illustration on page 108*]. In many ways vortex cores behave as if they were extremely thin strings bearing substantial tension. (In certain trapped-ion experiments it has even proved possible to "pluck" a vortex core like a guitar string, transmitting vibrational waves up and down its length.) The flows of normal fluid and superfluid around the cores can act against this internal tension, causing the cores to move, bend and twist.

William Vinen—one of the first to detect quantized vortices—also made the first serious study of superfluid turbulence. Using the attenuation of second sound as a probe, he found that a certain critical counterflow velocity (the relative velocity of normal fluid and superfluid) is necessary to generate an observable quantity of tangled vortex line. Vinen discovered a number of the other basic relations underlying superfluid turbulence. For example, when the heater flux (the power per unit of area emitted by the heater in a counterflow channel) is well above the value at which turbulence sets in, the density of vortex lines varies as the square of the counterflow velocity, and the temperature difference between the two ends of the channel varies as the cube of the heater flux. Vinen developed equations to describe the growth and decay of superfluid turbulence, and he proposed hypothetical mechanisms to explain some of the relations he had discovered.

Since then investigators have gained a much more detailed knowledge of the characteristics of superfluid turbulence. For example, it has been found that the density of vortex lines can vary depending on the direction in which it is measured. In 1983 my doctoral students Carlo Barenghi and Charles E. Swanson and I simultaneously transmitted second sound along the length of a channel and across its width and measured the degree to which each signal was attenuated; we also measured the overall temperature difference between the two ends of the channel. We showed that such measurements could enable investigators to explore the anisotropy of the tangle (the degree to which it appears different when it is looked at from different directions) and to determine whether the vortex tangle as a whole is drifting along the channel.

Precise results came between 1984 and 1986 from the thesis work of Swanson and Rabi Wang. They found that the vortex tangle seems denser when viewed along the channel than when viewed across it. At the lowest temperatures the tangle is nearly isotropic, and at the highest temperatures, near the point of transition to helium I, the tangle is nearly twice as dense along the channel as across it. In related work Wang has shown that the tangle as a whole drifts toward the heater at the same speed as the superfluid component of the liquid.

Investigators at various laboratories have also studied how the characteristics of the turbulence depend on the channel's shape and size. These studies have resolved a



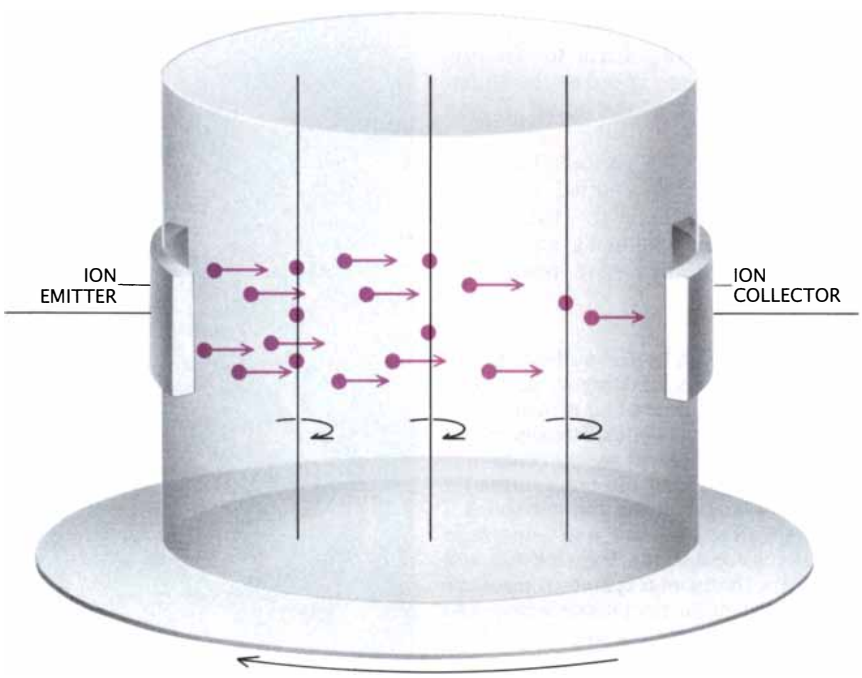
APPARATUS built by the author's group at the University of Oregon generates superfluid turbulence and measures the density of vortex tangles. The superfluid channel itself (*central square column*) is 38 centimeters long and one centimeter across. A heater at the top generates a counterflow of normal and superfluid helium. Second-sound transmitters and receivers (*horizontal cylinders*) measure the density of vortex tangles across the channel by gauging the degree to which the passage of second sound is obstructed by vortices. Another transmitter at the bottom probes the longitudinal density of vortex tangles by sending a second-sound signal up the length of the channel to the top, where it is reflected back to a receiver at the bottom.

good deal of confusion that had arisen among those within the field: for many years the results of experiments done at different laboratories seemed to contradict one another, and the contradictions only became more confusing as more data were gathered. Two relatively recent developments—a classification scheme for turbulent states and refinements of a concept known as scaling—have finally clarified the situation.

In 1982 James T. Tough of Ohio State University, in a comprehensive review of past research in superfluid turbulence, proposed that differences in channel geometry result in several kinds of turbulent state; he introduced a way to classify the different kinds of turbulence. The essential difference between turbulent states involves the way in which the amount of heat transmitted down the channel per unit of time varies as the temperature difference between the ends of the channel is changed. Tough observed that one feature in particular of a channel's cross-sectional geometry is crucial in determining what kind of turbulence will be generated in the channel: the aspect ratio, which is the ratio of the width of the channel to its height. In channels that are essentially square or circular (channels whose aspect ratios are close to unity) there are two possible kinds of turbulent state. In rectangular channels with large aspect ratios (that is, ratios of about 10 : 1) there is only one kind of turbulent state. Tough elucidated the various characteristics of each kind of state.

The other recent advance in the general understanding of turbulent superfluid flows concerns scaling, which is essentially a means of comparing the results of experiments carried out in systems that have the same geometry but are radically different in size. Scaling makes it possible to run tests on small laboratory models that will be relevant when the systems being investigated have been scaled up to industrial size. In classical fluid mechanics two pipes that are geometrically identical will display the same flow characteristics if they have the same Reynolds number: the flow velocity multiplied by the diameter of the pipe divided by the kinematic viscosity (the viscosity divided by the density). Classical scaling is applicable even when the flows are turbulent.

Attempts to apply classical scaling techniques to experiments involving quantized turbulence have not been particularly successful, and so Swanson and I, following a suggestion by Klaus W. Schwarz, now at the IBM Corporation's Thomas J. Watson Research Center in Yorktown Heights, N.Y., began in 1984 to ask ourselves what key facts are different in quantum and classical turbulence. The fundamental problem, we have decided, is that at a given temperature and pressure a quantized vortex must always have the same core radius, regardless of the scale of the flow: the core size is set by the laws of quantum mechanics, and it is no smaller in a smaller pipe. We have therefore devised a quantum analogue of the Reynolds number. A key term in our quantum-Reynolds number involves the logarithm of the ratio of the mean distance between vortices to the size of the vortex core, and another term involves the ratio of Planck's constant to the mass of the helium atom. The new scaling factor has so far proved highly successful.



ION-TRAPPING TECHNIQUE measures the density of vortices in a sample of superfluid helium. Helium ions (electrically charged atoms) are emitted by a radioactive source on one side of the sample (*left*) and collected on the other side (*right*). Along the way some of the ions are trapped in the center of vortex cores. The number of ions that are trapped by vortex cores indicates the density of vortices in the sample.

The compact scale of quantum turbulence allows it to serve as a model system for studying aspects of classical turbulence. For example, turbulent flows in rotating systems are of considerable geophysical interest, but there have so far been very few experiments involving rotating turbulent flows, perhaps because of the size and expense of the necessary wind-tunnel apparatus. Superfluid turbulence, in contrast, can readily be studied on a rotating turntable.

The influence of rotation on turbulent flow is complicated and rather forbidding. It is not an area in which intuition is particularly useful, since we live in an environment that is locally almost at rest. Yet Swanson, Barenghi and I have found that the properties of rotating quantum turbulence are surprisingly simple in the limits of large heat flux and slow rotation or small heat flux and fast rotation.

For example, in the limit of large heat flux and slow rotation, we have found that the total density of vortex lines is less than one would predict by adding the regular array of lines that should appear because of rotation (as in a spinning-bucket experiment) to the irregular tangle that should arise through turbulence. We have concluded that some lines are, in effect, doing double duty: some of the heat-induced vortex lines must be orienting preferentially in a direction parallel to the axis of rotation.

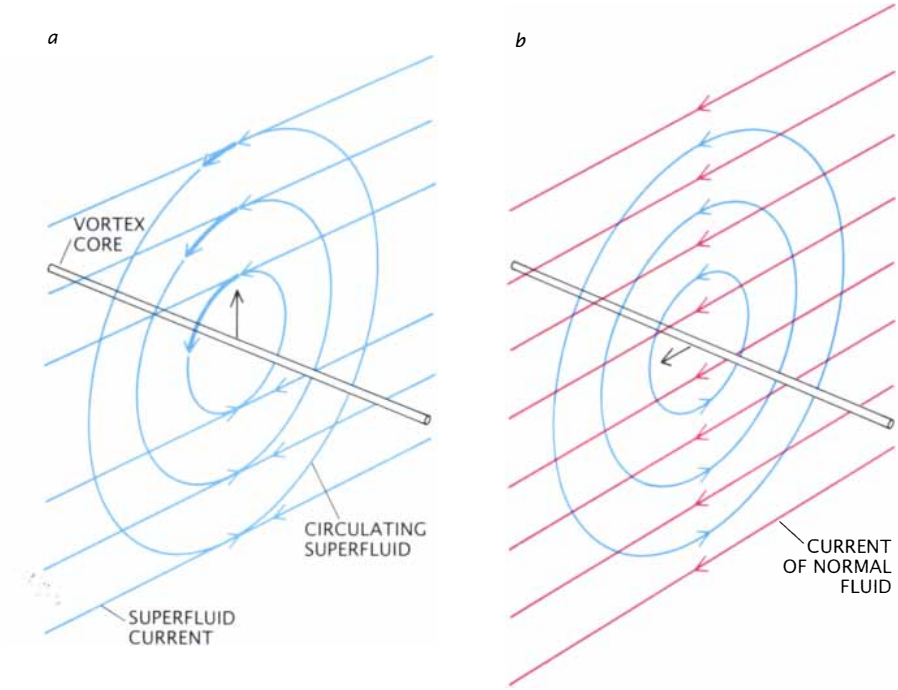
The alignment can be increased by increasing the rate of rotation, much as a gas of magnetic dipoles can be aligned by increasing the strength of an external magnetic field. Indeed, it appears that the two situations are mathematically quite similar: the equation relating the alignment of the vortices to the rate of rotation is nearly identical with the equation that relates the magnetic polarization of the

gas of dipoles to the strength of the external magnetic field. The analogy can be carried further. The tangling of the vortices due to heat flux tends to disorder their alignment in much the same way as thermal agitation due to an increase in temperature disorders a gas of magnetic dipoles. Thus the study of quantized turbulence in a rotating system has unexpectedly provided physicists with one of the few known physical analogues of the disorder that is caused by thermal agitation.

Another productive area of research is the use of powerful, fast computers to simulate the behavior of vortices numerically. Because it is impossible to observe the motions of individual vortices directly, computer simulations are currently the only method available for investigating superfluid turbulence on such a small scale. One objective of such simulations is to determine how a hypothetical vortex tangle will develop and move by directly solving certain basic equations of motion. The forces exerted on vortices by the flows of normal fluid and superfluid are well understood, and the extremely small radius of the core makes it possible to model a vortex as a thin filament.

Computer simulations of classical turbulence must model an entire fluid as a three-dimensional set of points moving through space. In the case of quantum turbulence, assuming the flow of the normal fluid itself is not turbulent, the situation is much simpler. One needs to follow only a one-dimensional set of points—the vortex filament—as it moves in three-dimensional space. Nevertheless, so far only Schwarz has carried out significant numerical simulations of quantum turbulence. Over the past eight years he has made good progress in simulating several special cases.

One case involves the homogeneous turbulence sometimes produced at temperatures above one degree. Schwarz has developed a set of rules that allows a rapid calculation of the properties of the tangle, and he has found that in many cases one can disregard the complicating features of quantum mechanics and consider instead only the laws of classical hydrodynamics. When the counterflow velocity is reduced below a certain value in these simulations, the internal tension of the vortex filaments pulls the cores into straight lines, which move to the side of the channel. Perhaps this mechanism underlies the experimental observation that turbulence disap-



MOTION OF VORTICES is influenced by a number of relatively well-understood forces. When a current of superfluid flows past a vortex core (*a*), it accelerates the circulation on one side of the core and slows the circulation on the other side. According to a fundamental principle of fluid dynamics, pressure is lower in a swift flow than in a slow one, and so a lift force (like the force on a moving airplane wing) pushes the vortex core in a direction perpendicular to the superfluid current. When a current of normal fluid flows past a vortex (*b*), it exerts a drag force that pushes the vortex core in a direction parallel to the current. A third principle of vortex motion (*c*) is not as well understood. When two vortex lines are brought close together, they can break apart and reconnect so that the broken ends of each line join the appropriate broken ends of the other.

pears rapidly when the heat flux in a channel is reduced below a critical value. Other recent results suggest that interactions between vortices and microscopic roughnesses on the walls of the channel may also underlie the appearance and disappearance of quantized turbulence at critical counterflow velocities. Such surface roughnesses are known to play a large part in the appearance of turbulence in classical systems.

In certain parts of these simulations, Schwarz has found that it is not possible to rely solely on the laws of classical hydrodynamics. When two vortex lines approach each other very closely, for example, classical hydrodynamics breaks down and quantum mechanics must enter the picture. For this case Schwarz has adopted a rule proposed many years ago by Feynman: each line splits in two at the point of closest approach and the lines reconnect, but in a new way—the ends of each line join to the appropriate ends of the other [*see illustration above*]. When such reconnections are included in the simulations, almost any initial distribution of vortices can evolve quickly into a complex tangle.

What physical mechanism could underlie reconnection? In classical hydrodynamics it is not possible for a vortex simply to break apart in the middle. An initial clue may have been provided by the recent work of Christopher Jones of the University of Newcastle upon Tyne and Paul Roberts of the University of California at Los Angeles. Jones and Roberts have examined the mathematical description of a vortex ring—a vortex line pulled into a simple loop—that shrinks until it is almost the size of a vortex core. They have found that at a certain definite size the quantized circulation about the vortex line disappears completely. Perhaps a similar sequence of events takes place when two sections of different vortex lines come close enough: perhaps for quantum-mechanical reasons the circulation stops in a limited area, enabling the lines to reconnect without any violation of the laws of hydrodynamics.

Looking to the future, one can foresee steady progress in the study of superfluid turbulence. Although there is no experimental probe that can determine the properties of flow near a single vortex core, it now seems possible to simulate the flow numerically in a reliable way. Some questions—what is the mechanism underlying reconnection, for example—can still be approached only on theoretical grounds, but there is progress on those fronts as well.

If all goes reasonably well, I believe we shall be in a position before too long to say that superfluid turbulence is well on the way to being genuinely understood—by which I mean that the results of an experiment that has not yet been carried out may soon be predictable by calculations (and perhaps partly by simulations) relying only on the fundamental laws of vortex dynamics. Superfluid turbulence may become the first topic in the study of turbulent flows to be approached with such a level of insight.

FURTHER READING

MUTUAL FRICTION IN A HEAT CURRENT IN LIQUID HELIUM II. W. F. Vinen in *Proceedings of the Royal Society of London,* Vol. A242, pages 493-515; 1957.

GENERATION OF SUPERFLUID TURBULENCE DEDUCED FROM SIMPLE DYNAMICAL RULES. K. W. Schwarz in *Physical Review Letters,* Vol. 49, No. 4, pages 283-285; July 26, 1982.

QUANTUM TURBULENCE. Russell J. Donnelly and Charles E. Swanson in *Journal of Fluid Mechanics,* Vol. 173, pages 387-429; 1986.

SUPERFLUID TURBULENCE. J. T. Tough in *Progress in Low Temperature Physics,* Vol. 8, pages 133-216; 1987.

THREE-DIMENSIONAL VORTEX DYNAMICS IN SUPERFLUID ^{4}HE: HOMOGENEOUS SUPERFLUID TURBULENCE. K. W. Schwarz in *Physical Review B,* Vol. 38, No. 4, pages 2398-2417; August 1, 1988.

Ever since the first phone call, we've been listening to the voices of tomorrow.

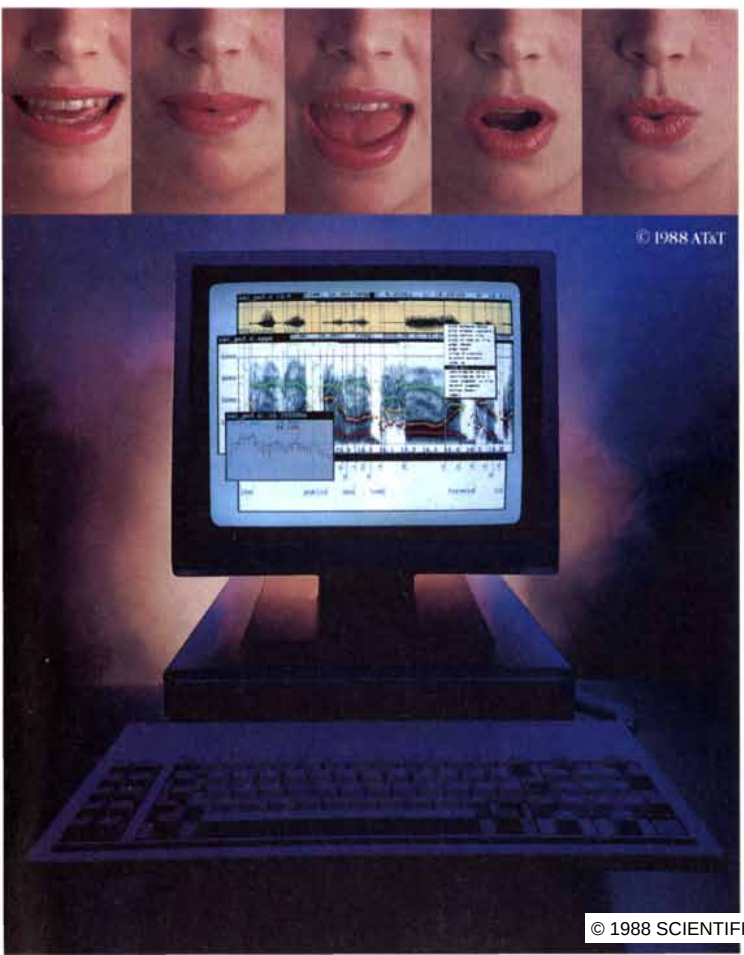
It started with the genius of Alexander Graham Bell. And from the beginning, AT&T has been committed to helping the people of the world communicate better.

We've been the world's most ambitious sound laboratory for more than a century. And the results speak for themselves—75 million times each business day.

Calls placed today over AT&T's Worldwide Intelligent Network are predictably clear. Noise, echo and static free. And our experiments in digital high fidelity transmission point to a future when your AT&T calls will ring truer still.

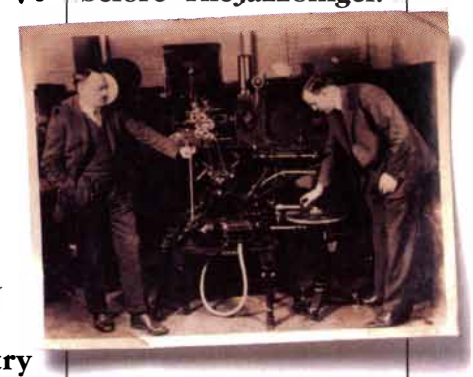
Because we've always believed that by improving the way things sound for the people of the world, we help bring them closer to one voice.

Funny, how the future seems to repeat itself.



1929
Sound Motion Pictures

AT&T brought sound to the motion picture "Don Juan"—the first full-length movie with a synchronized sound track—released a year before "The Jazz Singer."



Today
Cordless Clarity

Because AT&T sound quality is the industry standard, we've worked hard to achieve it in our cordless models. Today's AT&T cordless phones bring you clear sound which equals the quality you've come to expect from our corded models. The perfect combination of convenience and clarity.

Tomorrow
Interactive Synthetic Speech

Technology now exists which enables computers to "speak" words to us in any language that is entered on their keyboards—while machines which respond to verbal commands, act upon them and converse with us by simulating human speech are the already audible voices of tomorrow.

AT&T
The right choice.

Infrared Optical Fibers

New glass and crystalline fibers, promising greater transparency and transmitting longer wavelengths than silica fibers, have been applied to communication systems, medical diagnostics and optical-fiber lasers

by Martin G. Drexhage and Cornelius T. Moynihan

Egyptian glassmakers crafted an opaque glass fish with a blue head, white scales and yellow fins 3,000 years ago. The intentional addition of metallic oxides to silica-based glass—cobalt for blue, tin for white and antimony for yellow—colored the fish. Yet because their raw materials already contained traces of these metallic impurities and others, clear glass eluded the ancient Egyptians. As the science of chemistry grew in sophistication and better glassmaking techniques were developed over the centuries, other artisans succeeded in creating clear glass. As a mark of this success, the best lenses from the early 20th century are 10,000 times as transparent as the first glasses made in ancient Egypt. Since the 1960's the transparency of glass has been increased another 10,000 times by a refinement technique for producing pure silicon dioxide—a technique that made fiber optics possible. Now telephone conversations, computer data and television images are transmitted as light pulses traveling through strands of highly transparent, silica-based glasses.

In 1979 optical fibers made from silica and silica-based glasses reached their limit of transparency. A second generation of fiber materials are now emerging that transmit infrared light at longer wavelengths. Such fibers belong to three classes of materials: the halide-containing crystals, the chalcogenide glasses and the heavy-metal fluoride glasses.

All these infrared optical materials promise greater transparencies than those feasible with silica fibers, and they are being tested for many new applications. Infrared optical fibers can transmit simple infrared images or temperature information about objects in remote locations. The fibers can serve as flexible conduits for energy from lasers that can cut, weld or drill industrial components. The fibers have even served to convey infrared laser energy to unclog plaque deposits in animal arteries, a procedure that may soon be a routine alternative to coronary-bypass surgery and balloon-catheter angioplasty.

The application of infrared optical fibers has sparked much interest in the telecommunications industry as well. If the transparency limit of infrared fibers can be reached, they will have a distinct advantage over other optical fibers that operate at shorter wavelengths. When a light source sends a signal through a fiber to a receiver, part of the signal power from the source dissipates before it reaches the receiver. Indeed, the signal can be so attenuated that it is lost as it falls below the sensitivity limit of the receiver. Signals that travel over great distances need to be amplified by regenerators periodically, and the manufacture, operation and maintenance of such regenerators represent a significant expense. Signals transmitted over commercial silica fibers must be reamplified every 10 to 50 kilometers and signals transmitted over copper wires every four to six kilometers. In communications systems based on infrared optical fibers, on the other hand, the distance between regenerators could be hundreds or perhaps thousands of kilometers.

To attain this goal and to develop new short-distance fiber technologies, investigators have examined numerous glasses and crystals that are transparent in the infrared. The halide-containing crystals, the chalcogenide glasses and the heavy-metal fluoride glasses were chosen as the most practical candidates after careful consideration of the basic structure and interatomic forces that guide and attenuate light through a fiber.

The journey of all light waves through an optical fiber starts at a source, usually a laser or a light-emitting diode. The source is placed close to one end of the fiber. The fiber consists of two concentric cylinders: the core and the cladding. Light from the source illuminates the core. Some of this light will pass straight along the axis of the core, while the remainder will enter at an angle and will eventually hit the core-cladding interface. Light striking the interface at steep angles will breach the interface and cladding and will then be absorbed typically by a plastic coating meant to protect the fiber from mechanical damage. Light striking the interface at shallow angles will be totally reflected back into the core. This total reflection process occurs repeatedly, guiding light rays along the length of the fiber.

The total reflection is the result of the fact that light and all other electromagnetic waves travel at different speeds in different materials. The speed with which light travels through a material is represented by the index of refraction: the ratio of the speed of light in a vacuum (which is the fastest) to the speed of light in the material. Since light travels slower in the core than in the cladding, the core has a higher index of refraction. Whenever light travels from a material of high

MARTIN G. DREXHAGE and CORNELIUS T. MOYNIHAN have collaborated on a number of research projects. Drexhage directs a research group exploring applications of infrared transparent materials at the Galileo Electro-Optics Corporation in Sturbridge, Mass. He was born in Germany and received a Ph.D. in 1977 from the Catholic University of America. From 1977 to 1987 Drexhage worked for the Solid State Sciences Directorate at Hanscom Air Force Base. Moynihan is professor of materials science at the Rensselaer Polytechnic Institute, where he helped to set up the Center for Glass Science Technology. He earned his Ph.D. from Princeton University in 1962. Moynihan taught at California State University at Los Angeles from 1964 to 1969 and then at Catholic University until 1981.

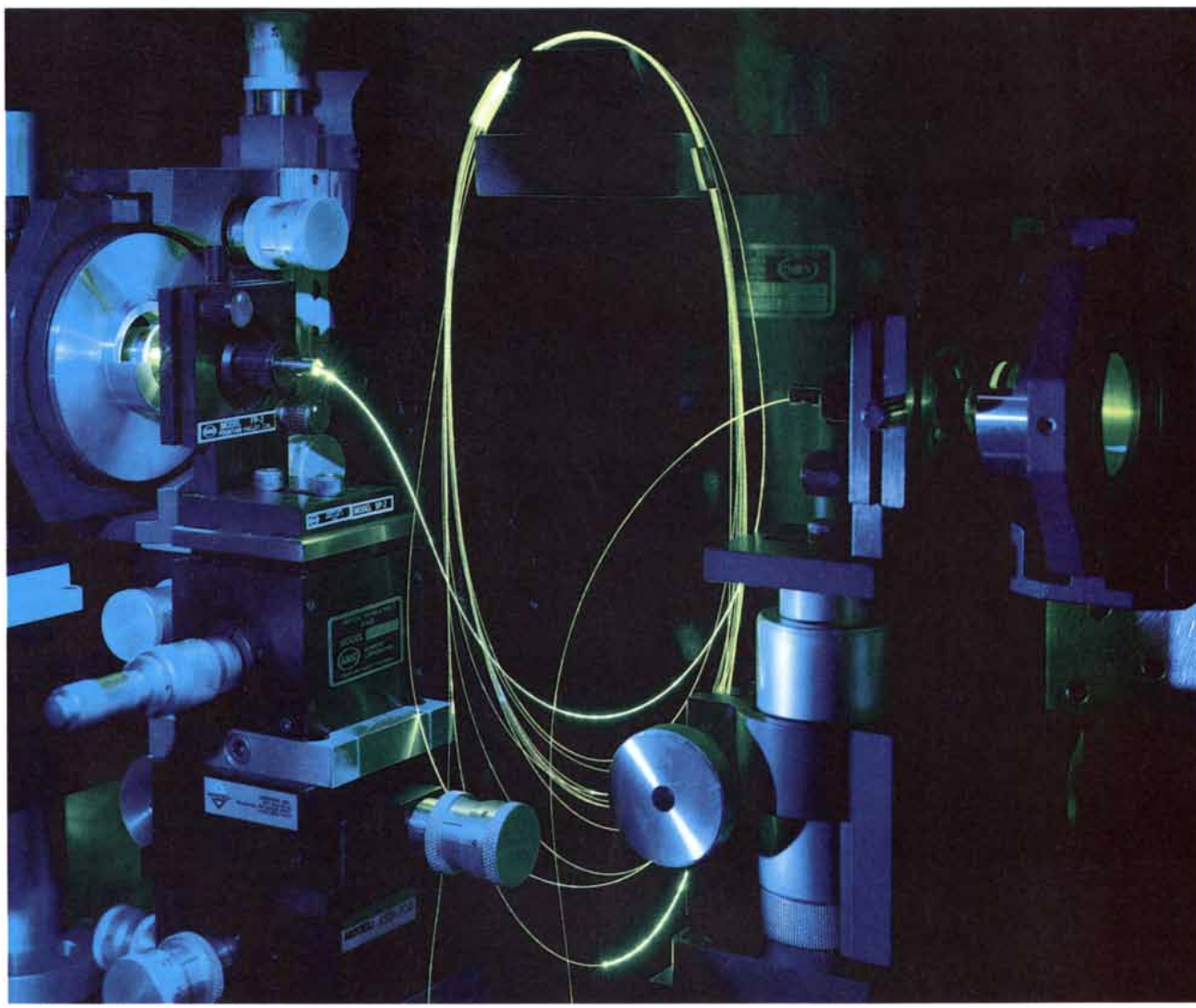
refractive index to one of low refractive index, some is totally reflected and some is transmitted. The fraction of light reflected depends on both the angle at which light strikes and the refractive-index difference between the two substances.

Fine-tuning the difference in refractive index between the core and the cladding and controlling the diameter of the core can enhance the guiding effect so that light signals stay focused and all signal components arrive at the receiver simultaneously [see "Communication by Optical Fiber," by J. S. Cook; SCIENTIFIC AMERICAN, November, 1973]. If all goes well, some of the light from the source emerges at the far end of the fiber core and illuminates a receiver, which is usually a photosensitive detector on a circuit wafer.

Even if light has entered the fiber and is guided along the core, internal scattering and absorption effects can impede propagation and attenuate the signal. The attenuation of light in a transparent solid is the result of three independent processes in the material: electronic absorption, light scattering and vibrational absorption. Although their magnitude may vary with the material, these three phenomena are indigenous to all transparent solids.

Electronic absorption governs attenuation at short wavelengths. A fiber is, after all, composed of atoms joined together by the electrons that form their chemical bond. If a light wave of appropriate energy enters the fiber, it may be absorbed by the bonding electrons and subsequently dissipated as heat. Since short wavelengths correspond to high energies of chemical bonds, short-wavelength light will be absorbed more strongly by this method than long-wavelength light.

Light scattering governs attenuation at longer wavelengths. One common form of intrinsic light scattering is Rayleigh scattering, which occurs in liquids, gases and many solids including glasses. The scattering arises from localized variations in the refractive index of the material caused by changes in density and composition. The intensity of the scattering decreases rapidly with increasing wavelength; its



FLUORIDE OPTICAL-FIBER LASER represents an important application of infrared optical-fiber technology. The device converts laser light from blue-green into infrared. A length of coiled fiber, made from a heavy-metal fluoride glass and doped with neodymium, is placed between two mirrors (*cylindrical structures at left and right*). The blue-green light stimulates the neodymium ions, which then emit infrared light. The infrared light is amplified as it repeatedly traverses the length of the fiber and is reflected from the end mirrors. One mirror, partially transparent, allows some infrared radiation to escape. The device was assembled at GTE Laboratories, Inc., by William J. Miniscalo, Leonard Andrews and Barbara Thompson.

magnitude depends on the material. Experiments carried out on many silica-based glasses suggest that among the more important parameters for determining the attenuation due to Rayleigh scattering are the index of refraction and the glass-transition temperature. This temperature marks the point at which molten glass solidifies, freezing in localized variations in the refractive index. Materials that have low glass-transition temperatures and low indexes of refraction should exhibit low Rayleigh scattering.

Vibrational absorption governs attenuation at the longest wavelengths. It is a complicated function of the effective charges, masses and sizes of the atoms that compose a solid. A bond between atoms in a crystal or a glass can be thought of as an attraction between positively charged ions (cations) and negatively charged ions (anions). Two ions joined by a chemical bond vibrate continuously like two weights connected by a stiff spring. If the weights are displaced by a periodic force that matches their vibrational period, energy will be efficiently transferred to the weights and will increase the amplitude of the vibration. The effect is known as resonance. If the atomic bonds are displaced by light at certain resonant wavelengths, the light energy is transformed into the vibrational energy of the ions. Signal strength declines because energy is absorbed. As the wavelength becomes shorter than the resonant wavelength, vibrational absorption decreases and the material becomes more transparent.

The resonant wavelength is determined by bond strength and ion mass. The resonant wavelength tends to be short when the mass of the ionic pair is small and the interatomic bond is strong. In silicon dioxide, for example, each silicon cation is surrounded by four oxygen anions. The silicon is bonded strongly to the oxygen, and the combined mass of the two ions is light compared with the ions in other infrared optical fibers. Therefore significant absorption occurs at short wavelengths in the infrared region, and silicon dioxide can transmit light only to about 2.5 microns.

In the disordered atomic arrangement that is characteristic of glasses and even in the well-ordered lattice of crystals, the onset of vibrational absorption occurs gradually as the resonant wavelength is approached. This gives rise to what is called the vibrational-absorption edge of the material. In this region the intensity of absorption in many substances is found to decrease exponentially with decreasing wavelength.

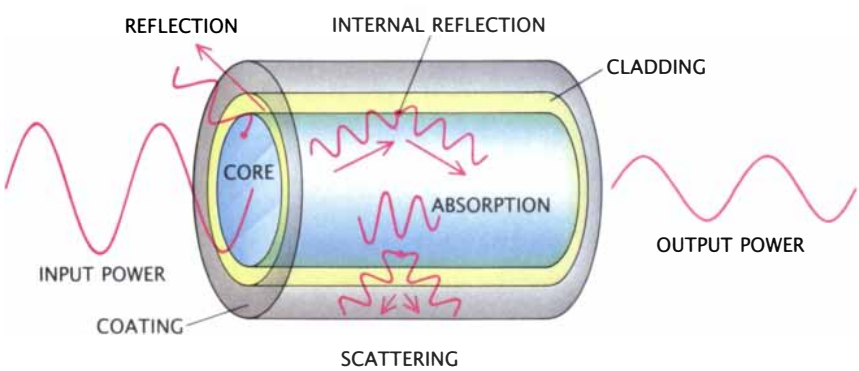
Electronic absorption, Rayleigh scattering and vibrational absorption are all regarded as intrinsic forms of loss because they are inherent in the material. Knowledge of the intrinsic properties allows one to make predictions concerning the maximum transparency attainable in a particular solid. Also to be considered in choosing materials are the extrinsic losses, which are the result of contamination and improper processing. Extrinsic factors include absorption due to impurities in the material, light scattering from large particulates or bubbles, variability in fiber diameter and localized deviations from the desired index of refraction.

The total power loss from intrinsic and extrinsic sources is represented by an attenuation coefficient. The coefficient is defined in decibels per kilometer. If a fiber has an attenuation of one decibel per kilometer, then 10 watts of optical energy injected into a fiber one kilometer long will emerge as 7.9 watts. Ordinary window glass has an attenuation of several thousand decibels per kilometer. Optical communication systems require that losses be held to a level of one decibel per kilometer in a fiber 50 kilometers long; still longer links require reduction of the loss to .01 decibel per kilometer. If windows were produced with that transparency, one could peer through a pane 200 kilometers thick.

In addition to low optical attenuation, other characteristics of a fiber material must be considered. The ideal fiber material should be strong, flexible, simple to fabricate and immune to chemical attack. Unfortunately low attenuation and good structural characteristics tend to be incompatible. Many of the same properties that favor long-wavelength transparency—such as low glass-transition temperature, weak interatomic bonds and anions heavier than oxygen—often produce materials that have undesirable physical, chemical and mechanical behavior. Nevertheless, there exist a number of crystals and glasses with improved long-wavelength transparency for which these tradeoffs are acceptable. Such materials can transmit infrared wavelengths with little Rayleigh scattering and electronic absorption. Therefore their transparency can in theory be enhanced to levels beyond those found in silica fibers.

The lowest attenuation yet achieved in silica fibers is .2 decibel per kilometer at a wavelength of 1.5 microns. This agrees very closely with the intrinsic transparency limit predicted for silica-based glasses on the basis of their Rayleigh scattering, electronic absorption and the vibrational-absorption edge associated with the silicon-oxygen bond. The low attenuation was achieved by careful processing of the glass and fiber so that all sources of extrinsic losses were virtually eliminated.

To obtain transparencies higher than those available in silica-based fibers, new materials—chosen on the basis of their intrinsic attenuation properties—must be refined to minimize extrinsic attenuation factors. These new materials will operate in the infrared beyond two microns.



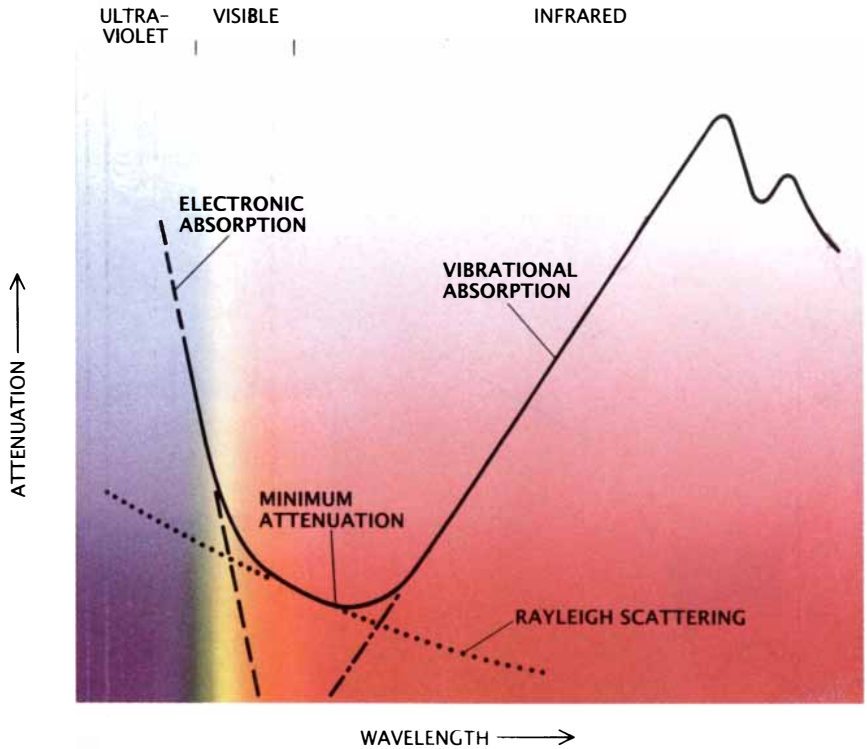
ATTENUATION, or power loss, in an optical fiber is caused by a variety of processes. A fiber consists of a core that conducts light and a cladding that reflects most errant light back into the core. Some light passes into the cladding and is absorbed by the coating; other light is lost on entry because it reflects off the end of the fiber. Within the core, light is absorbed by the atoms of the material and by impurities. Light is also scattered by microscopic fluctuations in the fiber's composition or density.

Crystalline materials form the first class of substances that may be effective as infrared optical fibers. In theory many two-component crystals—silver bromide, zinc selenide, sapphire and even sodium chloride—have low intrinsic attenuations. In practice, however, the use of single crystals as long-length optical fibers presents several problems. The growth rates of single-crystal fibers are quite low, often only a few centimeters per minute. This low growth rate increases the chance that the fiber diameter will vary. The simultaneous fabrication of a true optical-fiber structure, with a high-index core and lower-index cladding, has also proved difficult.

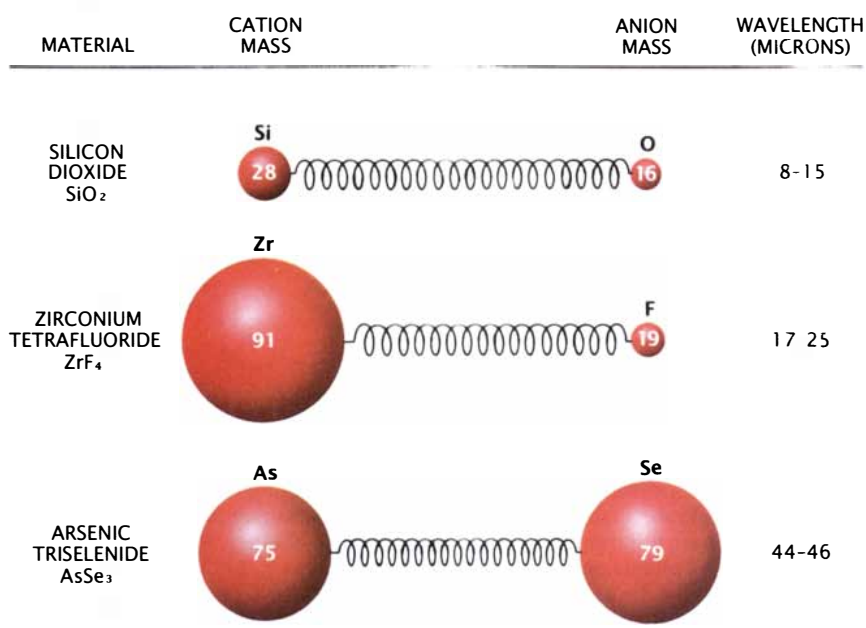
Much more promising are the polycrystalline materials, principally those based on thallium or silver halides. A composite fiber of thallium, bromine and iodine has been studied intensively. To be sure, the high refractive index (about 2.7) of the thallium halide crystal suggests that it might exhibit significant Rayleigh scattering. A vibrational edge in the far-infrared region, however, allows use of light at a wavelength where Rayleigh scattering is less significant. Theoretically thallium halide single crystals could be made that would have an attenuation factor of less than .01 decibel per kilometer near seven microns. Yet in the fabrication process the material develops a granular, polycrystalline microstructure, which scatters light. Such extrinsic scattering effects, coupled with impurity-related absorption, have limited the attenuation of polycrystalline thallium halide fibers to between 150 and 400 decibels per kilometer. This moderate level of attenuation, however, is retained over a broad range of wavelengths extending from about six to 15 microns.

The related polycrystalline material, silver halide, consists of silver, bromine and chlorine. Abraham Katzir of Tel-Aviv University has developed fibers whose attenuations are less than about 1,000 decibels per kilometer beyond six microns. Like the thallium halides, silver halide fibers can transmit the light from a carbon dioxide laser at a wavelength of 10.6 microns. The material is therefore effective for laser power-transmission applications such as laser-assisted surgery.

The chalcogenide glasses, the second class of infrared-transparent fiber materials, are fabricated by combining metals with the heavier elements in the oxygen family, the chalcogens: sulfur, selenium and tellurium. Arsenic trisulfide and arsenic



INTRINSIC ATTENUATION of light results from three mechanisms in fiber materials. Electronic absorption consumes energy when short-wavelength light excites the bonding electrons in atoms. Vibrational absorption occurs when long-wavelength light causes the atoms of a material to vibrate. At intermediate wavelengths, where neither absorption mechanism is significant, fluctuations in density or composition cause Rayleigh scattering. These mechanisms determine the attenuation curve.



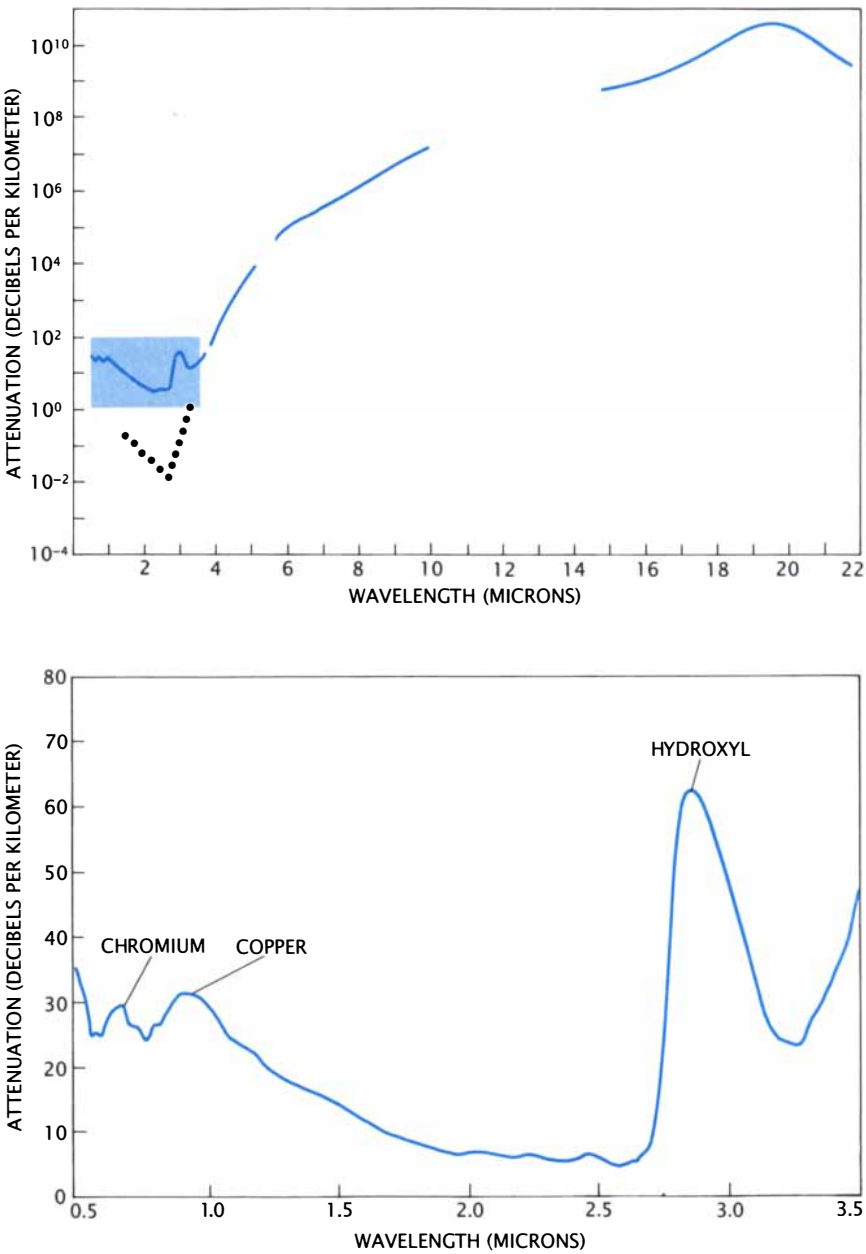
GLASS MATERIALS absorb light through atomic vibrations. Atomic mass and bond strength determine the wavelength at which vibrational absorption is strongest. Large masses and weak bonds between ions increase this wavelength. Consequently zirconium tetrafluoride, a major component of heavy-metal fluoride glass, and arsenic triselenide, a chalcogenide, attenuate signals most strongly in the far-infrared region. Silicon and oxygen ions have smaller masses and a stronger bond, and so silicon dioxide attenuates signals most strongly in the near-infrared region.

triselenide epitomize the properties of the chalcogenide glasses. The electronic absorption of arsenic trisulfide lies in the middle of the visible spectrum and that of arsenic triselenide is in the near infrared. Consequently arsenic trisulfide glass is red in color and arsenic triselenide glass is an opaque black.

These materials have high refractive indexes (between 2.4 and 2.7) and low glass-transition temperatures (between about 150 and 175 degrees Celsius); as a result they exhibit low Rayleigh scattering. Arsenic trisulfide is transparent to about 10 microns, whereas the glass based on selenium, which has more than twice the atomic weight of sulfur, transmits to a wavelength of about 14 microns.

In contrast to single or polycrystalline materials, chalcogenide glasses can be easily fabricated into optical fibers possessing the proper core, cladding and diameter that can stretch for many kilometers. Among the best chalcogenide fibers are those prepared from arsenic trisulfide by Terutoshi Kanamori of the Nippon Telegraph and Telephone Public Corporation in Japan. A minimum attenuation of 35 decibels per kilometer was obtained at 2.4 microns, although the projected intrinsic limit of these fibers is .01 decibel per kilometer at five microns. Much of the loss is a result of contamination from water molecules and other hydrogen-containing impurities. A more serious problem affecting chalcogenide fibers is the possible existence of a strong intrinsic electronic absorption in the infrared due to defects in the glass structure that may limit the minimum attainable attenuation to 10 decibels per kilometer.

Much of the current interest in chalcogenide fibers stems from their ability to transmit six-to-12-micron infrared light over a distance of a few meters. Peter Prideaux of the Galileo Electro-Optics Corporation (one of our laboratories) has fabricated chalcogenide fibers—hundreds of which can be assembled to produce a flexible fiber-optic bundle capable of transmitting simple images at infrared wavelengths or information about the temperature of an object. The intensity of the infrared light emitted by an object increases as the temperature rises. Fibers can guide the light to a detector, allowing thermometry to be carried out from remote locations. Gases, liquids or chemical reactions, which absorb or emit infrared light, could be monitored at a central location in a factory using a network of fibers.



HEAVY-METAL FLUORIDE GLASSES show most promise as materials for low-loss optical fibers. The broken curve (*top*) represents attenuation measurements on a fluoride glass in four different wavelength regions. The samples had a minimum attenuation of 4.3 decibels per kilometer. The dotted curve predicts the minimum attainable attenuation to be near .01 decibel per kilometer. The disparity between the prediction and these measurements is due to extrinsic absorption and scattering. The solid curve (*bottom*) highlights the low-attenuation region of a fiber prepared by Paul W. France at the British Telecom Research Laboratories. Near one micron some loss results from electronic absorption by metallic contaminants. The two peaks in attenuation are attributed to copper and chromium, which together with other impurity metals have a concentration of about 85 parts per billion in the glass. Hydroxyl ions (from moisture that enters the glass during fabrication) absorb light strongly near wavelengths of three microns. Hydroxyl ions present in a concentration of one part per million can attenuate a signal 10,000 decibels per kilometer.

The third class of infrared fiber materials consists of the heavy-metal fluoride glasses. Michel Poulain of the University of Rennes in France serendipitously discovered the first member of this family of materials in 1974. As part of his graduate studies, Poulain was attempting to fuse a crystalline compound from a mixture of zirconium, barium, sodium

Great holiday savings on this powerful computer with built-in MS-DOS® and DeskMate® software—only $799.

Start computing Christmas morning with the Tandy 1000 SL—now at a terrific low price. Or purchase the 1000 SL complete with a CM-5 color monitor for just $1098.95.

MS-DOS and DeskMate are built into ROM, so you can move quickly from start-up to MS-DOS programs, or go directly into our DeskMate multi-applications software.

With DeskMate, you're ready to start computing right away with ten popular applications to write letters, prepare budgets, file information and even draw colorful pictures. There's also a spell checker, address book, Hangman word game and PC-Link℠ on-line information service.

As an added bonus, there's a Music program as well as a sound-editor program that lets you record and edit music, voice or other sound effects onto your 5¼" diskettes. This latest breakthrough in computer-audio technology delivers outstanding speech and music capabilities which will generate new life into programs.

Come in to Radio Shack today. The Tandy 1000 SL is a gift the whole family will love.

Tandy® Computers: Because there is no better *gift* value.™

Radio Shack®
The Technology Store
A DIVISION OF TANDY CORPORATION

Sale ends 12/24/88. Tandy 1000 SL reg. $899. CM-5 monitor reg. $299.95. Prices apply at Radio Shack Computer Centers and participating stores and dealers. DeskMate communications require modem. MS-DOS/Reg. TM Microsoft Corp. PC-Link/SM Quantum Computer Corp.

GRANULAR MICROSTRUCTURES of thallium halide crystals scatter a large fraction of the light propagating through the fiber. The fibers therefore exhibit attenuations 1,000 times greater than silica fibers even though thallium halide single crystals should theoretically have greater transparency. Thallium halide fibers can transmit energy from lasers for surgical applications. This fiber, 75 microns in diameter, was made by James A. Harrington, who was then at the Hughes Research Laboratories.

and neodymium fluorides. When he cooled the sample, he noticed large clear fragments. Although he initially thought the fragments were crystals, they proved to be glass. He, Marcel Poulain (his brother) and Jacques Lucas then found a host of new glass compositions, which are now known as the heavy-metal fluoride glasses.

Virtually any metal in the periodic table can be incorporated into a heavy-metal fluoride glass. From the standpoint of infrared transparency and ease of glass fabrication, however, only a limited number of compositional types have warranted intensive study: the fluorozirconate, fluorohafnate and barium-thorium glasses.

Measurements on the transmission characteristics of heavy-metal fluoride glass indicate the existence of high transparencies over a broad spectrum of wavelengths that range from .3 to eight microns. The materials have moderate glass-transition temperatures (near 300 degrees C.) and silicalike refractive indexes (1.5). A long-wavelength infrared vibrational edge and low Rayleigh scattering suggest that intrinsic attenuations of .01 decibel per kilometer are possible in heavy-metal fluoride glasses.

For this reason the materials have been intensively examined as candidates for low-loss fiber communications systems in the U.S., the U.K. and Japan. The U.S. Naval Research Laboratory and the Nippon Telegraph and Telephone Public Corporation have fabricated seven-to-30-meter lengths of fluoride fiber that exhibit attenuations of from .7 to .9 decibel per kilometer at 2.5 microns. Such transparencies are far superior to those of other infrared-fiber materials. Yet they are still about seven times as great as the attenuation found in the best silica fibers and far above the intrinsic limit of fluoride fibers.

This limit can be approached by identifying sources of extrinsic loss in the fibers and controlling them. Detailed light-scattering and attenuation experiments have already revealed some of the causes of extrinsic losses. At wavelengths from .5 to two microns, some attenuation is due to electronic absorption by minute impurities of copper, chromium, nickel and iron. Contamination by only one part per million of iron, for instance, can lead to an attenuation of 15 decibels per kilometer near 2.5 microns. Hydroxyl groups, introduced into the fiber as moisture either in the starting materials or during the glass-melting process, attenuate infrared light strongly. Estimates show that one part per million of hydroxyl can cause an attenuation of approximately 10,000 decibels per kilometer at 2.9 microns. Although significant advances have been made in purifying the chemicals melted to form glass, extrinsic scattering from small crystals and bubbles continues to be a perplexing problem in fluoride fibers.

Even at current levels of transparency, heavy-metal fluoride glass offers technological opportunities. Dieter Pruss of Dragerwerk AG in West Germany has used fluoride fibers to monitor the uptake of anesthetic gases by patients undergoing surgery. These gases have unique infrared absorption bands, which can be transmitted through a fiber and detected. The concentration of the anesthetic can then be displayed, breath by breath.

In most applications infrared optical fibers serve simply to transmit light from one point to another. The material can also be employed to create special fibers that can alter the amplitude or wavelength of the light passing through them. One example of this technology is the optical-fiber laser. The fiber core is doped with small amounts of certain rare-earth ions. A section of fiber is placed between two partially reflecting mirrors. Laser light of a particular wavelength is injected into the fiber and excites the rare-earth ions to a high-energy state. As the excited atoms give up light and revert to a lower energy state, they emit light at a wavelength longer than the original stimulating light. This light is amplified as it bounces back and forth along the length of the fiber, repeatedly reflected by the end mirrors. One mirror is partially transparent to the emitted light, allowing some of it to escape. The fiber thus acts as a compact, solid-state laser light source, producing a wavelength much different from that of the original laser. The emitted wavelength can be controlled by changing the original wavelength and selecting different rare-earth dopants. Such devices, prepared from heavy-metal fluoride glass fibers, can be used to efficiently generate infrared laser wavelengths that cannot be obtained in silica-based fibers.

It is our belief that heavy-metal fluoride glasses offer the best prospects for achieving optical losses lower than those now attainable in high-quality silica fibers. Before such low-loss communication systems can be built, significant advances are required in purifying and manufacturing heavy-metal fluoride fibers. In the meantime infrared optical fibers with losses of from 10 to 100 decibels per kilometer can be used in medical diagnostic devices, for optical-power transmission, for remote sensing and as the key component in optical-fiber lasers. For the short term these applications have the brightest future.

FURTHER READING

INFRARED OPTICAL FIBERS. Tadashi Miyashita and Toyotaka Manabe in *IEEE Journal of Quantum Electronics,* Vol. QE-18, No. 10, pages 1432–1450; October, 1982.

SPECIAL ISSUE ON LOW-LOSS FIBERS. In [IEEE] *Journal of Lightwave Technology,* Vol. LT-2, No. 5; October, 1984.

THE SEARCH FOR VERY LOW LOSS FIBER-OPTIC MATERIALS. M. E. Lines in *Science,* Vol. 226, No. 4675, pages 663–668; November 9, 1984.

HEAVY METAL FLUORIDE GLASSES. Martin G. Drexhage in *Treatise on Materials Science and Technology,* Vol. 26, pages 151–243; 1985.

HALIDE GLASSES. Cornelius T. Moynihan in *Materials Research Society Bulletin,* Vol. 12, No. 5, pages 40–44; June 16, 1987.

Automobile drivers using a Head-Up Display (HUD) need not divert their eyes from the road to view information ordinarily shown in the instrument panel. The HUD focuses key driving data, including vehicle speed and turn signal indicators, through the windshield, so the data appears to be suspended over the front bumper. The system design is derived from HUDs used in today's advanced fighter aircraft. Developed by Hughes Aircraft Company, Delco Electronics, and General Motors, the automotive HUD uses a collimating optical system, with an electronically driven, high-performance display device. The world's first production cars equipped with a HUD will be the 1988 Cutlass Supreme Indianapolis 500 Pace Car and limited edition replica cars.

A palm-size supercomputer being developed by Hughes for the U.S. Air Force is designed to offer data processing speeds in excess of 100 billion operations per second. In a radical departure from conventional computer design, the new computer processes its signals in a three-dimensional network and will have capacities to rival today's room-size systems. Consisting of a number of stacked integrated circuit wafers, the 3-D computer utilizes tiny "microbridges" and feedthroughs to provide an unprecedented degree of interconnectivity and a high degree of parallel processing. The combination of ultra-high speeds and low power consumption in such tiny packages is expected to make 3-D computers ideally suited for image processing systems and signal processing of complex radar information.

Japan's first commercial communications satellite will have the highest communications capacity of any domestic satellite outside the United States. Once in orbit, the antenna shelf of the HS 393 JCSAT will remain stationary and pointed toward Earth to provide 24-hour communications. The satellite's propulsion unit and cylindrical skin of solar cells spin like a gyroscope to provide stability. Hughes is building two of the HS 393 satellites for Japan Communications Satellite Company.

One of the longest lived production runs of a high-power traveling wave tube (TWT) produced by the industry still represents state-of-the-art technology, and has enjoyed a high instance of on-time deliveries and quality workmanship. Hughes pioneered TWTs, and the Model 797H is one of the industry's most successful programs with the 4,000th unit of the landmark TWT recently being manufactured. In 1987, the program recorded a 100 percent test yield by shipping 27 tubes out of 27 tubes tested. In addition, the program has reduced manufacturing costs more than 30 percent by using cycle time management techniques and redesigning for ease of manufacture. The 797H's main application has been in a sophisticated radar system which has been in continuous production for over 20 years.

Hughes is seeking experienced engineers and scientists to design, develop, and produce Hughes' new line of body-stabilized HS 601 communications satellites. Openings are in the fields of: software, computers, and data processing systems; electrical components; microwave/RF communication systems development; on-board spacecraft electronics and control systems; satellite design, integration, propulsion, and electrical power system development; spacecraft manufacturing, systems test and evaluation; GaAs applications R&D. Send your resume to Michael Martinez, Hughes Space & Communications Group, Dept. S2, S4/A300, P.O. Box 92919, Los Angeles, CA 90009. Equal opportunity employer. U.S. citizenship required.

For more information write to: P.O. Box 45068, Los Angeles, CA 90048-0068

HUGHES

Subsidiary of GM Hughes Electronics

Evolution of Human Walking

Features of her pelvis show that a three-million-year-old hominid, Lucy, was as adept at upright walking as we are. Bipedality could date from the earliest phase of human evolution

by C. Owen Lovejoy

Asked to choose the most distinctive feature of the human species, many people would cite our massive brain. Others might mention our ability to make and use sophisticated tools. A third feature also sets us apart: our upright mode of locomotion, which is found only in human beings and our immediate ancestors. All other primates are basically quadrupedal, and with good reason: walking on two limbs instead of four has many drawbacks. It deprives us of speed and agility and all but eliminates our capacity to climb trees, which yield many important primate foods, such as fruits and nuts.

For most of this century evolutionary theorists have held that human ancestors evolved this strange mode of locomotion because it freed their hands to carry the tools their larger brains enabled them to make. Over the past two decades, however, knowledge of the human fossil record has expanded. Neither a unique brain nor stone tools are in evidence among our earliest known ancestors, the australopithecines of three million years ago and more. Yet these same ancestors do clearly show many of the hallmarks of bipedal walking.

How long had human ancestors been walking upright? Was bipedality fully developed in the hominids of three million years ago, or did they sometimes revert to using all four limbs for running or climbing? The answers can help to solve the puzzle of bipedality's role in early human evolution. If upright walking was well established by the time of *Australopithecus,* its advent could date back as far as the earliest hominids, whose lineage probably diverged from other primates some eight or 10 million years ago. The development of erect walking may have been a crucial initiating event in human evolution.

I have proposed that bipedality accompanied a set of behavioral adaptations that became the key evolutionary innovation of humanity's earliest ancestors. These adaptations included, in effect, the nuclear family: lasting monogamy together with care of the offspring by both parents. The male's contribution took the form of providing high-energy food, which expanded the mother's ability to nurture and protect each infant and also enabled her to give birth more often. Bipedality figured in this new reproductive scheme because by freeing the hands it made it possible for the male to carry food gathered far from his mate. These developments must have come long before the current hominid fossil record begins.

Upright walking should therefore have been perfected by the time of an australopithecine female whose fossil has become a test case for early walking. In 1974 the continuing search for human ancestors in the Afar Triangle of Ethiopia, led by Donald C. Johanson of the Institute of Human Origins in Berkeley, Calif., was splendidly rewarded by the recovery of the "Lucy" skeleton, known formally as A.L. 288-1. Although the skeleton is not quite complete, it preserves far more detail than any comparable fossil. In particular, it includes many of the lower-limb bones, one of the innominate bones that, in a mirror-image pair, make up the primate pelvis, and an intact sacrum (the fused vertebrae at the back of the pelvis). Upright walking is so dependent on this structure that an analysis of Lucy's pelvis can reveal how well she and her contemporaries walked.

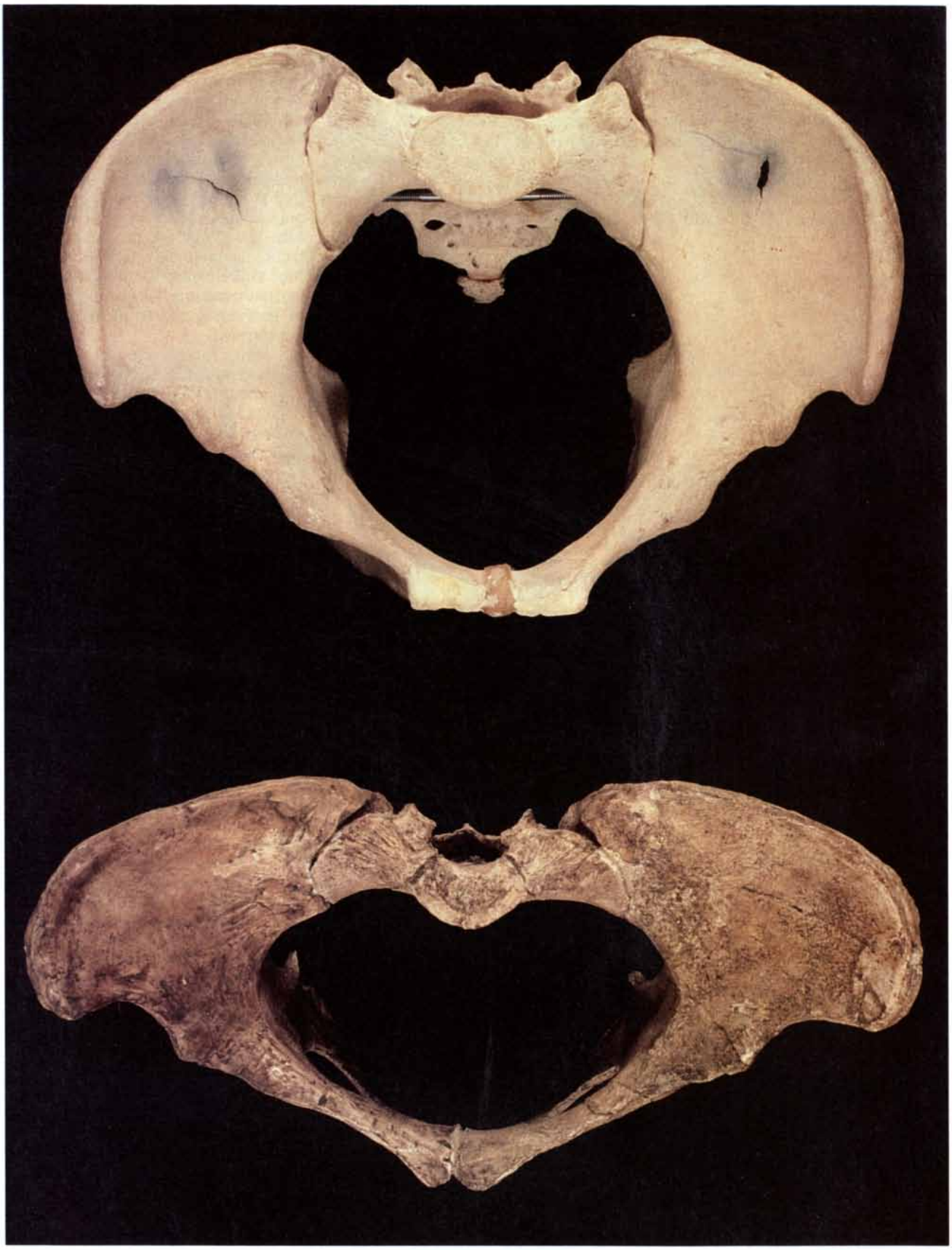
The distinctive pelvic features of a biped reflect the very different mechanics of two- and four-legged locomotion. In order to propel itself any terrestrial mammal must apply a force against the ground in a direction opposite to the direction of travel. It does so by extending the joints of its legs, which lie between the ground and the animal's center of mass. Lengthening a leg produces a "ground reaction" that propels the torso in a direction determined by the angle between the leg and the ground.

In the quadrupedal posture of most primates the center of mass lies well forward of the hind limbs. Hence extending the hind limbs generates a ground reaction that has a large horizontal component. Because the hip and knee joints of the hind limbs are tightly flexed at the start of each cycle, their extension can be prolonged and powerful.

Our upright posture, in contrast, places our center of mass almost directly over the foot. If we stand erect and lengthen our legs by straightening the knee and rotating the ankle, the ground reaction is directed vertically and we end up on tiptoe. In order to propel our upright trunk we must reposition our center of mass ahead of one leg. The trailing limb is lengthened to produce a ground reaction while the other leg is swung forward to keep the trunk from falling. The strength of the ground reaction is limited, because much of it is still directed vertically and also because the trailing limb is already near its limit of extension owing to our upright posture: the hip joint is fully extended and the knee joint nearly so.

With the new bipedal strategy there came new roles for most of the muscle

C. OWEN LOVEJOY is professor of anthropology at Kent State University; he also holds a staff position at the Case Western Reserve University School of Medicine and teaches anatomy at the Northeast Ohio Universities College of Medicine. He received his Ph.D. in physical anthropology from the University of Massachusetts at Amherst in 1970 and later helped to restore and analyze the Lucy skeleton. Lovejoy, who has written and lectured widely about human origins, also pursues an interest in forensic anthropology as an adviser to the Cuyahoga County Coroner's Office.



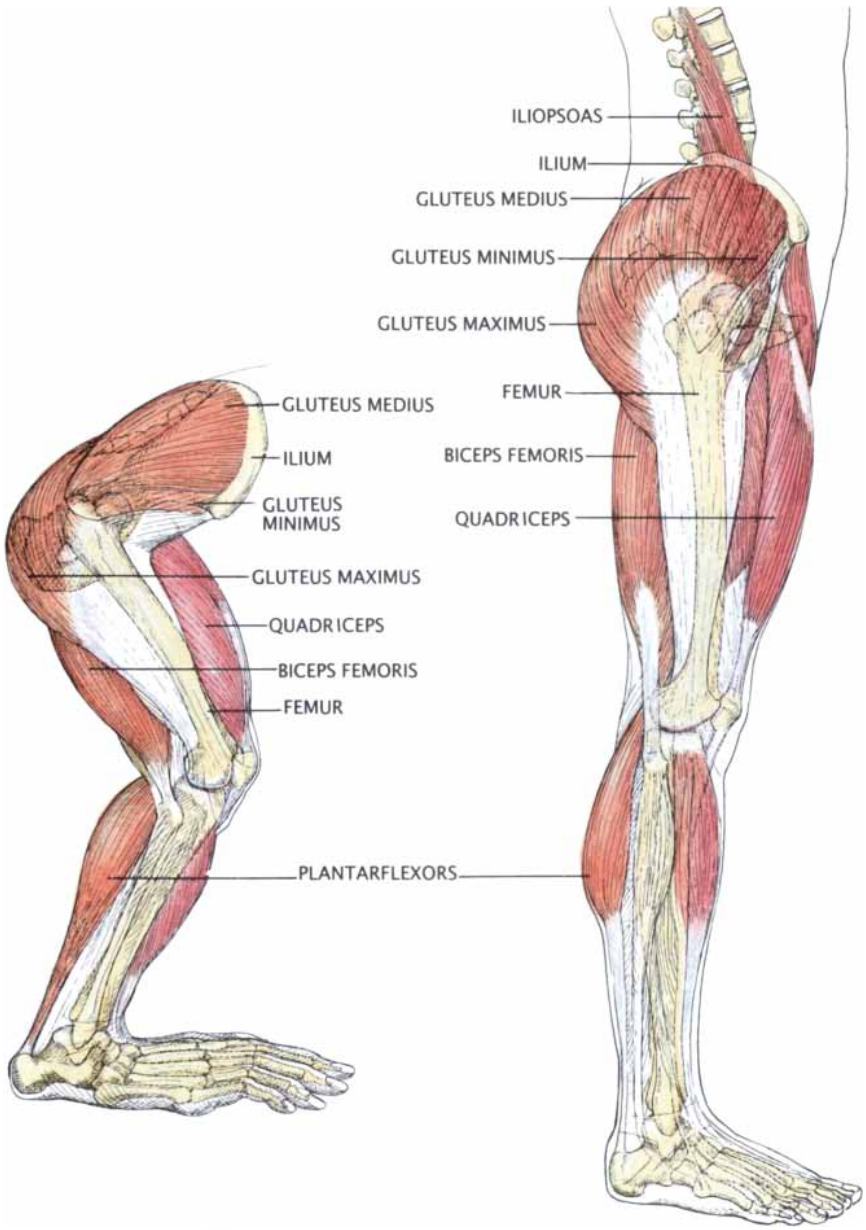
PELVISES of a modern human female (*top*) and Lucy (*bottom*) are separated by three million years of evolution but bear the same hallmarks of upright walking. The major change visible in this view—the more ovoid form of the human pelvis—accompanied an expansion of the birth canal, needed because of the increase in brain size since Lucy. The author and Barbara Brown restored the Lucy pelvis from the fragmented fossil; Larry Rubens of Kent State University made the photograph.

groups in the lower limb—roles that in turn required changes in the muscles' structure or position and hence in the design of the pelvis. A comparison of the human pelvis with that of our closest living relative, the chimpanzee, highlights these changes in mechanical design.

The need to stabilize an upright torso dictated the most dramatic change in musculature that has come with the adoption of bipedality: the transformation of the gluteus maximus, a relatively minor muscle in the chimpanzee, into the largest muscle in the human body. The gluteus maximus originates over much of the back of the pelvis and is attached to the back and side of the upper femur, or thighbone. As such it is defined as a hip extensor, and many classical anatomists believed it serves as the major propulsive muscle in upright walking. By straightening the hip, it was thought, the gluteus maximus contributes to the ground reaction imparted by the trailing leg.

Actually, because the hip is almost completely extended in the first place during erect walking and running, the muscle's contribution to ground reaction is limited. Its hypertrophy in human beings reflects a quite different function. When we run, our upright trunk tends to flex forward at each foot strike owing to momentum. The gluteus maximus has taken on the role of preventing the trunk from pitching forward.

A major modification of the pelvis has made the muscle's stabilizing task considerably easier. Each innominate bone in the pelvis is topped by a blade of bone called an ilium; most of the lower viscera are cradled in the space between the two ilia. In the chimpanzee and other primates the ilia are much longer than they are in humans. The long ilia have the effect of lengthening the torso; when these primates rear up, their center of mass lies well above their hip joints. In the language of engineering, their trunk has a long lever arm. A gluteus maximus working to hold such a trunk upright would tire rapidly. The dramatically shortened human ilium shortens the torso and brings the trunk's center of mass much closer to the hip joints, thereby reducing the muscle's mechanical disadvantage.



PELVIS AND LEG of a chimpanzee (*left*) and a human being (*right*) reflect the differing demands of quadrupedal and bipedal locomotion. The musculature of the chimpanzee pelvis is dominated by the gluteus medius and gluteus minimus, which help to propel the animal by extending its hip joint. They are joined in that task by the hamstrings, which include the biceps femoris. In humans the gluteus maximus dominates the pelvis; it serves the new function of stabilizing the upright trunk. (The shortening of the ilium lowers the trunk's center of mass and makes it easier to control.) Other major muscles, such as the gluteus medius and minimus, the hamstrings and the iliopsoas, also play new auxiliary roles in upright walking. Only two muscle groups—the quadriceps and plantarflexors—are left to provide propulsion.

The ilium is long in the apes to accommodate a second muscle group that was transformed as our ancestors began walking upright: the anterior gluteals, composed of the gluteus medius and the gluteus minimus. In the chimpanzee these muscles contract between attachment points near the top of the ilium and on the outside of the upper femur. Their position enables them to serve as powerful hip extensors during quadrupedal locomotion, and because the ilium is long, the muscles have a large range of contraction. Human beings can forgo this almost universal skeletal feature of other primates because hip extension contributes very little to bipedal locomotion. Our anterior gluteals have been freed to assume a new role.

This new role is best understood by imagining a head-on view of a person walking. Soon after the heel of the leading foot strikes the ground, the trailing leg leaves the surface and begins to swing forward. While it does so the trunk is supported by only one hip, which lies well to the side of the

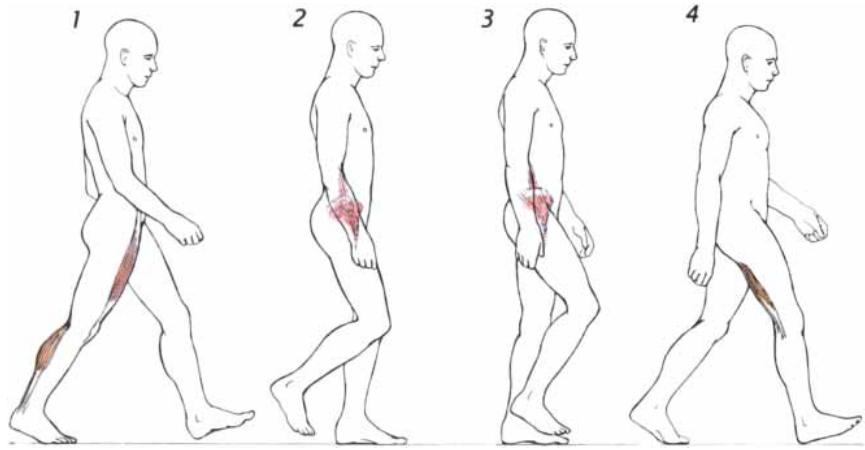
trunk's center of mass. On their own the pelvis and trunk would tip toward the unsupported side at each step, causing rapid fatigue; they are prevented from doing so by the action of the anterior gluteals, which are also referred to as abductors in human beings.

The transformation of the anterior gluteals from propulsive muscles to stabilizing ones required major changes in their position. A top view of the human and chimpanzee pelvises reveals a radical reorientation of the iliac blades in the human pelvis. In the chimpanzee the blades are flat and lie more or less in a single plane across the back of the torso. In humans each ilium has been rotated forward, carrying with it the upper attachment point of the gluteals. Their lower attachment point falls on the outside of the upper femur, where the bone forms a neck that angles in to meet the pelvis at the hip joint. The abductors are thus disposed laterally in humans, away from the hip joints, which puts them in position to balance the pelvis against the weight of the trunk.

The reorientation of the ilia required two other changes in pelvic design not dictated directly by the mechanics of bipedality. If the ilia had simply been rotated forward, the space between them would have been sharply narrowed, leaving no room for the lower viscera. In compensation the sacrum, which separates the ilia at the back of the pelvis, has grown wider and the ilia have changed in shape: they are dished, so that the bending that has reoriented the abductors takes place well to the side, leaving ample room within the pelvis.

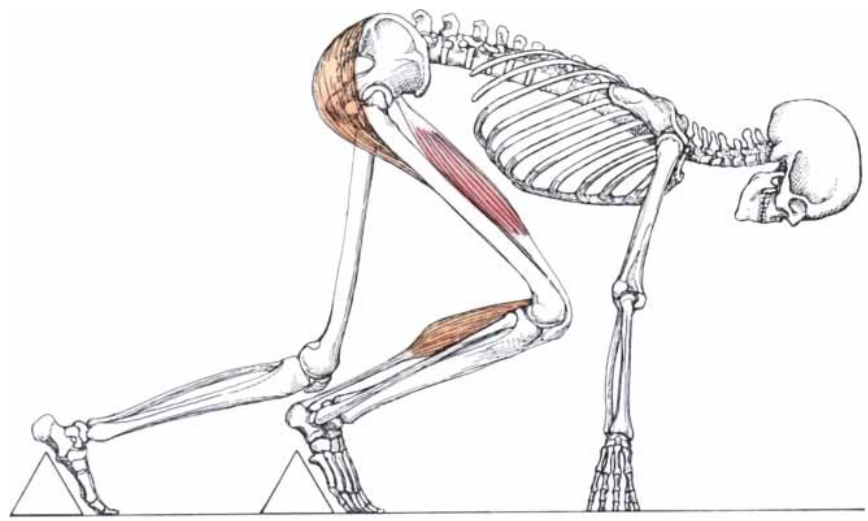
By increasing the distance between the hip joints, however, this widening of the central pelvis placed the abductors in a position of considerable mechanical disadvantage. The force the abductors must exert to offset the weight of the trunk depends in part on how far to the side of the trunk's center of mass each hip joint lies. The greater the separation of the hip joints is, the longer the trunk's lever arm will be and the harder these muscles will have to contract to offset its weight. They will be more likely to tire during walking, and the safety of the hip joint itself may be threatened, since the joint is subjected to both the weight of the torso and the abductors' force of contraction.

A front view of the human pelvis reveals the evolutionary solution. The abductors' own lever arm can be increased, and their work made easier, if their upper and lower attachment points are moved farther out from the hip joint. Two features of the human pelvis serve that purpose. The complex curvature of the human ilium includes an outward flare, which displaces the upper attachment point of the abductors to the side of the hip. In addition the human femoral neck is longer than that of the chimpanzee. The longer femoral neck serves to move the abductors' lower attachment point outward as well, adding to their leverage.

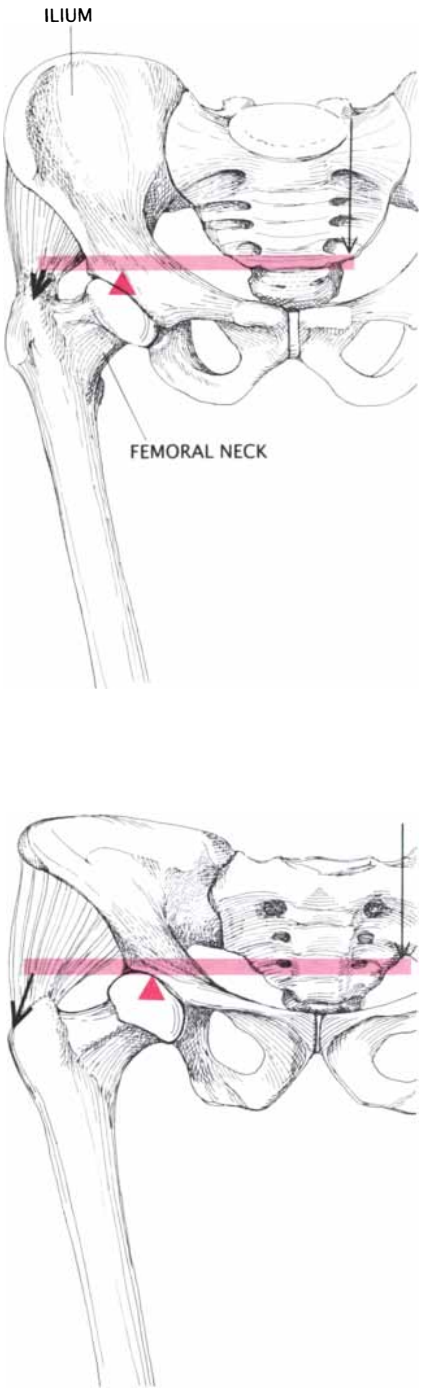


MUSCLE ACTIVITY during human striding is diagrammed. As the weight-bearing leg (here the right leg) becomes angled behind the torso (*1*), two muscle groups contract to extend it, generating a "ground reaction" that propels the body; they are the plantarflexors, which rotate the foot around the ankle, and the quadriceps, which straighten the knee. The foot then leaves the ground as weight is tranferred to the left leg. Contraction of the iliopsoas begins to tug the right leg forward (*2*) while the knee flexes passively (*3*). Near the end of the leg's swing the hamstrings contract to stop it, and the foot is planted (*4*). The left leg in turn generates ground reaction.

One set of muscles—the anterior gluteals—that help to propel chimpanzees has thus become co-opted to stabilize the human pelvis. A new role is also evident for another set of propulsive muscles in the chimpanzee: the hamstrings. They connect the lower pelvis to the back of the femur; in quadrupedal locomotion



SPRINTER on the starting block briefly recovers the advantages of being quadrupedal: the hip and knee joints are tightly flexed, preparing the limbs for prolonged and powerful extension, and the center of mass is positioned well forward of the legs, which gives the ground reaction a strong horizontal component. Ordinary walking or running sacrifices these advantages. An upright posture requires the hip and knee joints to be almost fully extended and places the body's center of mass almost directly over the legs. Both factors tend to limit the strength of the ground reaction.



ABDUCTOR MUSCLES (the gluteus medius and minimus) contract to counterbalance the torso when the human pelvis is supported on only one leg. The hip joint acts as a fulcrum, with the weight of the torso and unsupported leg bearing down on one side and the abductors acting on the other (*top*). The abductors are at a mechanical disadvantage: the hip joint lies well to the side of the torso's center of mass, giving the body weight a long lever arm. In the Lucy pelvis (*bottom*) the body-weight lever arm was even longer, but greater lateral flare of the ilium and a longer femoral neck placed the abductors farther from the hip joint, increasing their mechanical advantage.

they serve as powerful hip extensors, which contribute even more to ground reaction than the anterior gluteals do. In bipedal walking, in contrast, they serve not to extend the limb but to control it.

A biped must swing each leg forward rapidly when it is not bearing weight. Because the limb is carried almost fully extended in a biped rather than tightly flexed, as it is in a quadruped, its center of mass lies well away from the pelvis. Like a long pendulum, an extended leg has a large moment of inertia, and it takes powerful muscle impulses to start and stop its swing. The iliopsoas, a muscle that originates within the pelvis and extends forward to an attachment point on the femur just below the hip joint, contracts to tug the limb forward. Once the leg has completed its arc, its swing must be checked. The position of the hamstrings, which is largely unchanged from the position in other primates, enables them to contract and decelerate the limb.

In human beings, then, the demands of stabilizing the pelvis and controlling the limb occupy several muscle groups that serve for propulsion in the chimpanzee. Only two muscle groups, the quadriceps and the plantarflexors, are left in positions that enable them to produce a ground reaction. The quadriceps are a mass of four muscles that make up most of the front of the human thigh. They end in a stout tendon, which crosses the patella, or kneecap, and is anchored to the top of the tibia, the main bone of the lower leg.

As the weight-bearing leg becomes angled behind the torso during walking or running, this powerful muscle mass contracts and straightens the knee. The plantarflexors, which originate at the back of the lower leg and are attached to the heel by the Achilles tendon, contract in synchrony with the quadriceps and cause the foot to rotate about the ankle. The extension of the knee and the rotation of the foot together lengthen the trailing leg, producing a strong ground reaction.

How well developed was this set of muscular adaptations by the time of Lucy and her kin, according to the fossil evidence? The discovery included a largely intact sacrum, but the innominate bone that accompanied it had been broken and partially crushed; it consisted of about 40 separate pieces fused into a single mass by the matrix of stone in which it was preserved. Often a fossil in this condition can be reduced to its separate pieces and then reassembled like a jigsaw puzzle. The pieces of Lucy's innominate, however, could not safely be separated. Instead I took a cast of each piece and assembled the casts in proper anatomical juxtaposition; the restored innominate was then mirror-imaged to create its opposite number. The result was a complete pelvis of an almost three-million-year-old human ancestor.

The pelvis bears all the hallmarks of bipedality seen in our own. Its ilia are much shorter than those in the pelvis of an ape. The shortening would have lowered the trunk's center of mass and made it easier to keep upright. The ilia have also become bent around to provide lateral attachment for the abductor muscles that stabilize the bipedal pelvis when it is supported on one leg. The attachment points for the gluteus maximus, abductors and quadriceps can be seen, and they indicate that in Lucy these muscles had attained a size and disposition remarkably similar to our own arrangement. The same is true for the iliopsoas, the hip flexor that initiates the swing of the leg: a groove on the brim of the pelvis, ahead of the hip joint, matches the groove that indicates the muscle's course in the human pelvis.

In one respect Lucy seems to have been even better designed for bipedality than we are. Her ilia flare outward more sharply than those of a modern pelvis and her femoral necks are longer. Her abductor muscles thus enjoyed a greater mechanical advantage than these muscles do in modern females. Some of the abductors' advantage merely compensated for the slightly wider separation of her hip joints (which gave her trunk a longer lever arm). Yet accurate measurements of both the abductor and the trunk lever arms—possible because the Lucy pelvis is so complete—show that her abductor advantage is still greater than our own. Her abductors had to exert less force to stabilize the pelvis, which also reduced the pressure on the hip-joint surfaces.

Why should a three-million-year-old hominid have had this mechanical advantage over her descendants? The answer lies in the accelerated growth of the human brain during the past three million years. Lucy's pelvis was almost singularly designed for bipedality. The flaring ilia and long femoral necks increased her abductors' lever arm, but they yielded a pelvis that in top view was markedly elliptical, resulting in a birth canal that was wide but short from front to back. The constriction was tolerable because Lucy

predated the dramatic expansion of the brain; her infant's cranium would have been no larger than a baby chimpanzee's. The process of birth in Lucy and her contemporaries would have been slightly more complex than in an ape, but much easier than the modern human birth process [*see illustration on page 125*].

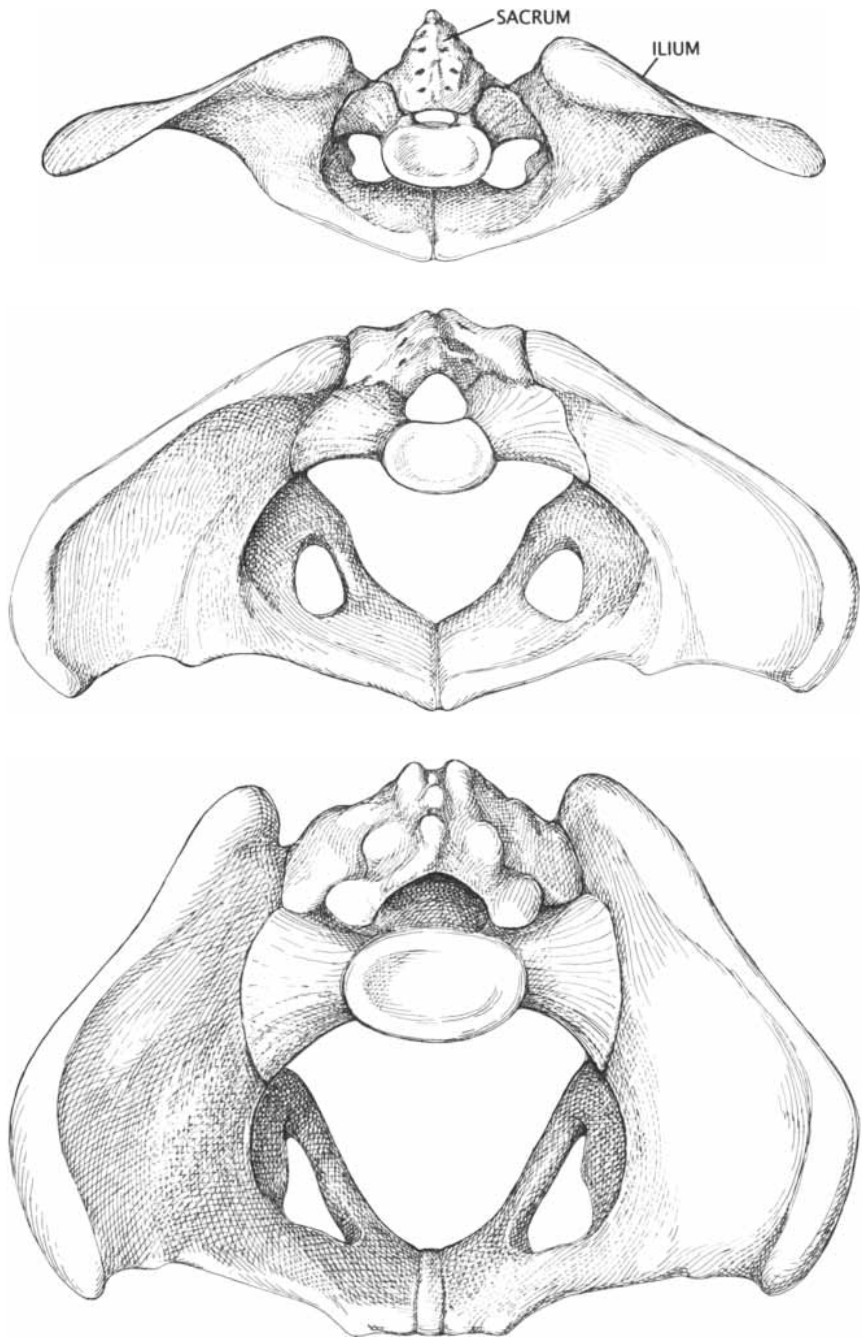
As human ancestors evolved a larger brain, the pelvic opening had to become rounder. The pelvis had to expand from front to back, but at the same time it contracted slightly from side to side. In the process the flare of the ilia was reduced, leaving us with a somewhat shorter abductor lever arm than Lucy's. (These changes are less pronounced in the modern male pelvis, where the abductors retain some of their former mechanical advantage.) Meanwhile the head of the modern femur has become enlarged to withstand increased pressure from the harder-working abductors. The difficulty of accommodating in the same pelvis an effective bipedal hip joint and an adequate passage for a large infant brain remains acute, however, and the human birth process is one of the most difficult in the animal kingdom.

The close resemblance of Lucy's pelvis to that of a modern human and its dramatic contrast to the pelvis of a chimpanzee make it clear that she walked fully upright. But was her bipedal progression truly habitual? Had she forsaken all other kinds of locomotion? The muscular rearrangements that enabled her to walk upright would not have allowed efficient quadrupedal movement on the ground. Perhaps, however, she often took to the trees and climbed, as most primates do, using all four limbs.

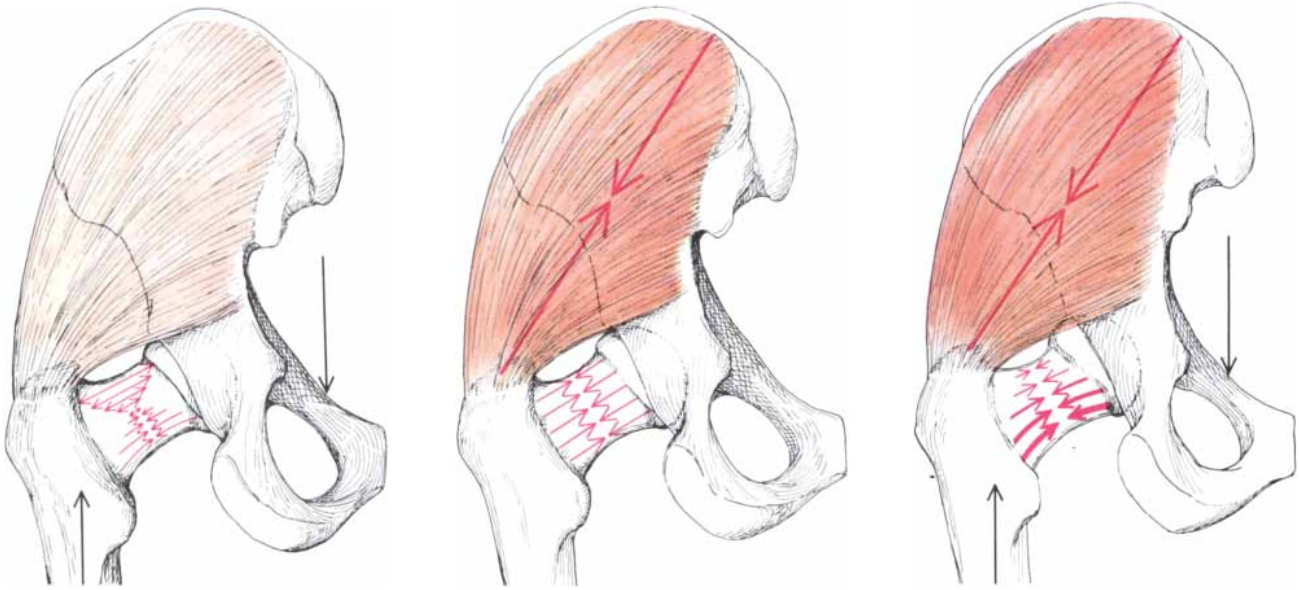
Basic evolutionary principles provide one kind of verdict on the possibility. A species cannot develop detailed anatomical modifications for a particular behavior, such as bipedality, unless it consistently employs that behavior. For natural selection to have so thoroughly modified for bipedality the skeleton Lucy inherited, her ancestors must already have spent most of their time on the ground, walking upright. Analysis of the Lucy fossil, however, can yield more direct evidence.

The analysis focuses on the neck of the femur, where much of the stress of locomotion is concentrated. When the leg is bearing weight, the hip joint transmits the weight of the torso to the femoral neck. The neck acts as a cantilevered beam: a beam that is anchored at one end to a supporting structure (the shaft of the femur) and carries a load at the other end. Cantilevering results in high bending stresses at the beam's anchorage—compression along the bottom of the beam and tension along the top—and the stresses increase with the length of the beam. A long femoral neck such as Lucy's reduces pressure on the hip joint by improving the leverage of the abductors, but the neck itself is subject to higher bending stresses.

The femoral neck of the chimpanzee is much shorter than the modern human one; nonetheless, it is robustly engineered to withstand the loads im-



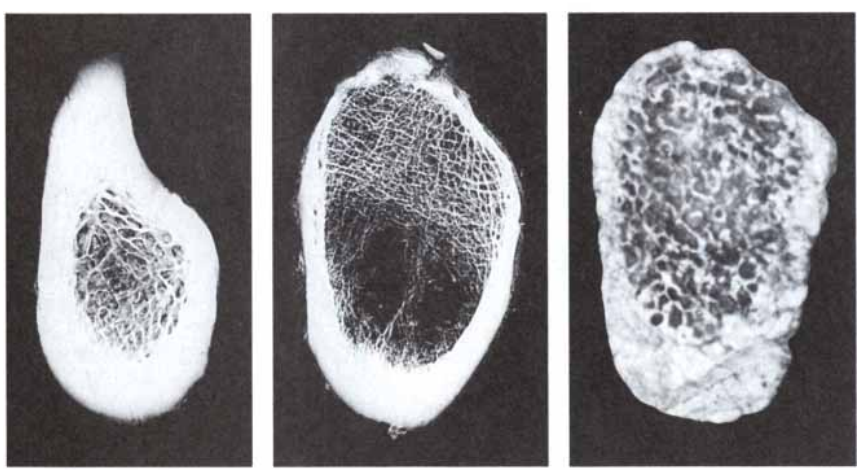
ROTATION OF THE ILIA took place as human ancestors began walking upright. In a quadrupedal ape such as a chimpanzee (*top*) the ilia (seen here from above) lie almost flat against the back of the torso. In Lucy (*middle*) they have become bent around, providing lateral attachment points for the abductor muscles, which stabilize the pelvis during walking. The bending takes place well away from the center of the pelvis, leaving room for the viscera; in addition the sacrum, which separates the ilia, has widened. These changes are retained in the modern human pelvis (*bottom*), which has also become longer from front to back to create a more ovoid birth canal.



NECK OF THE FEMUR (shown from the back) is subjected to stress from two sources during human walking. Body weight imposes bending stress: tension on the top of the neck and compression on the bottom (*left*). At the same time the abductors, acting almost parallel to the femoral neck, subject its entire diameter to compression (*middle*). The sum of the two stress patterns is a gradient of stress running from low stress at the top to high compressive stress at the bottom (*right*).

posed by the animal's terrestrial and arboreal acrobatics. A cross section of the bone reveals a central marrow-filled channel surrounded by a thick layer of dense bone. Dense bone is weaker under tension than it is under compression, and so the upper surface of the structure, which will be subjected to tension when the neck is bent, carries a markedly thicker layer of bone. With this ridge of thick bone (a bone "spike" in cross section), the chimpanzee femoral neck imitates the principle of an I beam: material is placed where it can best resist bending stresses.

Because the human femoral neck is longer than the chimpanzee's and must resist the combined force of body weight and abductor contraction, one would expect it to be even more robustly constructed. A cross section of the human bone reveals a surprise: the outer ring of solid bone is thick only at the bottom, and the rest of the neck is bounded by a thin shell of bone and filled in by a lattice of fine bone plates called trabeculae. Such porous bone, as one might expect, is weaker than solid material. The upper part of the femoral neck, where tensile stresses are presumably the highest, actually contains less bone than any other part of the structure. How can our femoral neck survive the greater stresses imposed by its length and function when it seems so much less sturdy than the femoral neck of the chimpanzee?



INTERNAL STRUCTURE of the femoral neck distinguishes habitual bipeds. Seen in cross section, the femoral neck of the chimpanzee (*left*) has a robust thickness of bone together with a reinforcing ridge (visible in this section as a spike) at the top. These features enable the chimpanzee femoral neck to withstand the high bending stresses imposed by climbing and leaping. The human femoral neck (*middle*) has only a thin layer of bone at the top. It is suited only to the stresses of upright walking and running, when the abductor muscles counteract tension on the top of the neck. A fossil femoral neck from a contemporary of Lucy (*right*) has the same structure as the human one; it was designed exclusively for bipedal walking.

The answer lies in the action of muscles that operate only in bipedal locomotion: the abductors. These muscles have lines of action that are not vertical but are sharply inclined, which makes them roughly parallel to the femoral neck. When they contract, they push the femoral neck into the hip socket, compressing the neck along its length. This compressive stress combines with the stresses that result from bending (tension on the top of the femoral neck and compression on the bottom). The effect is to eliminate tension at the top of the femoral neck and create a gradient of increasing stress running from the top of the femoral neck, where stress is now minimal, to the bottom, where stress is very high but purely compressive. The bottom of the human femoral neck has a robust layer of solid bone, and even the porous bone that fills in the rest of the section

is reasonably strong as long as it remains under compression.

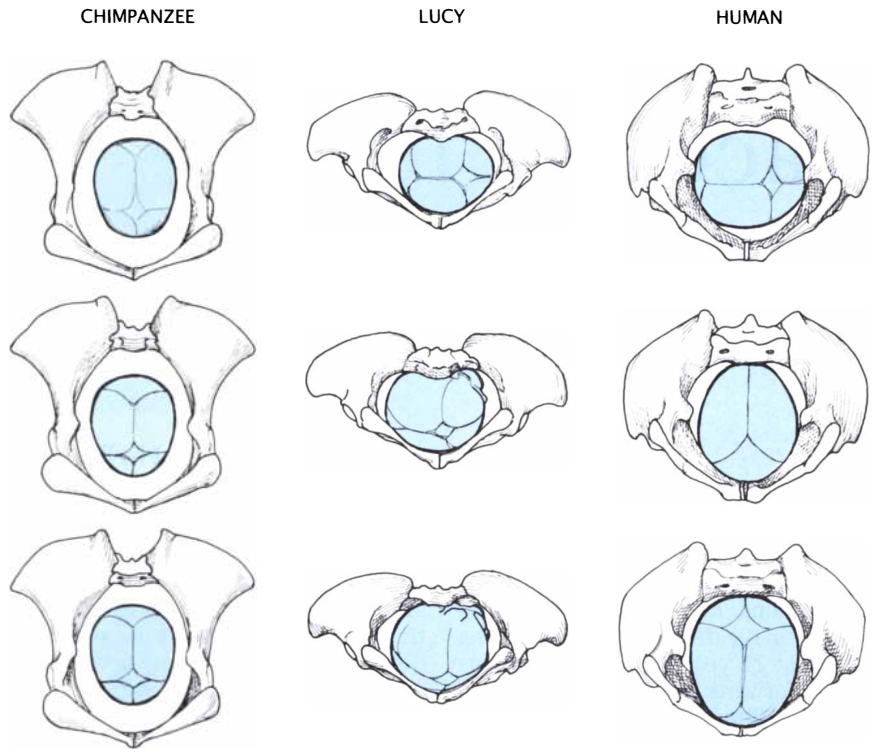
Other muscles work with the abductors to keep the femoral neck under compression when it is loaded. The most important of them is the piriformis, which originates on the front of the sacrum and extends to the outer end of the femoral neck. That orientation enables the muscle to increase the femoral neck's level of compression. The synchronized action of all these muscles when body weight is supported on one leg makes it possible for this seemingly fragile bone to cope with its load.

Because of its distribution of bone, however, the femoral neck is indeed vulnerable if the abductors and other muscles do not act in the proper synchrony. The femoral neck is a primary site of fracture in old age, and not just because bone quality is reduced in old people. These "broken hips" are also a product of reduced muscular coordination. Thus the design of the human femoral neck requires the muscular action of bipedal walking. The bone is poorly engineered for climbing and arboreal acrobatics, where it would be frequently subjected to bending stresses without being compressed at the same time by the abductors.

The femoral neck in *Australopithecus,* because it was even longer than that of modern humans, was subject to even greater bending stresses. If these human ancestors had often taken to the trees, stressing their femoral neck without coordinated compression by the abductors, the bone would have had to have been even more robust than it is in the apes. Was it? The same site where Lucy was found also yielded several femurs that had broken during their long burial, affording a view of the neck's internal structure. Each specimen clearly shows the human feature of thin bone on the upper part of the femoral neck. Lucy's femoral neck, then, was suited exclusively for bipedality. She was not just capable of walking upright; it had become her only choice.

I have concentrated on the pelvic anatomy of Lucy because the hallmarks of bipedality are so vivid there. A review of the rest of her skeleton and of other *Australopithecus* skeletons would reveal equally dramatic modifications that favor bipedality and rule out other modes of locomotion. The knee, for example, is adapted for withstanding greater stress during complete extension than the knee of other primates, and its design brings the femur and the tibia together at a slight angle, so that the foot can easily be planted directly under the body's center of mass when body weight is supported on one leg. The ankle is also modified for supporting the entire body weight, and a shock-absorbing arch helps the foot to cope with the added load. The great toe is no longer opposable, as it is in quadrupedal apes, but runs parallel to the other digits. The foot is now a propulsive lever for upright walking rather than a grasping device for arboreal travel. The arms have also become less suited to climbing: both the limb as a whole and the fingers have grown shorter than they are in the apes.

Lucy's ancestors must have left the trees and risen from four limbs onto two well before her time, probably at the very beginning of human evolution. I have suggested an explanation of why bipedality, with its many disadvantages, appeared long before our ancestors could have put their freed hands to use in carrying tools or weapons: it was part of a novel reproductive strategy that included provisioning by the male, a strategy that enabled the first hominids to flourish and diversify. The explanation will continue to be debated, but the evidence is conclusive that this curious form of locomotion was among the first anatomical characteristics to mark the ascent to cognitive life.



BIRTH PROCESS has competed with bipedality in shaping the modern human pelvis. In the chimpanzee pelvis (shown from the back) the head of the fetus descends without difficulty through the inlet (*top*), midplane (*middle*) and outlet (*bottom*) of the birth canal. In Lucy the birth process was somewhat more difficult: her short, flaring ilia were well suited to bipedality but resulted in a birth canal that was broad but constricted from front to back. Her infant's cranium could pass through only if it was first turned sideways and then tilted. The much larger brain in the human infant demands a rounder birth canal. The necessary lengthening of the pelvis reduced the flare of the ilia and hence the mechanical advantage of the abductor muscles; even so, the human birth process is complex and traumatic, requiring a second rotation of the fetal cranium within the birth canal. The illustration is based on one by Robert G. Tague of Louisiana State University and Linda Budinoff of Kent State.

FURTHER READING

HUMAN WALKING. Verne T. Inman, Henry J. Ralston and Frank Todd. Williams & Wilkins, 1981.

LUCY: THE BEGINNINGS OF HUMANKIND. Donald C. Johanson and Edey Maitland. Simon and Schuster, 1981.

THE ORIGIN OF MAN. C. Owen Lovejoy in *Science,* Vol. 211, No. 4480, pages 341–350; January 23, 1981.

THE OBSTETRIC PELVIS OF A.L. 288-1 (LUCY). Robert G. Tague and C. Owen Lovejoy in *Journal of Human Evolution,* Vol. 15, No. 4, pages 237–255; May, 1986.

Obstacles to Developing Vaccines for the Third World

Six vaccines are already in use there. Many others could be produced within 10 years. Yet those who have the know-how to make the needed vaccines have lacked incentives to apply it

by Anthony Robbins and Phyllis Freeman

Global immunization programs sponsored by the United Nations have made astonishing strides in the developing world during the past 25 years. Unfortunately this rapid forward movement is in danger of stalling. The loss of momentum would be tragic, given what has been achieved so far. The programs totally eradicated smallpox a decade ago and more recently have significantly reduced the annual death and disability caused by six major diseases: measles (which still kills two million children in the Third World every year), diphtheria, pertussis (whooping cough), tetanus, polio and tuberculosis. In 1974 an estimated 5 percent of the children in the developing world were being immunized against those six diseases; today, thanks to the World Health Organization's Expanded Program on Immunization (EPI), almost 50 percent of the children are completely immunized—that is, they receive all recommended doses of all six vaccines. The EPI's aim is to reach every child by 1990.

ANTHONY ROBBINS and PHYLLIS FREEMAN first began studying the obstacles to the development of new vaccines in 1984, while they were working for the U.S. Congress. Robbins is professor of public health at the Boston University School of Medicine. He has a bachelor's degree from Harvard College (1962), an M.D. from the Yale University School of Medicine (1966) and a Master's of Public Administration from Harvard University (1969). Freeman is associate professor and chairman of the Law Center of the College of Public and Community Service at the University of Massachusetts at Boston. She earned her law degree from Northeastern University in 1975 and in 1986 was a scholar in residence at the U.S. National Academy of Sciences' Institute of Medicine.

Yet much remains to be done. In particular, new or improved vaccines are needed for the many other infectious diseases that cause unnecessary death and disability in the developing world, where four out of five children live. Children are the prime target of vaccination programs because they are the chief victims of infectious diseases. In Africa, Asia and Latin America acute respiratory infections (such as the pneumonias caused by *Streptococcus pneumoniae* and *Hemophilus influenzae* type b bacteria and by parainfluenza and respiratory syncytial viruses) and diarrheal diseases (such as infections caused by rotavirus and by *Shigella, Vibrio cholerae* and certain types of *Escherichia coli* bacteria) annually kill some eight million children under the age of five. These infections combine with other diseases, notably measles, malaria, tetanus, meningitis and typhoid fever, to kill an estimated 14 million children younger than five every year and to disable many millions more.

Strangely enough, it is not a lack of scientific know-how that is impeding the development of the needed vaccines. Many are already on the drawing board in laboratories around the world. Indeed, in 1986 the Institute of Medicine of the U.S. National Academy of Sciences identified 19 priority infections for which, from a scientific standpoint, new or vastly improved vaccines could feasibly be produced by 1996 [*see illustration on page 130*].

Rather, the obstacles to the testing, mass production and distribution of the needed vaccines are economic and political. The UN is not equipped to produce the vaccines, and the handful of manufacturers—most of them in the developed world—that have the technological skill and production capacity needed for making human vaccines have little interest in that task.

During the past four years we have reviewed the vaccination needs of the Third World and have conferred with public health officials in many developing nations as well as with leaders of the EPI and most of the world's vaccine manufacturers. Our survey indicates that the current dilemma can be resolved. The solution requires, however, that the public and those involved in vaccine development and procurement gain a full understanding of why the EPI has been successful in the past, what the current challenges are and where there is room for cooperation between those who have and those who need.

Although finding a way to produce and deliver the vaccines required by the Third World will not be easy, it must be done. Vaccination programs are less expensive, easier to implement and sometimes more effective than other public health approaches now being pursued to varying degrees. For example, one could prevent many infectious diseases by controlling the factors that promote their spread, such as a high prevalence of disease-carrying insects (vectors), overcrowded living conditions and unsanitary sources of water for drinking and cooking. Yet vector-control programs, including those aimed at eliminating the mosquitoes that carry malaria, often are not completely successful. Moreover, many countries cannot yet afford the capital investment needed to improve housing, sanitation and drinking-water purity.

Relying on the delivery of effective treatments to the Third World is similarly problematic. Such drugs as antibiotics are highly effective against certain diseases or their secondary complications, for instance against

shigellosis (which results in an estimated 650,000 deaths per year, primarily in children) and against the bacterial infections that often accompany infection by the respiratory syncytial virus. In addition oral rehydration therapy, which replaces the fluids and salts lost in diarrhea, is an inexpensive and easy way to counter many diarrheal diseases; it is necessary only to mix the contents of a prefilled packet with boiled water and feed the solution to the child.

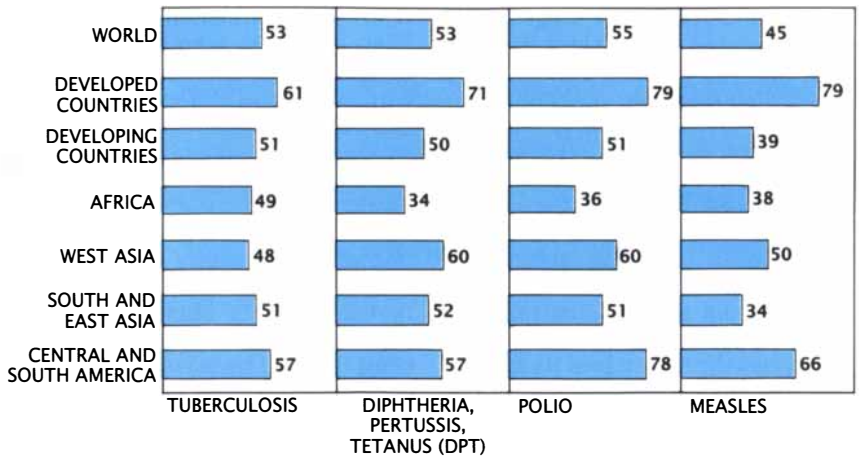
Yet antibiotics and various other drugs are expensive, in part because many more doses are required than is the case with vaccines. Furthermore, the medical-care services that are necessary to identify sick children, make laboratory diagnoses and provide appropriate treatment have yet to be established in many developing nations. Also missing in many areas are massive public-education programs that could teach families to apply therapies that require little medical skill, such as oral rehydration therapy.

The difficulty of implementing vector-control, sanitation, housing and treatment programs highlights another reason increasing the range of vaccines available to the Third World is an attractive approach to decreasing death and disability: an infrastructure—the EPI—is already being built for delivering vaccines to children. The infrastructure could readily be called on to deliver new vaccines for additional diseases. In the future the same infrastructure can also be extended to support other services that are important to child survival, such as family planning.

How has the EPI attained its impressive successes, and why can it not expand its vaccine arsenal by simply following the same strategy? To answer the first question, the EPI has been successful in part because, after some initial duplication of effort, the WHO has collaborated closely with the United Nations Children's Fund (UNICEF), which has a history of serving children in the developing world. Together the organizations have helped the public health authorities in the participating nations to set up systems for obtaining equipment, training the personnel needed for vaccine delivery and ensuring that vaccines reach their destination in good order. These two UN agencies have been further supported since 1984 by the World Bank, the UN Development Program and the Rockefeller Foundation. In 1984 all five organizations formed the Task Force for Child Survival to bolster funding for and increase the stability of the EPI's global immunization activities.

The administrators of the EPI also found a way to obtain vaccines at unusually low prices: they chose vaccines that had been sold in the developed world for many years. The manufacturers had long since recouped their research and development costs, and so they could consider selling their products at close to the cost of production. We limit the use of the word research to the discovery of a candidate vaccine. Development is the costly and time-consuming effort that includes producing small but high-quality batches of a candidate vaccine for testing, running clinical trials to demonstrate safety and effectiveness and to determine appropriate doses, meeting licensing requirements and generating the capacity to produce

CHILD BEING IMMUNIZED in China is one of millions in the developing nations who have been protected against six dangerous infectious diseases in the past 15 years as part of the United Nation's Expanded Program on Immunization (EPI). The participating agencies, including the World Health Organization and the United Nations Children's Fund (UNICEF), hope to also provide new or improved vaccines against many other pathogens that pose a major threat in the Third World.



PERCENT OF INFANTS (up to a year old) in various parts of the world who in 1986 were receiving all prescribed doses of the EPI's vaccines against tuberculosis, diphtheria, pertussis (whooping cough), tetanus, polio and measles was estimated recently by UNICEF. The table has been redrawn from one in UNICEF's *State of the World's Children 1988* and is based on data supplied by the WHO, UNICEF and the Population Division of the UN; it excludes China, where vaccination rates are generally higher than they are elsewhere in the developing world. In 1974 fewer than 5 percent of children were thought to be receiving all doses of all six vaccines. Today some 50 percent are completely immunized, saving 1.4 million lives per year.

large, quality-controlled batches for mass distribution. It is at the development stage that many promising vaccines become stalled.

Having recouped its investment does not necessarily lead a company to sell its products at cost, but several corporations and public institutes have been willing to do so for any of a number of reasons. For some manufacturers, such as the government-owned National Institute of Public Health and Environmental Protection in the Netherlands, the reasons include a long-standing commitment to public health; those manufacturers do not expect to earn a profit from vaccines sold either domestically or abroad. Socialist nations that take part also commonly do so for humanitarian reasons and, as can be true for all participating manufacturers, to enhance their image in the developing world. For commercial firms, recognition and appreciation for their role in the EPI can facilitate other business activities, such as the sale of more profitable pharmaceuticals or the purchase of raw materials.

Many manufacturers also cooperate because they find the enterprise profitable. Selling large quantities of long-standing vaccines to the EPI can be remunerative even if the selling price is relatively low, because manufacturers who have the capacity to produce more vaccine than they sell at home or in other markets in the industrialized world realize economies of scale. Those manufacturers can increase their production and do quality-control tests with existing personnel and equipment, thereby lowering the cost of each unit of vaccine produced. By selling vaccines to the EPI at close to cost, they recover a substantial fraction of their production costs and earn a greater profit on the higher-priced packages they sell in commercial markets.

Commercial wisdom holds that it is inefficient to produce a vaccine for a population base of fewer than 40 million. As a result certain manufacturers whose domestic markets are small, such as Smith Kline RIT of Belgium, have even been willing to expand their production capacity in order to win the EPI's business and so optimize production efficiency.

Although the EPI was able to capitalize on the varying motivations of many vaccine manufacturers to obtain low prices before 1983, in that year it managed to reduce its costs even more by initiating a bidding system. In particular, UNICEF, which buys most of the vaccines administered in Africa and some of those administered in Asia, and the Pan American Health Organization (PAHO), which buys vaccines for Latin America, solicit bids for the individual vaccines from several manufacturers and contract with the lowest qualified bidders. (The WHO itself now focuses more on science, administration and the training of personnel than on obtaining vaccines.)

Since the inception of the bidding program, all the interested manufacturers have been situated in Canada, Europe and Japan. The domestic market in the U.S. is so large that pharmaceutical firms here feel little need to compete elsewhere or to work with the EPI. They have also been discouraged from selling vaccines to the EPI at low cost because Congress has criticized them for not offering similarly low prices in their own backyard.

In 1988 the bidding system is enabling UNICEF and PAHO to buy vaccines at the extraordinarily low average price of five cents a dose, which translates into a cost of about 50 cents for the purchase of all doses of all six vaccines. The EPI spends more than 10 times as much per child on nonvaccine items: transportation, the personnel administering the vaccines and refrigeration equipment that protects the vaccine in a "cold chain."

The current procurement system, then, dealing as it does with long-used vaccines, emphasizes buying surplus goods. It takes advantage of all available incentives for makers to sell their products at, near or even below the marginal cost of production. The system continues to be highly successful and—once the delivery systems are fully in place—should enable the EPI to meet the goal of delivering the six current vaccines to all the children of the world. At that point the EPI will provide more than one billion doses per year.

The current program does face challenges. The EPI and the Task Force for Child Survival are concentrating on solving the problems of how to make inroads into areas where children are not yet receiving the six existing vaccines and how to ensure that vaccines continue to be delivered to each new generation of children after coverage is finally established worldwide. The EPI also constantly confronts the threat of rising prices. Moreover, existing vaccines have some shortcomings. For example, they lose potency when they are exposed to heat, which is a risk in developing countries, where refrigeration is often inadequate. The EPI badly needs vaccines that are heat-stable and ones that would increase the effectiveness and efficiency of the program in other ways, such as a measles vaccine that can be administered at or soon after birth to provide earlier protection.

The further problem of how to provide additional vaccines received little attention until recently, and yet it

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

■ **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. If you don't, 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.

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

Some of your future selections

■ **Islands,** H. W. Menard. "Breathtaking photographs, lavish illustrations...and lively narrative.... It is a book that only Menard, [with his] encyclopedic knowledge of the oceans...could have written." —*Science*

■ **The Timing of Biological Clocks,** Arthur T. Winfree. "Winfree stands unique in the science of biological time-keeping by visiting all fields, from population dynamics...to biochemistry.... Brilliant insights."
—*Nature*

■ **Einstein's Legacy,** Julian Schwinger. "Delightful.... An ideal gift for the curious non-expert." —*Nature*

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, Dept. A, P.O. Box 646, Holmes, PA 19043

PATHOGEN	POTENTIAL EFFECTS	CASES PER YEAR (AND DEATHS)	INDUSTRIAL DEMAND
Dengue virus	Fever, shock, internal bleeding	35,000,000 (15,000*)	Small, (travelers to endemic areas)
Intestinal-toxin-producing *Escherichia coli* bacteria	Watery diarrhea, dehydration	630,000,000 (775,000*)	Small
Hemophilus influenzae type b bacterium	Meningitis, epiglottal swelling, pneumonia	800,000 (145,000*)	Great
Hepatitis A virus	Malaise, anorexia, vomiting, jaundice	5,000,000 (14,000)	Small
Hepatitis B virus	Same as hepatitis A; Chronic cirrhosis or cancer of liver	5,000,000 (822,000)	Moderate
Japanese encephalitis virus	Encephalitis, meningitis	42,000 (7,000*)	Small (travelers)
Mycobacterium leprae	Leprosy	1,000,000 (1,000)	None
Neisseria meningitidis bacterium	Meningitis	310,000 (35,000*)	Some (during epidemics)
Parainfluenza viruses	Bronchitis, pneumonia	75,000,000 (125,000*)	Great
Plasmodium protozoa	Malaria (with anemia, systemic inflammation)	150,000,000 (1,500,000*)	Moderate (travelers)
Rabies virus	Always-fatal meningitis and encephalitis	35,000 (35,000*)	Small
Respiratory syncytial virus	Repeated respiratory infections, pneumonia	65,000,000 (160,000*)	Great
Rotavirus	Diarrhea, dehydration	140,000,000 (873,000*)	Great
Salmonella typhi bacterium	Typhoid fever (with platelet and intestinal damage possible)	30,000,000 (581,000*)	Small (travelers)
Shigella bacteria	Diarrhea, dysentery, chronic infections	250,000,000 (654,000*)	None
Streptococcus Group A bacterium	Throat infection, then rheumatic fever, kidney disease	3,000,000 (52,000*)	Small
Streptococcus pneumoniae bacterium	Pneumonia, meningitis, serious inflammation of middle ear	100,000,000 (10,000,000*)	Small to moderate
Vibrio cholerae bacterium	Cholera (with diarrhea, dehydration)	7,000,000 (122,000*)	Small (travelers)
Yellow fever	Fever, jaundice, kidney damage, bleeding	85,000 (9,000*)	Small (travelers)

PATHOGENS listed here were identified by the Institute of Medicine of the U.S. National Academy of Sciences as ones for which new or greatly improved vaccines are needed in the Third World and could feasibly be developed and licensed by 1996. Many manufacturers, however, are reluctant to invest in vaccine development for the Third World because the process is costly and the prospect is poor for recouping such expenses from sales there. The number of cases and deaths due to each pathogen are estimated based on the institute's 1986 report, *New Vaccine Development, Establishing Priorities.* The asterisk indicates diseases for which children account for roughly half of the deaths or more. Virtually all deaths from dengue fever, parainfluenza, respiratory syncytial virus, rotavirus and pneumococcal meningitis occur in children. Two kinds of vaccines are already being made for hepatitis B.

is crucial. Indeed, it is where creative thinking and new strategies are now needed the most. Not only would new vaccines save countless lives but also their availability would help to sustain the enthusiasm of political leaders in the developing nations for establishing the systems needed to meet the EPI's goals of providing every child with vaccines and maintaining high levels of coverage.

The UN has not totally ignored the problem of new vaccines. The WHO does sponsor research into the creation and testing of new vaccines against infectious diseases in the developing world through five different programs based in Geneva. The programs influence vaccine research all through the world. Yet in 1987 the WHO's combined budget for vaccine research and development was only $10 million per year, much less than the $30 to $50 million the U.S. Institute of Medicine estimates is needed to develop a single vaccine.

The obstacles to the development and distribution of the needed vaccines are many, but among the most important is the fact that the decision to develop new vaccines is left almost entirely in the hands of a few institutes or commercial manufacturers in the developed world. The engineering knowledge and skills are concentrated in these few institutes and firms, several of which have already chosen not to participate in the EPI. Commercial manufacturers in particular cater to the needs of the prosperous, industrialized nations and find it necessary to recover development costs before considering selling their products at close to the cost of production.

For instance, vaccine producers in the U.S., Canada, Western Europe and Japan are focusing a great deal of attention on improving the diphtheria-pertussis-tetanus (DPT) vaccine's pertussis component, which in rare instances causes seizures or brain damage. A new vaccine is certain to be expensive—from 10 to 100 times the current price to the EPI. The extra expense, if the EPI could afford it, would not increase the vaccine's efficacy, reach new populations or offer protection against new diseases.

In spite of this bleak picture, the old strategy of waiting for research and development costs to be recovered and then buying in bulk at a steep discount may still work for those vaccines that have a large market in the developed world or that have some market for travelers. There may be a long wait, however. When

For the cost of a morning coffee break, you can break the cycle of poverty for one small child.

It may not seem as though 72¢ will go very far these days. But for a small child like this, it can lead her out of desperate poverty and into a future of hope.

And right now, you can reach out and make that future happen...by sponsoring one child overseas through Foster Parents Plan.

For only 72¢ a day, you can make it possible for your Foster Child and family to have adequate food and clothing, a decent place to live, needed medical care, education...*and the chance for a better life.*

Because while you're helping, the child's own family will be working with Foster Parents Plan to develop the skills and resources they need to help themselves.

You'll be amazed at the difference you can make. And you'll see the difference for yourself.

In photographs. Detailed progress reports. And heartwarming letters from your Foster Child. You'll be doing so much for that child...and for so little.

Won't you help now? Simply mail the coupon today or call toll-free. It could be the biggest break a child ever got.

Foster Parents Plan.®

Help so complete, it touches a child for life.

Founded in 1937, Foster Parents Plan is the leader in combining family assistance programs, community development, and personal communications between Foster Parent and Foster Child. These comprehensive programs provide long-term solutions to the unique problems facing Foster Children and their families. As a non-profit, non-sectarian and non-political organization, we depend on Foster Parents to make our work possible. A copy of our financial report, filed with the New York Department of State, is available on request. Your sponsorship is tax-deductible.

To start helping even faster, call toll-free:

1-800-543-2400

Detach and mail or call toll-free today. (N6)

Send to:
Foster Parents Plan, Inc.
P.O. Box 877, Dept. K, Providence, RI 02901

I want to become a Foster Parent to:

☐ **The child who has been waiting the longest,** or as indicated:

☐ Boy ☐ Girl ☐ Either

☐ Any age ☐ 6–8 ☐ 9–11 ☐ 12–14

☐ Bolivia	☐ Guatemala	☐ Kenya	☐ Philippines
☐ Burkina Faso	☐ Honduras	☐ Liberia	☐ Sri Lanka
☐ Colombia	☐ India	☐ Mali	☐ Sudan
☐ Ecuador	☐ Indonesia	☐ Nepal	☐ Thailand

☐ Enclosed is a check for $22 for my first month's sponsorship of my Foster Child. Please send me a photograph, case history, and complete Foster Parents Sponsorship Kit.

☐ I am not yet sure if I want to become a Foster Parent, but I am interested. Please send me information about the child I would be sponsoring.

☐ Mr. ☐ Mrs.
☐ Miss ☐ Ms.

Address Apt.#

City State Zip J 345

WATERWAY in East Pakistan has been identified as a source of cholera. Throughout the developing world, water taken for drinking or cooking from lakes, streams and rivers contaminated by body wastes and other sources of infectious organisms facilitates the spread both of cholera and of such diseases as hepatitis, rotavirus, shigellosis and typhoid fever. Poor sanitation, overcrowding and an abundance of disease-carrying insects (vectors) also promote disease. Programs to purify drinking-water supplies, improve sanitation and living conditions and eliminate vector populations can prevent much disease, and treatment programs can often prevent disability and death, but the authors note that for the short term the expansion of immunization programs is the least expensive and easiest way to save lives.

new vaccines come on the market, they tend to be expensive, not only because of the need to recoup standard development costs but also because new—and hence costly—technologies have often been employed to produce them. The new technologies are not always crucial. Many times they are chosen in preference to an older technology because they may result in a vaccine that has somewhat fewer side effects or is slightly more effective, but they often do so at the public health cost of making the product inaccessible to the Third World.

One product that is already on the market is a vaccine against hepatitis B. In fact, there are two types of hepatitis B vaccines: both consist of antigens (fragments of an organism that elicit the production of protective antibodies) from the virus, but one type is produced by genetic engineering and the other is made with an older technology that relies on isolating antigens from the blood plasma of infected individuals. The price of one such plasma-based vaccine, made in Korea with a technique developed in New York, is now about $1. This is still eight times costlier than the most expensive vaccine (measles) now bought by the EPI, but if the vaccine meets the WHO's standards of safety and efficacy, it may be adopted by the EPI as it waits for the prices of the potentially superior genetically engineered vaccines to drop significantly. The first of the genetically engineered vaccines to be licensed in the U.S. costs as much as $130 for a three-dose series.

Vaccines for AIDS and rotavirus, which are under intensive study because the diseases affect many people in the industrialized nations, may also become available in the future and affordable sometime later. A malaria vaccine too seems possible, since the demand from travelers living in industrial countries may well combine with the very large population exposed in the Third World to encourage commercial development. Other vaccines that may meet the needs of a smaller subset of travelers (those against dengue virus and hepatitis A infections, for instance) are less likely to be produced or, if they are produced, to be affordable to the Third World.

The waiting strategy will certainly not make available vaccines that have no market in the industrial world. As things stand now, several important vaccines that are scientifically feasible, according to the Institute of Medicine, are likely to be neglected by commercial firms for lack of a market in the industrialized nations. Such vaccines include ones against shigellosis, leprosy and infections caused by *Streptococcus pneumoniae* (in infants) and enterotoxigenic *E. coli* (*E. coli* that produce intestinal toxins). The same may be true for improved versions of the vaccines for measles, polio, cholera, typhoid, Japanese encephalitis and yellow fever.

Thus market forces militate against the production of critically needed vaccines for the Third World, and even the vaccines that have a chance of being produced are for the most part unlikely to be affordable any time soon, if ever. In some cases the EPI may have to consider buying potentially inferior products (such as the plasma-based vaccine for hepatitis B instead of a genetically engineered one) because they are the least costly. Even that option is not always open: manufacturers who have employed relatively low-cost technologies to make vaccines may remove their less expensive products from the market when they introduce versions made with a more advanced technology.

The situation is discouraging but not beyond redemption. Indeed, at least four approaches are worth considering. In one approach, which essentially accepts the status quo, the UN would raise the money to buy vaccines at close to their market price (in other words, a price that includes development costs and profits) in the hope that the promise of a new, lucrative market would encourage the industry to make the needed investment in development. For this approach to be successful the EPI would have to predict the number of doses it would buy, something that could probably be done. On the negative side, UNICEF and PAHO would have to commit themselves to buying given quantities at top dollar for a long time—a commitment they probably could not make.

A second option is for the UN to create a public institute to develop and manufacture its own vaccines, bypassing commercial makers. If the institute were equipped with advanced technology and staffed by the finest epidemiologists, molecular biologists, fermentation engineers and other professionals from around the world, it might well produce rapid results.

Such an institute could pursue new technologies that are particularly appropriate to Third World public health needs without having to earn rapid profits. For example, workers might be

able to insert genes for antigens from different bacteria and viruses into a single carrier organism—such as the vaccinia virus that once constituted the smallpox vaccine or the bacillus Calmette-Guérin (BCG), which is used as the vaccine against tuberculosis—to produce a single vaccine capable of eliciting immunity to a wide range of infections. The institute staff might also be able to advance the development of stable cocktails, or mixtures, of antigens that would immunize against several diseases at the same time, thereby improving the ability of the developing countries to provide complete coverage for their children.

As the UN's health programs expand, the possibility that the organization could operate such an institute increases, but the disadvantages of the solution are as tangible as its appeal. The creation of such a center would be both costly and time-consuming, and the politics of international cooperation are brutal. The difficulty the UN Industrial Development Organization (UNIDO) has had in establishing two new biotechnology centers is instructive. In the first five years not a single product has been put forward. Success in vaccine development and production is too critical to the health of the world to rely at the start on an international center alone.

A third alternative is to establish development and production units in countries or regions of the Third World with large populations. Such units would be dedicated to creating, developing and producing vaccines against the diseases that are most prevalent or destructive in the region. The concept is likely to appeal to the participating nations of the EPI, which are eager to develop new technologies. Funding for the transfer of technology is also a traditional development activity and might attract money that would otherwise not be available for vaccination programs.

A major roadblock to this alternative is the fact that the World Bank and other agencies that might provide loans to establish such units are increasingly concerned with meeting banker's standards (seeing a good return on an investment) when they finance new industrial enterprises. The development and production of vaccines to meet the health needs of the Third World is unlikely to turn a profit.

Even if financing can be obtained, the full benefits of this approach will take years to realize. It will take time for any institute to produce a product, time for the institute to address the range of needs in its region and more time still until a network of such institutes can meet all the vaccination needs of every region in the developing world.

The idea is nonetheless worth trying as a long-term strategy, and actually it is being pursued on a small scale by PAHO and the Rockefeller Foundation. Convinced that the aim of investment in regional vaccine production should be advances in public health and science rather than financial gain, PAHO and the Rockefeller Foundation are exploring the creation of one or more centers for developing and producing vaccines in Latin America. If the centers are successful in demonstrating that otherwise competitive nations can cooperate to solve a regional health problem, international-aid organizations may decide to help establish similar centers in Africa and Asia.

There is a fourth approach, which seems to us to be a rapid and short-term strategy for obtaining tangible results before the end of this century. As in the first alternative, the UN would raise money to pay manufacturers, but the funds would be earmarked specifically for the development of particular vaccines deemed important for large parts of the Third World. For many of the same reasons that now lead manufacturers to take part in the EPI, the institutes and companies that develop the products would then sell them to the EPI at or near the cost of production (with no development costs added on). The advantage of this approach over the first one is that a finite amount of money would have to be raised for each vaccine, and the EPI would receive a long-term commitment for a supply of vaccine at a low price.

Raising money for vaccine development is not out of the question. There are already mechanisms in place to fund vaccine research; for instance, the World Bank and the UN Development Program support such research. Those mechanisms would simply have to be extended or copied to support a program of vaccine development.

On the other hand, the UN and other agencies that run international-aid programs would have to make the difficult decision to spend money outside the Third World—in industrial nations, where most manufacturers with good track records in translating laboratory discoveries into full-scale production capacity are currently situated. This would be a necessary concession now, to ensure speedy, consistent and high-quality vaccine manufacture. With time, however, it should be possible to transfer much of the resulting technology to the developing nations so that they can build on any technology they have already acquired and make their own vaccines.

There are signs that the fourth approach may actually be tested soon. Key public and private manufacturers around the world have expressed some willingness to undertake the development and production of EPI vaccines on a contractual basis and then to sell the products at close to cost (although several manufacturers still express concern about whether the volume of purchases would be adequate). For some companies such an arrangement would be an attractive way to explore advanced technologies and have the needed tools subsidized.

There has also been some discussion at the UN of creating a revolving fund through which the development arm of the UN and other aid programs would pay the initial development costs. Under this plan UNICEF and PAHO would pay only the cost of production for the resulting vaccines and a small surcharge that would serve to replenish the fund so that more development could be pursued.

Actually any of the four approaches, or some combination of them, could succeed. The task now is to mobilize international agencies to confront directly and soon the lack of new vaccines, and to convince those who have the needed scientific knowledge and technical skill to apply such assets on behalf of all children—both in the developing nations and in the industrialized ones.

FURTHER READING

NEW VACCINE DEVELOPMENT, ESTABLISHING PRIORITIES, VOL. 2: DISEASES OF IMPORTANCE IN DEVELOPING COUNTRIES. Committee on Issues and Priorities for New Vaccine Development, Institute of Medicine. National Academy Press, 1986.

NEW SCIENTIFIC OPPORTUNITIES AND OLD OBSTACLES IN VACCINE DEVELOPMENT. Kenneth S. Warren in *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 83, No. 24, pages 9275-9277; December, 1986.

ESTABLISHING HEALTH PRIORITIES IN THE DEVELOPING WORLD. Julia Walsh. United Nations Development Program. Adams Publishing Group, 1988.

THE STATE OF THE WORLD'S CHILDREN 1988. James P. Grant, United Nations Children's Fund. Oxford University Press, 1988.

VACCINES. Stanley A. Plotkin and Edward A. Mortimer. W. B. Saunders Co., 1988.

Neural Networks at Work

They watch over factories, credit applicants, sleepy pilots

Watchful networks, trash mashing, genes for sale, physical attractions

On the exhibit floor of the Sheraton Harbor Island hotel in San Diego this past July, spectators attending the second annual International Conference on Neural Networks gathered to watch a toy train whir around an oval track. Overhead a television screen flashed black-and-white snapshots of the cargo: small plastic bottles of Tylenol. A computer-graphics display indicated that one bottle was missing its cap and another had a skewed label.

Appearances to the contrary, the device being demonstrated is no kid stuff. "What you are seeing is the first automated factory inspection system that uses a neural network to classify objects," explained Jim Belilove, the soft-spoken chief operating officer of Global Holonetics in Fairfield, Iowa. While most neural-network researchers continue to play with what critics call "toy problems"—tasks whose value is more theoretical than practical—the company has developed a quality-control application that can face the rigors of the factory floor. "We've looked at candy bars, syrup bottles, sponges, tire treads, batteries—over 100 different inspection tasks," said David E. Glover, Global Holonetics' chief scientist.

The neural network can switch from one task to another easily because it is not so much programmed as it is "trained": instead of listing rules for telling good bottles from defective ones, one simply "shows" it examples. Interconnections among the network's many simple processors, which are linked rather like brain cells, are modified so that a particular input will produce a desired result. The computation is spread over many connections; thus the network relates an input pattern to an output in a statistical rather than exact manner. The network can therefore process incomplete data, such as half of an image, just as the brain can—an analogy that has been overextended, fanning both hype and keen interest in the field.

The key to Global Holonetics' success, however, is less the network itself than it is the novel way the system processes the video-camera image beforehand. It exploits advanced optics—spatial light modulators and lasers—to speedily compress a 65,000-pixel image into 32 numbers containing salient information about an object's shape and orientation.

Neural networks enter at the next stage: assigning the number pattern to a category. Conventional statistical methods may also be able to do the task, but "we needed the system to be simple," Glover said. "With a neural network a person on the factory assembly line doesn't have to be a programmer to train the machine to examine and classify new objects." Several Fortune 100 companies are now testing the system, which will go on sale early next year.

The networks are good not only at image recognition but also at a broad range of tasks that involve classifying variable and complex data. In the early 1970's EEG Systems Laboratory, a Government-sponsored research institute in San Francisco, recognized that adaptive networks might be capable of analyzing electroencephalograms, tracings that combine a large number of electrical signals from different parts of the brain. The laboratory revamped an early network developed at McDonnell Douglas for analyzing satellite images and biomedical data, and by the late 1970's it had a program that could detect subtle EEG patterns that are linked to particular neurological states.

By analyzing EEG's of epileptics who were preparing to undergo surgery to eliminate seizures, the network correctly predicted the outcome of the operation 87 percent of the time. In another experiment the network monitored Air Force test pilots who were carrying out a difficult visual-motor memory task and detected EEG patterns that foretold the onset of fatigue. "We compared the network with standard statistical tests and the network was better," said Steven Bressler, a neurophysiologist at EEG Systems.

If adaptive networks can take brainwave patterns and predict when a pilot will run out of the right stuff, can they also take the pulse of the economy and predict the outcome of business decisions? Nestor in Providence, R.I., is developing a mortgage-insurance underwriter for a large private mortgage-insurance company and an equities-trading program for Morgan Stanley. The mortgage-insurance underwriter not only mimics the best human underwriters but also can be trained on records of past loans to predict which new loan applicants are most likely to default.

Murray Smith, president of Adaptive Decision Systems in Andover, Mass., has developed a system that judges whether a customer is a good or a bad credit risk. The system, commissioned by a financial-services company, was trained on 250,000 case histories and is being tested in the field. An executive at the unnamed company said there was concern that employees would resist the technology because they cannot see how the network makes its decisions, but its advantage is that it can be easily updated to reflect a changing economy.

The most ambitious business application to date may be the Airline Marketing Tactician from BehavHeuristics in Silver Spring, Md. The program helps airlines to find the most profitable way to allocate seats to various fare classes on a particular flight, taking into account a complex pattern of factors, such as flight time, competition, connections and no-shows. A U.S. carrier will test the program on its own data, said BehavHeuristics' chairman, Kenneth R. Stephens.

Another application that is being pursued vigorously is process control for industry. AI Ware in Cleveland installed a network in the chemical plants of a Fortune 500 company. The network analyzed how changes in the ratio of ingredients alter the properties of the product, thereby enabling the company to design new products. GTE Laboratories in Waltham, Mass., is planning to install a network in a fluorescent-tube plant. "We want to find out why the factory works well at certain times and not at others," said Richard Sutton of GTE.

Well over 100 large companies are

now doing active research in adaptive networks, estimates Edward Rosenfeld, publisher of the newsletter *Intelligence,* which follows the field. And the Defense Advanced Research Projects Agency says it will undertake a major research initiative in neural networks. An agency spokesperson said, however, that funding will be less than the $400 million over eight years recommended in a recent agency study—in spite of assistant director Jasper Lupo's pronouncement at the San Diego meeting that neural networks are going to be "more important than the atom bomb."

The products now appearing on the market hardly live up to this level of hyperbole. "The reality today is that neural nets do only somewhat better than statistical approaches," said Smith of Adaptive Decision Systems. But, mindful of the technology's promise, he believes companies should pursue applications of neural networks anyway as "a strategic investment, because someday they'll do things statistics can't do." —*June Kinoshita*

Disappearing Act

Can degradable plastics ease the landfill crisis?

An old saying from the north of England has it that "where there's muck there's brass." Inventive commercial solutions to the nationwide shortage of landfill capacity seem to be proving the wisdom of that maxim. Currently the U.S. produces 148 million tons of municipal waste per year, and the total is expected to increase another 20 percent by the year 2000. As the garbage mounts, previously uneconomical ways to reduce it are becoming more attractive.

Plastics, which account for about 7 percent of municipal waste and are its fastest-growing component, are a case in point. Current plastics do not degrade readily. One way to make a degradable plastic is to mix a small proportion of starch—typically less than 10 percent—with a conventional polymer. Bacteria and fungi can then break down the plastic into dust.

Amko Plastics in Cincinnati is making degradable shopping bags from such a plastic. Ampacet in Mount Vernon, New York, is producing similar resins that are used to make degradable trash bags. Agri-Tech Industries in Urbana, Ill., is making degradable plastic films through a different technology, developed by the Department of Agriculture. The recipe calls for from 20 to 80 percent starch, together with ethylene acrylic acid copolymer and other ingredients.

Plastics containing starch do not degrade completely. The British company ICI, through its subsidiary Marlborough Biopolymers, is producing experimental quantities of a completely biodegradable plastic made from a biological polymer known as PHBV. The polymer is harvested from bacteria that have been fed a special sugar. The material is also suitable for surgical implants. Although the process is chemically elegant, it is costly: the product now sells for $15 per pound.

EcoPlastics in Toronto relies on sunshine instead of biology to break down its product. It has developed chemical resins that can be mixed with the major types of plastics to create a product that slowly degrades outdoors, on exposure to ultraviolet light. Anthony E. Redpath, president of EcoPlastics, acknowledges that photodegradable plastics are not likely to mitigate the landfill crisis, but they could reduce litter. The company hopes to see its resin used in such products as candy wrappers and soft-drink containers.

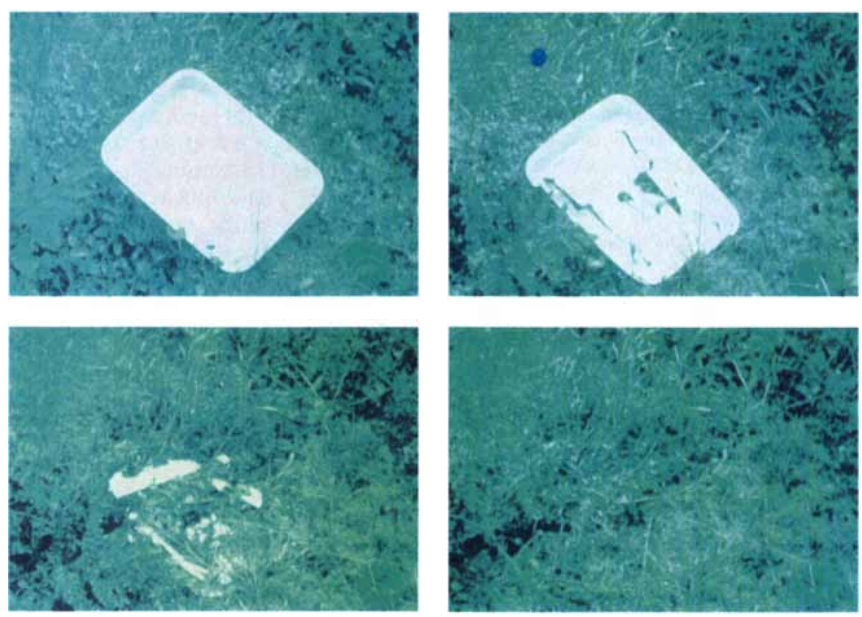
Government is encouraging such technologies: many jurisdictions are passing laws banning certain uses of nondegradable plastics. The Society of the Plastics Industry says it does not oppose the introduction of degradable plastics, but it points out that even supposedly biodegradable materials such as paper persist for decades in landfills. Degradable plastics are therefore unlikely to be a panacea. Moreover, Margaret Rogers, a spokeswoman for the society, points out that biodegradable plastics are likely to attract bacteria. Hence they may not be suitable for uses that put them in contact with food.

Many observers think recycling will make a bigger contribution to solving the landfill crisis. The recycling of plastics has made much less progress than paper or metal recycling: plastic is less valuable than metal, and separating plastics by type—generally necessary for recycling—is difficult. But two-liter soda bottles made of a plastic called PET are already extensively recycled by companies such as St. Jude Polymer in Frackville, Pa. St. Jude is preparing to expand its business to other plastics. Research at Rutgers University now suggests that waste plastics need not always be separated before recycling: "commingled" plastics can be turned into useful planks and boards. Ironically, Rogers points out, if degradable plastics become very common, they could eventually create a problem by weakening recycled plastic. —*Tim Beardsley*

Gene Therapy in Gestation

Three companies hope to turn it into medical reality

No doctor practices gene-transplantation therapy today. That situation may change within the next five to 10 years, if three new bio-



Sunlight attacks a meat tray made from EcoPlastics' photodegradable resin in time-lapse photographs spanning about 60 days. Such plastics could reduce litter.

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.

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. The date of the last issue on subscriptions appears at the right-hand corner of 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

technology companies have their way.

Gene therapy aims to treat genetic defects and some life-threatening diseases that have other causes by weaving a new gene into a patient's genetic makeup. "The science has reached the point where we think it's going to work," says Douglas J. Jolly, director of research at Viagene in San Diego, Calif., one of the recently formed companies. Gene therapy may eventually make it possible, for example, to treat infants suffering from severe immune deficiencies or people with sickle-cell anemia by removing bone-marrow cells, inserting the normal genes and returning the cells to the patient, where they can supply the proteins that were missing.

But even though early results look promising, workers are still far from simple gene-therapy programs. "It's hard to find basic research money to do that kind of thing and it's way too basic for large companies," Jolly adds. Consequently the research has moved to the domain of start-up companies, namely Viagene, Somatix in Cambridge, Mass., and Genetic Therapy Inc. (GTI) in Gaithersburg, Md.

Investigators at all three companies are trying to develop improved ways of using retroviruses to deliver normal genes to bone-marrow cells. (A retrovirus copies the genes it carries into the host cell's genetic material; hence a retrovirus bearing a therapeutic gene can serve as a vector for introducing the gene into a patient's cells.)

Each company is also pursuing specific applications of gene therapy. For example, Somatix, cofounded by Richard C. Mulligan of the Whitehead Institute for Biomedical Research, is exploring ways to incorporate specific genes into tissue to make it continuously secrete a therapeutic agent such as Factor VIII, a blood-clotting protein that is absent in hemophiliacs.

GTI, which is concentrating on inserting therapeutic genes to fight a range of cancers, differs from the other two companies in that it is affiliated with a Government laboratory—the National Heart, Lung, and Blood Institute—rather than a university. The unusual arrangement follows directly from a two-year-old law that allows national laboratories to commercialize their inventions by joining with private companies; although the Government holds all patent rights to jointly developed inventions, the industrial partner has an exclusive license to market the product.

The law came at just the right time for Wallace H. Steinberg, head of the venture-capital firm Healthcare Investment in Edison, N.J., who had long hoped to start a company devoted to genetic therapy. He helped to arrange a Cooperative Research and Development Agreement (CRADA) between the research institute and the fledgling GTI. W. French Anderson, a laboratory chief at the institute, became the principal investigator for the partnership. Even though Anderson remains a full-time Government employee and draws no income from GTI, his staff will work closely with it.

In October, for example, the National Institutes of Health was scheduled to review an application by Anderson and his colleagues to conduct the first genetic-manipulation trials. They have proposed to insert a marker gene into cancer-fighting white blood cells in order to track the proliferation of the cells in cancer patients, which will indicate the progress of therapy. GTI investigators are likely to help run the tests, Anderson says.

In spite of the commercial activity, questions dog gene therapy. Critics continue to challenge the ethics of modifying a patient's gene structure. Workers counter that because they are not tampering with genes that are passed on in sperm and eggs, gene therapy differs little from administering drugs. Nevertheless, if the techniques prove to be effective and safe, Morgan says, gene therapy could signal "a redesign in the nature of health care." *—Elizabeth Corcoran*

Body Shop?

Advertising's infatuation with physiology has cooled

"Arousal by the communication," reads the titillating if somewhat misleading heading in an industry reference manual. "Arousal" here refers to the emotional stimulation of consumers, and the entry goes on to list the spectrum of physiological measures by which advertisers have tried to quantify it: blood pressure, heart rate, body temperature and so on. These techniques, the manual's editors noted in 1978, have attracted interest because of the "aura of scientism" that surrounds them. In this decade the aura has evaporated.

The advertising industry's affair with psychophysiology began in the 1950's, when psychoanalytic theory sparked the imagination of marketers. A 1954 book on marketing strategy lamented "how naive we have been in basing marketing research on the assumption that human beings always

are able and willing to reveal their true feelings." Thus began the search for objective means of testing consumer response, particularly some involuntary reaction that would reveal in the laboratory what one paper called "the true state of the organism."

One of the first such measures was pupil dilation, which Eckhard H. Hess of the University of Chicago had established in the early 1960's as a basis for distinguishing "differential interest" (see "Attitude and Pupil Size," by Eckhard H. Hess; SCIENTIFIC AMERICAN, April, 1965). Marketers greeted the technique with enthusiasm, although reports from a 1967 American Marketing Association conference showed that it was a considerably more sensitive gauge of male responses to scantily clad women than of consumer tastes in advertising. Brain waves, voice pitch, respiration rate and galvanic skin response (the electrical resistance of the skin) came into vogue at about the same time. There was talk in the 1970's of adding lie detectors to the battery of marketing-research tools.

Marketers also tried to exploit physiology in advertising design. A 1975 article in *Journal of Advertising Research* put forth the idea that ads appealing to the right hemisphere of the brain—that is, ads having primarily emotional content—may be more effective on the left side of a page, whereas "advertisements stressing the scientific aspects of a particular brand would be better placed on the right side"—where, presumably, they would be more likely to engage the brain's analytical left hemisphere. A 1977 article that appeared in *Journal of Advertising* encouraged advertisers to tailor their copy to people having a particular physique.

In this decade, after 20 years of research, the advertising industry still cannot point to a correlation between physiological indicators and sales. Small wonder that confidence in psychophysiology has plummeted. "The methods produce interesting results," says Jay Dean, a vice-president at the New York advertising agency of Young & Rubicam, "but nobody really knows what to do about them."

Only one technology retains any credence in advertising research: an eye-tracking device that plots not arousal but how a person looks at an ad. Indeed, many people now berate the "petri dish" approach to advertising. Notes a planner at Chiat/Day, another New York agency, "A person's eyeballs are no more or less articulate than the person himself." —*Karen Wright*

THE AMATEUR SCIENTIST

The café-wall illusion, in which rows of tiles tilt that should not tilt at all

by Jearl Walker

The café-wall illusion involves a checkered arrangement of black and white tiles, rectangular or square, separated by thin lines of mortar. What is surprising is that the horizontal lines appear to converge toward the left or the right. The direction of convergence alternates: in one row the tiles seem to be wider on the left side and in the next row they seem to be wider on the right.

The illusion came to the attention of vision scientists in the early 1970's, when it was discovered on the wall of a café in Bristol, England. In 1979 Richard L. Gregory and Priscilla Heard of the University of Bristol reported the first detailed study of the illusion. They related it to a much older illusion, known as the Münsterberg figure for the psychologist Hugo Münsterberg, who wrote about it in 1897. In the Münsterberg version too the top and bottom edges of individual tiles appear to tilt and the direction of tilt alternates from row to row, but the convergence is weaker. The white regions in particular are noticeable, seeming to flare vertically at one end. In both figures the tilt is an illusion: neither has any tilted elements. You can eliminate the illusion of tilt by sighting almost along the plane of the page. (You may see the accompanying illusions better if you slip a sheet of dark paper under the page to mask the printing on the other side.)

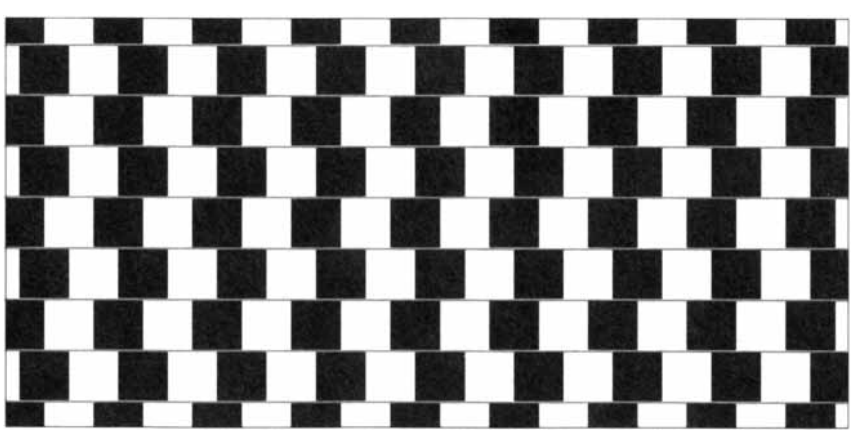
Gregory and Heard, and others after them, noticed that the café-wall illusion is pronounced when the black and white tiles contrast sharply in brightness and the layer of mortar is narrow and of intermediate brightness. When the mortar is just as bright as the white regions, the illusion is weaker or even missing; if it is even brighter, the illusion is certain not to appear. When the mortar is just as dark as the black regions, the arrangement produces the Münsterberg illusion, with its weaker convergence; if it is even darker, there is no convergence. The tiles can be colored instead of black and white, but unless the colors contrast in brightness, the illusion is lost. Identical tiles must be staggered in the array, but the illusion does not appear if they are offset so precisely that they form a chessboard.

Gregory and Heard likened the array to certain patterns invented by James Fraser in the early part of this century. Fraser showed how a narrow row of slightly tilted bright and dark lines gives the illusion of tilt to the row itself. The illusion is called the twisted-cord effect because each row of lines looks something like two cords spiraling around each other. The twisted-cord effect has never been satisfactorily explained. It probably has to do with the fact that the visual system's detectors responsible for determining small-scale (or "local") orientations may influence the perception of large-scale (or "global") orientations. The influence seems to be stronger when many identical, or only slightly different, locally tilted elements are viewed.

Is the tilt perceived in the café-wall illusion due to that kind of control of global orientation by local orientations? If it is, what are the locally tilted elements? The offset of the black and white tiles in adjacent rows might seem to serve the purpose, but the tiles are too large—and besides, such an explanation does not account for the requirement that the mortar supply a particular degree of contrast. For a long time investigators tried to identify some local tilt that could set up the global convergence of the rows.

Gregory and Heard suggested that the local tilt is generated when the visual system locates the borders of a mortar line. To counteract the incessant slight eye movements that shift borders over the retina, some mechanism must lock in the positions of the borders so that the pattern appears to be stationary. Where the mortar separates tiles of the same color (either white or black), the locking mechanism fixes the borders properly. But where the mortar separates contrasting tiles and when the mortar is intermediate in contrast, the locking mechanism may be less accurate, with the result that the borders are displaced slightly. Along the length of a mortar line the visual system then senses a periodic variation in just where the borders of the mortar are, and the line looks like a twisted-cord array.

In 1979 Bernard Moulden and Judy Renshaw of the University of Reading tackled the Münsterberg illusion with an idea from Hermann von Helmholtz, one of the 19th century's pioneers in vision research. When a bright white region adjoins a dull black one, their border is perceived as being shifted into the black region, an effect called irradiation. Such edge migration could account for the apparent vertical flaring of the end of a white region that is



A café-wall array

surrounded on three sides by black tiles. The opposite end, with white tiles above and below it, does not undergo this perceptual widening. The variation in the positions of the top and bottom edges of the white tiles thus creates local tilt.

Still another approach was taken in 1983 by Mark E. McCourt of the University of California at Santa Barbara. He considered an effect called brightness induction, in which the horizontal variation of bright and dark tiles in the café-wall array subtly alters the apparent brightness along the horizontal mortar lines. The mortar between dark sections of tile seems to brighten slightly and the mortar between bright sections seems to darken. The apparently darkened mortar connects two dark tiles along a slant, and the apparently brightened mortar connects two bright tiles along a parallel slant. The region along a mortar line then resembles a twisted-cord array.

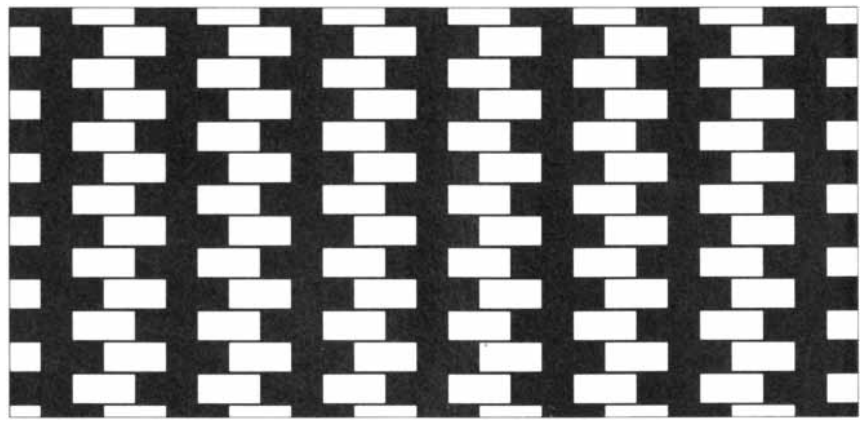
A supporting experiment revealed that if you darken and brighten the mortar lines so that they actually vary periodically in brightness, as in the bottom illustration at the right, the café-wall illusion is stronger than normal. Also, if the mortar is replaced with bright and dark lines that actually slant between tiles of equal brightness, the illusion is stronger still.

M. J. Morgan of University College London and Moulden may have hit on the best explanation of the café-wall illusion in 1986. They relied on a vision model that has been gradually crafted in the past 25 years. Particularly influential in the development was the late David Marr, who worked at the Massachusetts Institute of Technology. I shall first outline a simple version of the model, concentrating on the retinal processing and leaving out the mathematical details, and then return to the illusion.

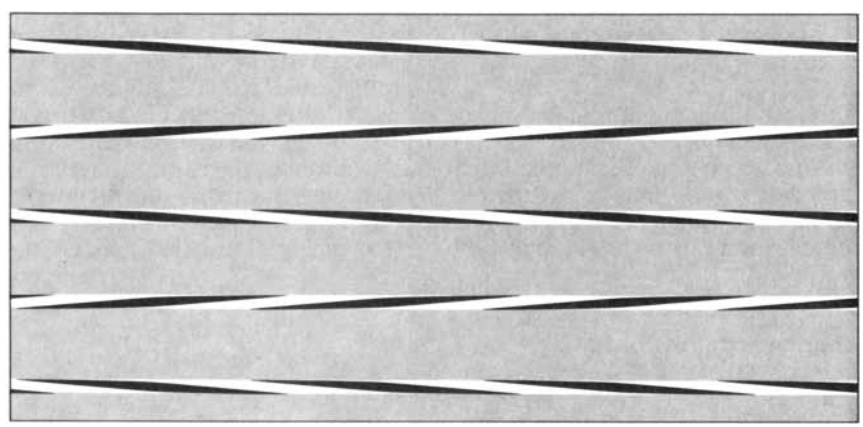
Marr argued that the early stages in the visual pathway construct a "raw primal sketch" of the reality being viewed. The sketch lacks the richness of the actual scene but has blobs, bars, terminations and segments of edges that roughly map real objects. To construct the sketch, the early visual processing by the retina and the brain identifies the sharp changes in brightness that correspond to edges.

The processing begins in the retina when photoreceptors absorb photons and then send signals through bipolar cells to ganglion cells [*see top illustration on next page*]. Each ganglion cell responds to a particular set of photoreceptors that are spread over a small circle, the cell's receptive field. (The illustration shows only a few of the photoreceptors.) The field is partitioned into a center and a surround that respond antagonistically to determine the net response of the cell when the field is illuminated.

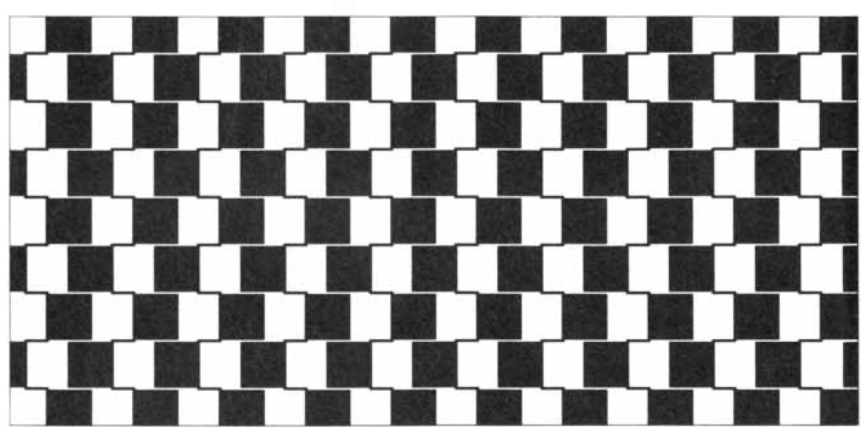
The strength of a signal from a ganglion cell is the rate at which it fires impulses deeper into the visual system. An unilluminated cell fires slowly, at its "resting rate." Suppose only the center photoreceptors are illuminated. One type of cell, called an on-center cell, increases its firing rate while another type, called an off-center cell, turns off. If only the surround is illuminated, the changes are just the opposite: an on-center cell turns off while an off-center cell fires faster than its resting rate. If both the center and the surround of either type of cell are illuminated equally, the cell may



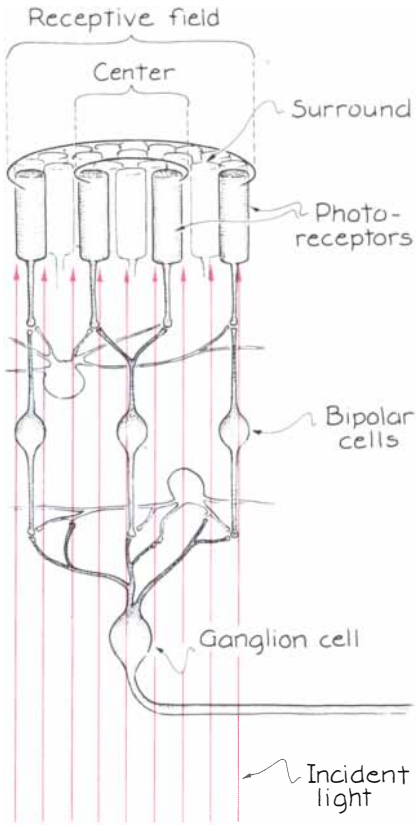
The Münsterberg array



The twisted-cord array



A modified café-wall array



A simplified diagram of the retina

fire just a little more frequently than its resting rate.

Suppose an edge separating bright and dark regions falls across the retina [*see illustration below*]. In the even illumination on the bright side, both on-center and off-center cells fire either at their resting rate or somewhat faster. On the dark side, both types of cells fire at only their resting rate. The cells that straddle the edge are more active. Any on-center cell whose center is in brighter light than one side of its surround fires vigorously; so does any off-center cell whose center is in dimmer light than one side of its surround. The signals from the two types of cells are sent separately along the visual pathway but are eventually brought together when oriented edges are added to the raw primal sketch. If the on-center cells are active in one place on the retina and the off-center cells are active nearby, there must be an edge between the two sets—and so an edge is added to the sketch.

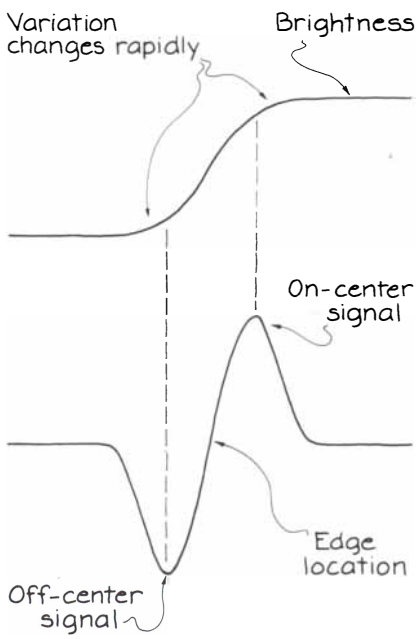
Marr contended that the output of a ganglion cell is a measure of how rapidly the variation in brightness changes across the receptive field. If the illumination is uniformly bright or dark or changes gradually, the cells fire at their resting rate or only slightly faster. But if the variation in illumination changes abruptly, as it does at an edge, certain cells are much more active, namely the on-center cells on the bright side of the edge and the off-center cells on the dark side. When you see a narrow dark line, two edges are detected and the interior of the line activates off-center cells, whereas just outside the line on-center cells are activated. When you see a narrow bright line, the distribution of activity is the opposite.

Morgan and Moulden applied the Marr model to the café-wall and Münsterberg illusions. Consider the café-wall illusion when the mortar has intermediate brightness. The uniformly bright or dark interiors of the tiles generate weak ganglion signals. These regions are gray in the illustration at the right at the bottom of this page. Stronger signals come from the edges of the tiles and the mortar. Where the mortar separates bright sections of tile, the relative darkness of the mortar activates off-center cells. The line of off-center activity connects with the edges of images of dark tiles, where there are also strong off-center signals. These regions are black in the illustration. Where the mortar separates dark sections of tile, the relative brightness of the mortar activates on-center cells. The line of activity connects with the edges of images of bright tiles, which also generate strong on-center signals. These regions are white in the illustration. (Do not take the gray, black and white of the illustration literally. The shading is meant to represent signal strengths; the true shading that you perceive is probably added at a later stage of the visual system.)

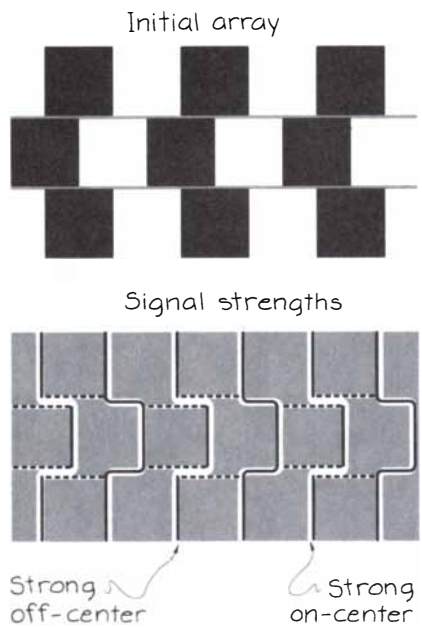
Note that in the illustration the regions of strong on-center and off-center signals form a repeated pattern that resembles two or three steps. The steps climb to either the left or the right, forming a slightly tilted composite. The tilt means that when edges are added to the raw primal sketch, they will be slightly tilted along the mortar line. The repetition means that the local tilt generates a global tilt, as in the case of the twisted-cord array. The illusion works not because there is local tilt in the array but because the processed signals make the raw primal sketch of the mortar line look as if it came from tilted elements.

You may find the explanation to be a long reach, but it accounts quite neatly for the illusion's dependence on the brightness of the mortar. If the mortar is just as bright as the bright tiles, the off-center signal from where the mortar separates bright tiles is missing, and the resemblance between the ganglion-processed patterns of the array and a twisted-cord array is weaker. If the mortar is even brighter, those regions have strong on-center signals, and no illusion appears.

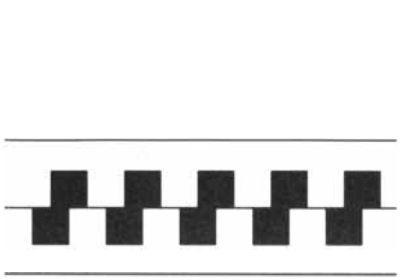
If the mortar is as dark as the dark tiles, the illusion of convergence is weaker and perhaps even absent, because there are no on-center signals from where the mortar separates dark tiles. If the mortar is darker still, off-center signals are generated at the line



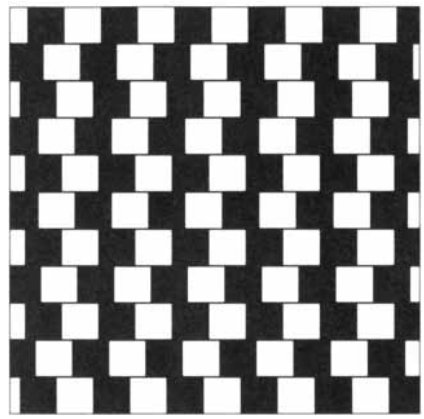
Ganglion-cell signals at an edge

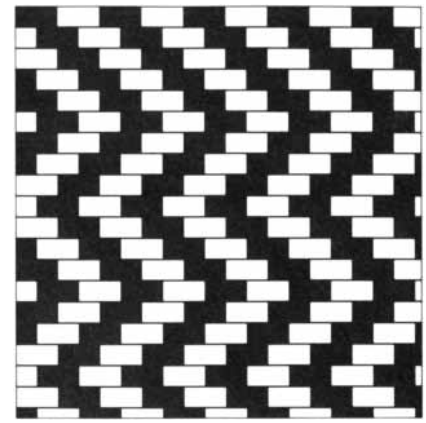


Ganglion-cell signal strengths



Modified Münsterberg illusions





The kindergarten illusion

of separation and the illusion is certain not to appear.

The irradiation seen in the Münsterberg illusion may also incorporate another artifact resulting from edge location. In 1984 Morgan, Moulden, G. Mather of University College London and R. J. Watt of Reading suggested that irradiation is due to a nonlinear response to light in the early part of the visual system. Although the site was not specified, presumably it is at the photoreceptors, the bipolar cells or their myriad interconnections. The term "nonlinear" means that if the brightness of the light is changed by a small increment, the corresponding change in the visual system's response is different depending on whether the light is bright or dim. The nonlinear response alters the ganglion cell's "measurement" of the rate at which the brightness varies across its receptive field, and as a result the edge seems to be shifted slightly into the dark region.

There are several variations of the Münsterberg illusion; two, published in 1978 by R. H. Day of Monash University in Australia, are shown at the left above. The first one is a stripped-down version of the illusion, in which the middle line appears to tilt with respect to the top and bottom lines, which are actually parallel to it. In the second version the offset of the tiles varies from row to row.

Day also published an older array, seen at the right above, that produces what is called the kindergarten illusion. He added a note about an unpublished observation made by Gregory, who pointed out that although the kindergarten array resembles the Münsterberg array, when the two are constructed with equally bright colored tiles the kindergarten array retains some sense of convergence, whereas the Münsterberg array does not. Perhaps some other mechanism plays a role in the powerful kindergarten illusion.

Another illusion of convergence was published in 1980 by Steve P. Taylor and J. Margaret Woodhouse of Cardiff, Wales [*see illustration below*]. When the same line serves as the border of two adjacent square outlines, no illusion appears (*left*), but when adjacent borders are formed by two distinct lines, there is a sense of convergence (*middle*). A stripped-down version is also shown (*right*); the array resembles a twisted-cord array even though there are no explicitly tilted elements. I find that these bare-bones designs exhibit tilt (with respect to long, free-standing parallel lines) even if they are arranged in a column or placed on a page at random. In 1985 Paola Bressan of the University of Padua demonstrated that the illusion of convergence disappears if the lines in the stripped-down version are made too thin or too thick.

I think the illusion must have to do with how the lines are mapped in the visual system's raw primal sketch. If they are too thin, they are represented by horizontal lines. If they are too thick, their details are more faithfully mapped with edges that retain their horizontal structure. But if they are of intermediate size, they are mapped with tilted bars that approximate their shape without the details. Later the details are added, so that you finally perceive the stacked lines, but the impression of tilt lingers.

You might search for the various illusions on tiled walls or in graphic designs and modern art. I wonder, incidentally, how many times a graphic design has had to be drawn wrong—or a tiled wall redesigned—in order to look right!

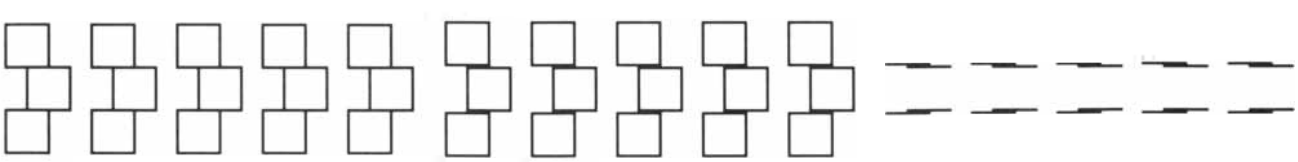
FURTHER READING

BORDER LOCKING AND THE CAFÉ WALL ILLUSION. Richard L. Gregory and Priscilla Heard in *Perception,* Vol. 8, pages 365–380; 1979.

VISION: A COMPUTATIONAL INVESTIGATION INTO THE HUMAN REPRESENTATION AND PROCESSING OF VISUAL INFORMATION. David Marr. W. H. Freeman and Company, 1982.

THE MÜNSTERBERG FIGURE AND TWISTED CORDS. M. J. Morgan and B. Moulden in *Vision Research,* Vol. 26, No. 11, pages 1793–1800; 1986.

FURTHER STUDIES OF THE CAFÉ WALL AND HOLLOW SQUARES ILLUSIONS. J. Margaret Woodhouse and Steve Taylor in *Perception,* Vol. 16, pages 467–471; 1987.



The Taylor-Woodhouse illusion

COMPUTER RECREATIONS

On making and breaking codes: Part II

by A. K. Dewdney

"What was enciphered on a machine might all the easier be deciphered on a machine."

—Andrew Hodges, *Alan Turing: The Enigma*

Until World War II concealment of messages by means of codes and ciphers had been a tedious manual operation. It was during the war that electromechanical machines such as the German Enigma took over. Now computers do the encoding. They swiftly scramble a message in ordinary language (called the plaintext) according to a secret "key" and transmit the scrambled message (called the ciphertext) over telephone or other electronic-communication links. Computers that are supplied with the same key unscramble the ciphertext just as quickly. This, in any event, is the idea behind the Data Encryption Standard (DES), a method of encrypting information issued by the U.S. National Bureau of Standards in 1977 and currently used in commercial and perhaps military communication systems.

In what follows I shall be presenting a complete description of the DES that happens to be its most public appearance to date. I can do so without reservation, because the security of any ciphertext generated by the DES depends entirely on keeping its associated key secret. Yet recent developments in cryptology, the science of making and breaking codes and ciphers, have centered on keys that are partly public and partly private: a private key is buried in a well-known problem the public cannot solve.

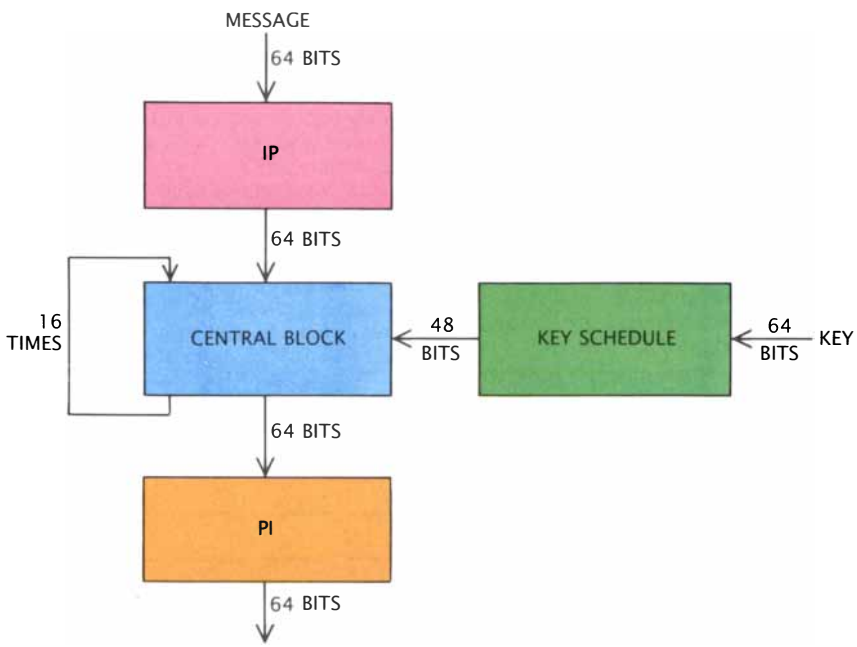
Regardless of what is allowed to be common knowledge and what is kept secret, any computer-based cryptosystem is built around an encoding and a decoding algorithm. The encoding algorithm of the DES was originally designed at the IBM Research and Development Corporation. Like the Enigma machine described in last month's column, the DES employs a sequence of scrambling operations that individually are rather simple but collectively are complicated. In the Enigma machine the component operations were effected by mechanical wheels; in the DES they are executed by either software "modules" or microchips.

Because computers are involved in the encryption process, the symbols that are scrambled are not alphabetic letters (as in the Enigma code) but bits, or binary digits. The DES handles only one string of 64 bits at a time. Hence in order to apply the algorithm one must first translate the plaintext message into 64-bit strings. Almost any convention for turning a message in ordinary language into a string of bits is appropriate. For example, since five bits suffice to count to 31 in base 2, the base-2 representations of 0 through 25 can be chosen to stand for 26 different alphabetic characters. Such a scheme would allow 12 contiguous letters of English to be represented as a 64-bit string (with four bits left over).

A plaintext message being encoded by the DES can be visualized as a river that divides and redivides in an extremely complicated manner. Indeed, the best way to keep track of the individual operations is to follow them on a flow chart [*see illustration on this page*]. Individual modules are like waterfalls, churning the code waters to an almost unbelievable degree. Each 64-bit string of plaintext encounters three types of modules on its rocky ride to the output channel: permutation, left-shifting and substitution modules. The operations done by each type of module can be summarized neatly in tables.

A permutation table indicates how the order of a string's bits is to be altered by a given permutation module. A simple example shows how to read such tables. Suppose a permutation module operates on the binary string 1011 according to the following (horizontal) table: 3 1 2 4. This means the string's third bit goes to the first position, its first bit goes to the second position, its second bit goes to the third position and its last bit stays in the last position. Hence by applying the table the number 1011 is permuted into 1101.

The DES employs several different permutation tables: the so-called ini-



The Data Encryption Standard (DES) in outline

tial permutation table (IP), the E bit-selection table (E), a pair of permuted-choice tables (PC1 and PC2), the permutation table (P) and finally the inverse permutation table (PI). These tables are listed in the illustration on this page.

Some of the tables do not contain 64 entries, because their input or output strings have fewer than 64 bits. The E bit-selection module, for example, accepts a 32-bit string and expands the string to 48 bits. An inspection of the corresponding table reveals, not surprisingly, that some entries appear more than once. The permuted-choice modules, PC1 and PC2, have the opposite effect: they shrink the strings that pass through them. Accordingly some entries are missing in the corresponding tables.

Home encrypters can transfer the tables, as well as the DES itself, to their computers, but they will have to do so without much advice from me. Since space is short in this month's column, I have only enough room to mention that assignment statements to be used with permutation tables will have the form

$$T(k) \leftarrow M(P(k)),$$

where the array M holds the bits of the string to be permuted, the array P contains the table, T is a temporary array and k counts off the new bit positions. In this way the computer can be programmed to place the $P(k)$th entry of M in the kth position of T. Once all entries in T have been made, M can be set equal to T.

The DES locks messages into ciphertext with a secret key: an arbitrary 64-bit string the DES user supplies to the so-called key-schedule module at the beginning of the encoding process. The key can be thought of as a 64-bit tributary river that narrows to 56 bits in the PC1 module, divides into two 28-bit streams that then come together at the PC2 module and join the main message river sequentially as 16 48-bit rivulets—one for each major iteration of the DES.

While they are in the key schedule, the two 28-bit streams also encounter crosscurrents that shift their bits to the left. As can be seen in the table at the top of the illustration on the next page, in the first, second, ninth and 16th iteration the DES calls for shifting the bits of the two strings to the left by one bit. Otherwise the strings are shifted to the left by two bits.

Such left shifts are easy to program. The 28 individual bits of a key segment are assigned bit by bit to a temporary array in such a way that the ith element of the temporary array takes on the value of either the $(i+1)$th or the $(i+2)$th bit of the segment. Of course, the last element (or two) of the temporary array must be set equal to the first bit (or two) of the key.

The DES also incorporates eight substitution modules each of which converts a six-bit string into a four-bit one. The bits of the number undergoing conversion tell one which row and column indexes to look up in the corresponding tables: the first and last bits determine the row index, and the middle four bits determine the column index [*see illustration on next page*]. For example, the six-bit number 111010 yields the row index 10 (the outside bits) and the column index 1101 (the inside bits). These binary numbers represent respectively the decimal numbers 2 and 13. The entry in table S1 for row 2 and column 13 is the decimal number 10. In binary form 10 happens to be 1010, which would be the output of the S1 module if it had the string 111010 as input. (Most home cryptologists will probably need two conversion procedures in order to change a binary number to a decimal number and to do the reverse.)

So much for the bits and pieces. How do they all fit together? A 64-bit string enters the IP module and is permuted as the module's corresponding table describes. It then passes on to the central block, where the main encoding algorithm scrambles it hopelessly in a kind of recirculating eddy, before it passes on to the PI module, where it is permuted according to a table that happens to be the inverse of the IP table.

The central block takes as input not only the output of the IP module but also the 48-bit keys (numbered from 1 through 16) that are derived in the key-schedule module from the supplied 64-bit key. The same basic operations are carried out in the central block 16 times, but each time with a different 48-bit key. Specifically, the 64-bit string from the IP module is split into right and left 32-bit halves. A copy of the right half is then blended with key 1 in the so-called F module. The result is added modulo 2 to the left half. (Modulo-2 addition is like ordinary addition except that 1 plus 1 equals 0.) The 32-bit sum and the original 32-bit right half then change places for the next iteration (with key 2 this time), the sum becoming the right half of a 64-bit string and the right half becoming the left half. At the completion of 16 such blending-and-swapping iterations the two 32-bit strings are put together once again into a 64-bit whole.

The F operation deserves a diagram of its own [*see bottom illustration on page 145*]. A copy of the right half of the current 64-bit string enters the E bit-selection module, where it is expanded into 48 bits. It is then added modulo 2 to the 48-bit key appropri-

IP	E	PC1	PC2	P	PI	
58	32	57	14	16	40	1
50	1	49	17	7	8	2
42	2	41	11	20	48	3
34	3	33	24	21	16	4
26	4	25	1	29	56	5
18	5	17	5	12	24	6
10	4	9	3	28	64	7
2	5	1	28	17	32	8
60	6	58	15	1	39	9
52	7	50	6	15	7	10
44	8	42	21	23	47	11
36	9	34	10	26	15	12
28	8	26	23	5	55	13
20	9	18	19	18	23	14
12	10	10	12	31	63	15
4	11	2	4	10	31	16
62	12	59	26	2	38	17
54	13	51	8	8	6	18
46	12	43	16	24	46	19
38	13	35	7	14	14	20
30	14	27	27	32	54	21
22	15	19	20	27	22	22
14	16	11	13	3	62	23
6	17	3	2	9	30	24
64	16	60	41	19	37	25
56	17	52	52	13	5	26
48	18	44	31	30	45	27
40	19	36	37	6	13	28
32	20	63	47	22	53	29
24	21	55	55	11	21	30
16	20	47	30	4	61	31
8	21	39	40	25	29	32
57	22	31	51		36	33
49	23	23	45		4	34
41	24	15	33		44	35
33	25	7	48		12	36
25	24	62	44		52	37
17	25	54	49		20	38
9	26	46	39		60	39
1	27	38	56		28	40
59	28	30	34		35	41
51	29	22	53		3	42
43	28	14	46		43	43
35	29	6	42		11	44
27	30	61	50		51	45
19	31	53	36		19	46
11	32	45	29		59	47
3	1	37	32		27	48
61		29			34	49
53		21			2	50
45		13			42	51
37		5			10	52
29		28			50	53
21		20			18	54
13		12			58	55
5		4			26	56
63					33	57
55					1	58
47					41	59
39					9	60
31					49	61
23					17	62
15					57	63
7					25	64

Permutation tables employed by the DES

ate for the current iteration number. Thereafter the 48 bits are separated like so many rivulets into eight equal parts of six bits each. The parts flow through the substitution modules S1 through S8, where each is reduced to four-bit numbers. Hence when the eight rivulets rejoin as one stream, the overall size of the string has shrunk to 32 bits. At this point the P module scrambles the string once again.

If the reader has been brave enough to keep up with the discussion, he or she might be wondering how one could possibly recover the plaintext message from the ciphertext after all the turbulent mixing the plaintext goes through in the DES. Is it as difficult as making water run uphill? Actually that is the trick: run the main operations of the central block backward. After directing a 64-bit string of ciphertext through the IP module (as one does when encrypting), have the string retrace the 16 iterations in reverse order—from the bottom of the central block's blending-and-swapping cascade to the top. The bits of the key must also be fed into the key schedule in reverse order. After another permutation by the PI module the original message should gush like a geyser.

Is the DES really secure? Some critics say it is not. Even before the standard was issued, Whitfield Diffie and Martin E. Hellman of Stanford University warned that the DES was vulnerable to a brute-force cryptological attack by a well-funded organization, one able to build parallel-processing computers that can test 10 billion keys per second.

The need to distribute secret keys among senders and receivers of messages is another weak spot in the DES; it makes it easier for an unfriendly agent to somehow get a copy of the key and use it to decode messages. For this reason Diffie and Hellman proposed 12 years ago that "public" keys be distributed among the members of a communication network.

ITERATION NUMBER	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
NUMBER OF LEFT SHIFTS	1	1	2	2	2	2	2	2	1	2	2	2	2	2	2	1

		COLUMN														
	ROW	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
S1	0	14	4	13	1	2	15	11	8	3	10	6	12	5	9	0	7
	1	0	15	7	4	14	2	13	1	10	6	12	11	9	5	3	8
	2	4	1	14	8	13	6	2	11	15	12	9	7	3	10	5	0
	3	15	12	8	2	4	9	1	7	5	11	3	14	10	0	6	13
S2	0	15	1	8	14	6	11	3	4	9	7	2	13	12	0	5	10
	1	3	13	4	7	15	2	8	14	12	0	1	10	6	9	11	5
	2	0	14	7	11	10	4	13	1	5	8	12	6	9	3	2	15
	3	13	8	10	1	3	15	4	2	11	6	7	12	0	5	14	9
S3	0	10	0	9	14	6	3	15	5	1	13	12	7	11	4	2	8
	1	13	7	0	9	3	4	6	10	2	8	5	14	12	11	15	1
	2	13	6	4	9	8	15	3	0	11	1	2	12	5	10	14	7
	3	1	10	13	0	6	9	8	7	4	15	14	3	11	5	2	12
S4	0	7	13	14	3	0	6	9	10	1	2	8	5	11	12	4	15
	1	13	8	11	5	6	15	0	3	4	7	2	12	1	10	14	9
	2	10	6	9	0	12	11	7	13	15	1	3	14	5	2	8	4
	3	3	15	0	6	10	1	13	8	9	4	5	11	12	7	2	14
S5	0	2	12	4	1	7	10	11	6	8	5	3	15	13	0	14	9
	1	14	11	2	12	4	7	13	1	5	0	15	10	3	9	8	6
	2	4	2	1	11	10	13	7	8	15	9	12	5	6	3	0	14
	3	11	8	12	7	1	14	2	13	6	15	0	9	10	4	5	3
S6	0	12	1	10	15	9	2	6	8	0	13	3	4	14	7	5	11
	1	10	15	4	2	7	12	9	5	6	1	13	14	0	11	3	8
	2	9	14	15	5	2	8	12	3	7	0	4	10	1	13	11	6
	3	4	3	2	12	9	5	15	10	11	14	1	7	6	0	8	13
S7	0	4	11	2	14	15	0	8	13	3	12	9	7	5	10	6	1
	1	13	0	11	7	4	9	1	10	14	3	5	12	2	15	8	6
	2	1	4	11	13	12	3	7	14	10	15	6	8	0	5	9	2
	3	6	11	13	8	1	4	10	7	9	5	0	15	14	2	3	12
S8	0	13	2	8	4	6	15	11	1	10	9	3	14	5	0	12	7
	1	1	15	13	8	10	3	7	4	12	5	6	11	0	14	9	2
	2	7	11	4	1	9	12	14	2	0	6	10	13	15	3	5	8
	3	2	1	14	7	4	10	8	13	15	12	9	0	3	5	6	11

The left-shift table **(top)** ***and substitution tables*** **(bottom)**

The public-key distribution scheme depends on a very large prime number, p, and a base number, a. The number a is deliberately chosen so that a^n modulo p, where n equals 0, 1, 2 . . . , gives all the integers between 1 and $p-1$. All members of the communication network know the numbers p and a, and each member must submit a key that will be listed in a network directory. Member i, for example, selects a personal "half key," x_i, which she locks in a safe. She then computes the number k_i given by the equation

$$k_i = a^{x_i} \pmod{p}$$

and sends it in to be published in the directory under her name. The same thing is done by each member.

If member j wishes to communicate secretly with member i, he looks up k_i in the directory, then uses his own private half key, x_j, to compute the combined key, k_{ij}, by raising k_i to the power of x_j modulo p. In other words,

$$k_{ij} = k_i^{x_j} \pmod{p}.$$

Knowing that member j has sent her a message, member i similarly computes the key k_{ji} by raising k_j to the power of x_i modulo p. Readers might enjoy confirming that both members in fact are using the same key, that is, k_{ij} equals k_{ji}.

An inquisitive outsider might break into a public-key system by solving the discrete-logarithm problem with reasonable speed. If the number y in the formula

$$y = a^x \pmod{p}$$

is called a discrete power of a, then x could be called the discrete logarithm of y base a. The discrete-logarithm problem simply asks us to find x, given y, a and p.

Is there no method faster than simply trying all possible values of x, raising a to that power modulo p each time and then comparing the result to y? Theorists think there is not, but no one has been able to prove it. If the discrete-logarithm problem turns out to be as intractable as it seems, then a rather large value of x, say one that is a few hundred digits long, should discourage computational interlopers.

Diffie and Hellman expanded their public-key distribution scheme into a public-key cryptosystem by envisioning a special kind of function called a trapdoor one-way function. Here is how the system works. Member i of a communication network selects an

arbitrary random integer, k_i, as her private key but publishes in a public directory her special encoding algorithm, E_i. She also formulates another algorithm, D_i, for decoding, which she keeps secret as well. Member j does the same thing, publishing his encoding algorithm, E_j, but keeping his special number, k_j, and his decoding algorithm, D_j, secret.

Later, when member j wants to send a confidential message to member i, he simply looks up her encoding algorithm, E_i, in the public directory and applies it to encode a plaintext message m. In short, he transmits the ciphertext y, where

$$y = E_i(m) .$$

It is now a simple matter for member i to decode the message by using her own secret algorithm, D_i, on y.

Such a cryptosystem depends critically on encoding algorithms that compute "one way" functions, so that knowing E_i makes it no easier to find D_i. The system sounded marvelous, even though no functions were known to be one way at the time Diffie and Hellman proposed it.

That was when the RSA cryptosystem came on the scene. RSA stands for Ronald Rivest, Adi Shamir and Leonard Adleman, who in 1978 developed what is still voted as the public-key cryptosystem most likely to succeed. In the RSA cryptosystem each member of a communication network selects two prime numbers and an encoding exponent. Member i, for example, selects the prime numbers p_i and q_i as well as an encoding exponent, e_i. She publishes the product $n_i = p_i \times q_i$ and her encoding algorithm, which converts a numerical message m as follows:

$$E_i(m) = m^{e_i} \pmod{n_i} .$$

Member i also formulates her private decoding algorithm,

$$D_i(x) = x^{d_i} \pmod{n_i} .$$

As it happens, the algorithm D_i exactly undoes the work of the algorithm E_i. The two algorithms are used precisely in the manner envisioned by Diffie and Hellman in their public-key cryptosystem. Why does it work? How does member i find the number d_i?

The magic number d_i is easy to compute when one knows both the numbers e_i and $f_i = (p_i - 1) \times (q_i - 1)$. To find it, one solves the modular equation $e_i \times d_i \pmod{f_i} = 1$. In other words, the product of d_i and e_i must be equal to the sum of a multiple of f_i and 1. Additionally, e_i must be relatively prime to f_i: the two numbers must have no common divisor except 1.

Hence when member i receives a message encoded in her own algorithm E_i, she receives in effect a message m in the form

$$m^{e_i} \pmod{n_i} .$$

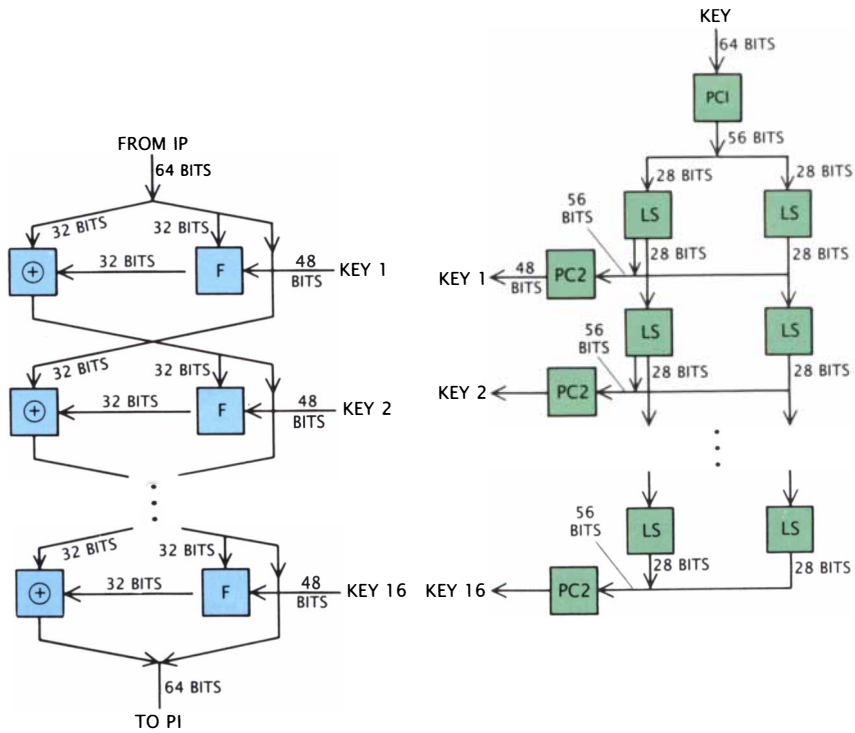
If she now raises this expression to the d_i power, she forms

$$m^{e_i \times d_i} \pmod{n_i} .$$

The new exponent of m, $e_i \times d_i$, as I mentioned above, is equal to the sum of some multiple of f_i and 1. According to a theorem of Euclid, however, a number raised to the sum of f_i and 1 modulo n_i exactly equals m itself: the plaintext stands naked in the mathematical daylight!

The question of whether the RSA cryptosystem can be broken has an answer similar to that given for the public-key system. If someone develops a really fast factoring algorithm, one that can chop very large numbers quickly into their prime components, there would be little trouble breaking the RSA cryptosystem. As soon as the number n_i is published, the computational interloper simply breaks it into the factors p_i and q_i and finds the related number d_i, with which he or she can decode any ciphertext from member i. Yet theorists believe this kind of factoring may be one of the problems destined to be forever intractable.

Owing to lack of space, I shall comment on the July column next month.

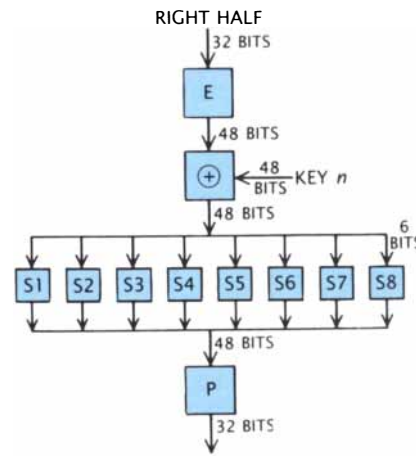


***Expanded versions of the central block* (left) *and the key schedule* (right)**

FURTHER READING

DATA ENCRYPTION STANDARD. National Bureau of Standards, Federal Information Processing Standards Publication 46. National Technical Information Service, 1977.

THE FIRST TEN YEARS OF PUBLIC-KEY CRYPTOGRAPHY. Whitfield Diffie in *Proceedings of the IEEE*, Vol. 76, No. 5, pages 560–577; May, 1988.



Expanded version of the F module

BOOKS

The noble family, the most ancient rocks, views sacred and profane, what scientists do

by Philip Morrison

GRASS SYSTEMATICS AND EVOLUTION, edited by Thomas R. Soderstrom, Khidir W. Hilu, Christopher S. Campbell and Mary E. Barkworth. Smithsonian Institution Press ($45).

From the color frontispiece (a lovely and appropriate Dürer) to the list of participants (about 150 experts from all the continents), this handsome report of a 1986 symposium in Washington, D.C., is a model—which is not to say that all its 33 papers lie within the reach of a nonspecialist reader. They do not, and yet so much here is worth reporting that the heavy volume could not go unnoted.

The old symbiosis between humankind and the grasses has certainly fed the botanists' interest in "that noble family of plants," whose ubiquity and diversity are matched only by the orchids, another worldwide family that counts its species in the tens of thousands. Orchids show off their flowers, artful traps baited with sight and scent for insect pollinators, but their seeds are tiny, wind-sown, unseen. Grasses are strategists of the opposite school: their pollen is flung to the winds, if by no means randomly, outfitted with subtle adaptations for making best use of flowing air. The fruits of grasses are dispersed by diverse means. Some burs stick to fur, some sport long, silky hairs to catch the wind, some plants dry up as a whole and then come loose to tumble across the prairie, some seeds have oil droplets for the tending ants, and some—which we live by—bear fruit that is very good to eat: wheat, rice, maize, barley.

This symposium examined the systematics, genetics and evolution of the grasses. The herbarium sheets, stiffened with dry stalks that faintly shed the fragrance of green meadows after a century on the shelf, remain canonical, but what is striking is the variety of techniques now being vigorously applied, even in the venerable domain of systematic botany.

Nothing seems more modular than the blade of grass, the living letter from which a wide, verdant prairie is spelled out. That theme of modularity informs many of these up-to-date papers. The grass plant itself can be understood as a union of stalk joints, each an element in terms of which the entire branching plant can be analyzed. The distinct characters of some 750 genera of grasses are a combinatorial burden even for experts. An Australian organization has placed these coded descriptions on line and has devised a computer program that will assemble a new diagnostic key to order for any classifier who knows what kind of traits appear to stand out in the specimens that lie before him.

Chromosomes are modules too, and their patterns offer much insight, particularly in this family where multiplication of chromosomes is common, so that some 12-ploid species rejoice in close to 300 of them. The photosynthesizing chloroplasts, without which grass would not be green, hold chromosomes all their own, small neat rings of double-strand DNA open to mapping and comparison across the family. About 50 distinct enzymes found in the grasses have been assayed for biochemical diversity. These proteins vary under genetic control but broadly retain their catalytic function. The rather inbred commercial hybrid maizes of the Corn Belt show a degree of enzyme variability about half that of the most complex Mexican races.

Fossil grasses have been known since 1820. On such strong evidence alone we can put the origin of the grasses back 35 million years; they might have been present as rarities much earlier. The horse in America coevolved with the grassy plains, as did other hoofed grazers; there is even a fossil rhinoceros from the Middle West with fossil grass fruits under the tongue bones. Arguing from fossil grazers and from modern diversity, one evolutionist here dates the rise of the grasses to abundance to as much as 50 million years ago, in tropical South American lowlands where there was open, unforested habitat.

One paper was riveting to a general reader. Its author, botanist Hugh Iltis of the University of Wisconsin at Madison, has sought for 20 years (as apparently his father did before him) the origin of that remarkable grass, maize, which has long been the staff of life for the Americas and is unable to reseed itself without human aid. About five years ago he put forward a novel theory, now compellingly reviewed.

All agree that maize has one forebear: the complex branched annual grass teosinte, whose lush growth to a 12-foot height is familiar in season over wide areas of central Mexico. That is not far from where the oldest archaeological remains of maize (not teosinte) are found, dated 7,000 years back, before protofarmers had even settled down to farm.

The problem is easily put. Teosinte bears up to hundreds of small ears, but when the grains ripen, their stone-hard little fruit cases do not open. Although the ripe ears are indestructible, none has ever been found associated with evidence of human use. No humans eat the ripe grain now, or ever ate it. (The green stems are sugary and attractive, sold today all over Mexico and found as chewed quids in the old caves.) How did teosinte become maize? The genetic difference is small; the plants are entirely interfertile.

A well-argued theory—once shared by Iltis—requires an entire series of mutational steps to modify the stable two-rowed, few-grained, hard-shelled ears of teosinte into the soft-husked ears of maize—now many-rowed and giant. That proposed sequence is elaborate and chancy. The ears of maize are themselves terminal branches like the male tassels of teosinte, not close to the main stems like the lateral ears of the old plant. Iltis' new scheme is startling: the feminine ears of maize are simply the sexually transformed male tassels of its ancestor! The male spikes are far more variable; the sexual development is under hormonal control. The lower structures close to the main branches feminize, the high distant tips become male. Even now feminized tip tassels do appear in maize under fungus infection, and perhaps even in response to the stress of short days and cold nights. Maybe no mutations at all were needed.

It may have taken thousands of years for the new maize ears, now of some use, to respond to deliberate

selection under the "watchful but not overly concerned eyes of incipient Mexican proto-farmers who had other foods to depend on." Slowly the complex husk properties that make for a high yield emerged, until the bountiful crop had become a great staple. That long delay in the history of civilization was still felt when Cortés came. Iltis draws fascinating contrasts with the unbranched simple grasses of the Middle East, nourishing if not easy to harvest even in the wild, and their swift response to easier reaping under unwitting seed selection by a hand-held sickle. In the Old World, sowing and harvesting preceded the few mutations that encouraged full domestication of the grass; in Mexico domestication followed on a sudden big change in the plant.

Why have we never found the "semi-fragile ears" that would show the slow human-selected ear transition from teosinte to maize? They will not be in early agricultural settlements, Iltis argues, but in prefarming country, where the people still gathered seeds knowingly, although not of teosinte. A hundred miles west of the town of Cuernavaca lie winter-dry savanna-covered limestone hills full of caves and sunny slopes, where the form of teosinte that is genetically most like maize grows luxuriantly every summer; "there in a limestone shelter lies a precious secret." The archaeologists have not yet looked in the right place.

It is the conjectural sudden beginning that dazzles most. Perhaps, long ago, the Mexicans endured an ill-timed darkening, an ashfall that led to a year without a summer. Yet everywhere around them, at the end of that hungry season, many of the useless tassels of common teosinte were spontaneously sprouting visible and edible grains. "Who but an unreconstructed pre-Columbian atheist would then not have been tempted to call them 'God's grain,' the literal translation of the Nahuat 'Teo-cintle'!"

OASIS IN SPACE: EARTH HISTORY FROM THE BEGINNING, by Preston Cloud. W. W. Norton & Company, Inc. ($29.95).

So far the beginning is witnessed by astronomy alone: meteorites and moon rock are all we have for samples to establish the 4.6-billion-year age now accepted for the beginnings of sun, earth and moon. Barring a few tiny mineral grains that support an earlier date, the oldest earth rocks we know are in a landscape pictured here at Isua, a stony lakeshore amidst low, dark hills on the southwestern edge of the Greenland ice cap. These sediments, much heated and folded since their origin, were first reliably dated in the mid-1970's: the rocks are 3.8 billion years old. They were laid down on the surface, but their constituent minerals had earlier long been held hot and at high pressure far underground. Surface water was already abundant enough to wear away and transport much rock debris; the weight of atmosphere was sufficient to impress liquidity on those waters, although free oxygen must have been low, because the iron-rich rocks are dark, never rust red.

Here and there around the world comparable "grandfather rocks" have been found. A world map plots the half-dozen locations where such antiquities now lie. For the present these are our protocontinents. Name their time the Isuan. Around them grew the continental nuclei, larger areas on which Isuan locales now form mere patches. That growth took three or four geocenturies (the author's graceful term for 100 million years). Preston Cloud shows us detailed maps to support a close account of each of the four best-known examples of what came next, the "granite-greenstone terrains" he names Pilbaran.

The orbital photograph of Pilbara itself, in the northwestern Australian desert, is striking: oval masses, not very unlike the big domes of Yosemite, lie there, granite islands in a sea of volcanic greenstone. Which came first: the granite rising slowly through earlier volcanic outflow or the volcanics engulfing old granite domes? We had better allow for both cases. These volcanics are mainly marine; they are rich

***Albrecht Dürer's "Das grosse Rasenstück," from* Grass Systematics and Evolution**

in pillow lavas like those we see today whenever lava enters the sea. Nowhere yet had mountains or plains stretched to continental size. Perhaps a fifth of the continental area was in place, in half a dozen fragments. Copious heat flowed out of the young earth, and with it gold and platinum and other rarities from the metallic depths.

These very early lands were mere minicontinents; Neptune reigned then, albeit over less impressive seas. The main formation of continental shields, up to 60 percent or more of all the area that now stands high above the ocean depths, proceeded. The action was extended in both time and space. The southern African example was the first to be laid down, completed three billion years ago; the wider Canadian shield came later, say 2.6 or 2.7 billion years ago. Huge low areas were eventually inundated by deep intrusions of granite, but now with a different chemistry and in consequence a reduced heat flow. The grandest of present geologic processes could begin—the sarabande of the continents.

Worldwide differentiation of the earth's crust was at hand. Two types of crust appeared: the thinner, denser, basal oceanic crust and its overrider, the thicker, lighter plates of continental size. Both crustal types rest on a weak, yielding upper mantle wherein patterns of convection feed worldwide rifts, "like a wet softball bursting its seams," to nudge the plates around. Thermal activity before this time was strong enough but complex; its energies were spent on a regional scale. Slow, deep currents organized worldwide may date from this first appearance of ocean floor and highland plate, preconditions for plate tectonics. (So far hard evidence is not at hand; this chapter of a consistent narrative is based on plausible connections.)

The first great revolution in earth history brought that entire first era of shield forming, called the Archean, to a close. Its end was hardly abrupt; it was "more comparable to the agricultural or industrial revolutions than to the American or French." The long era that came next, 25 geocenturies ago, bears the name Proterozoic. It did not bring first life, whatever the Greek name says; even in Isua life is already hinted by the carbon-isotope ratio. By Pilbaran times, as indeed ever since, we have secure fossil evidence to confront the eye—better, the microscope—of any skeptic. All subsequent eras are paleontological rather than geological: they are marked and described by the development of life forms. To be sure, we know in our bones how intimate are life and the mineral kingdom, to say nothing of the green-plant source of the oxygen we breathe. During the entire long run of the Proterozoic the fossil signs of life are largely chemical—not skeletal or prints or tracks.

The extensive narrow-banded iron formations, undisturbed in their alternations over wide stretches of space and time, the signs (such as vast areas of rust red beds) in many sediments of an atmosphere steadily growing more oxygenated and the complicated layered microbial structures known as stromatolites—all testify to the powers of abundant early life, however silent, still and lowly. Small single cells, often colonially aggregated into knots and mats, grew at all the water margins, although nowhere on the broad uplands.

Step by step the organisms became biochemically more sophisticated; certain inventive microbial pedigrees fused and the new symbiotic cells grew bigger. In time there appeared the first intricate prints of mobile, soft-bodied multicellular marine creatures large enough to be noticed individually. Such longer-lived creatures could endure in the shallows, because increasing ozone filtered out more and more of the solar ultraviolet. The modern era was at hand, geologically like today in event and process except for the effects of evolving life. This era of ours began about seven geocenturies back, when even the trilobites were geocenturies in the future. The entire period is named the Phanerozoic (the time of visible life) and is much subdivided to clarify its oft-told tale, as plates drift and close and mountains and oceans come and go.

Long with the U.S. Geological Survey, Dr. Cloud has been a major contributor to this unraveling. He has made a well-illustrated, closely argued and beautifully written book, the first introductory geology to devote central attention to these decisive ancient events: the first growth of continents and ocean floor, the coevolution of life and air, the surfacing of the ores, the origin of the continental dance, even the perplexing glaciations of the deepest past. Such issues outweigh the pettier transient features of the past few geocenturies, such as life on the land, the Atlantic Ocean, the Himalayas and the fateful parting between dinosaur and mammal.

The thick volume does attend to earth history as a whole. It begins with a somewhat Cloudy summary of the still disputable initial astronomical context and an account of geological techniques. It closes with a fresh and compact account of the later steps in the evolution of life and the lands. Yet it is the marvelous finds and many-sided arguments that underlie the current grasp of earth history before hardshell life arrived that sharply distinguish this excellent and readable if not easy book. Copyediting and photograph reproduction are a little disappointing, but the quality of the text and the wealth of apt illustrations compensate for small annoyances.

MARKINGS: AERIAL VIEWS OF SACRED LANDSCAPES, photographs by Marilyn Bridges, essays by Maria Reiche, Charles Gallenkamp, Lucy Lippard and Keith Critchlow. Aperture Foundation Inc., 20 East 23 Street, New York, N.Y. 10010 ($29.95). **AT WORK IN THE FIELDS OF THE BOMB,** photographs and text by Robert Del Tredici. Harper & Row, Publishers ($35; paperbound, $15.95).

In a light plane a few hundred yards above the downs of Oxfordshire, pilot-photographer Bridges circled the White Horse of Uffington as the low sun brought a textured glow to the broad fields but left deep blacks in every fold. Only a dozen fluid strokes compose the entire 365-foot figure of the animal extended in full gallop. Each stroke is in fact a controlled patch scoured through the rich grass to expose the white chalk below, and by tradition renewed "every seven years during Whitsuntide," beginning perhaps 20 centuries back. The rich black-and-white composition is representation enough to evoke the openness of the sunny place, yet abstraction enough to carry a viewer some way into the sense of ancient purpose that is made manifest here. The artist has brought us many such photographs, archaeological wonders taken "from a witch's distance," although she flies no fireside broom but a loud-throbbing Cessna. "I need to fly," she writes, and as swift as flight and shadow she has brought us back the sense of what she felt. Sometimes it was a lonely fear that the ancient funereal marks sent up; sometimes the artist felt instead a warm community with those artists long past.

Marilyn Bridges has flown her cameras low over the stones of forested Yucatán, over sites in our own country from luxuriant Ohio to the bare Mojave, across the high Peruvian desert at Nazca, over Stonehenge and Carnac and Silbury Hill; from all four provinces of the old people she has gathered 80-odd images to offer a display that both reveals and conceals. The land-

scapes include many sights far less familiar than those named. The ornate ruins of the Mayan world, from Tulúm at the very shore to the great cleared pyramids of inland cities and an unexcavated temple or two still capped by the intruding trees, are here with particular force, at a scale neither intimate nor remote that always rivets the eye. Maria Reiche, who has spent a lifetime puzzling over the folded linearities of Nazca, reminds us in her essay how much we are apt to project our own disparate meanings into such silent forms. Neither witch nor angel has ever seen views like these; the magic we share now with Bridge's vision aloft outdoes the uncertain mysteries of dolmen or kiva.

Robert Del Tredici is another photographer who has taken his camera and his point of view on a long journey. High aloft or underground, his images too are informed by a single theme. These aerial views are not of sacred marks but of the jumbled factories of an ominous industry three decades old. Here one desert patch is strangely meager, for it is only an excluded space between the two fences surrounding a well-guarded bomb-assembly plant. We are moved again by the fear of death, but one that is subtle, implied, potential.

Del Tredici went to Japan, across Europe as far as Lapland and to the outside of Moscow Hospital No. 6, wherein lay the radiation patients from the fiery Chernobyl reactor. He traveled coast to coast across the U.S. and Canada to bring us these 100 black-and-white images from the "fields of the bomb." Some pictures are the work of others, including a fast-camera shot of the Trinity fireball and a page from Hiroshima's agony. But mostly these are evocative images without explicit drama, commonplace industrial scenes of ditches, steel drums and skeins of piping. From the long wall of sandy tailings that lies near the uranium mines at Eliot Lake in Ontario, south to the final assembly plant for all U.S. nuclear weapons, Pantex in Amarillo, we visit the links of the long chain of devoted human effort that has cleverly, laboriously and fatefully armed our country and its counterparts. All the omen is implicit in the pictures (the imagination having been freed by the ordinariness of the scenes), aside from the comments of a few people who ascribe their long, slow killing disease to radiation, and of some witnesses of 1945.

Some of the scenes gleam with the extraordinary: two big Trident submarines side by side on the Thames in Connecticut, one afloat, its sister high and dry on the stocks and another still in disjointed pieces; a long, wide steel-lined vacuum tunnel under a Nevada mesa, through which radiation from a test bomb passes to distant instruments sheltered from the explosion debris by huge, swift-acting gates of foot-thick metal; the mushroom cloud of a recent test in the open desert that set off an unprecedented 5,000 tons of surrogate chemical explosive.

There are some 50 people here in portrait. They level with you, the despairing as well as the complacent. Among them are the general officer who piloted that first bomb to Hiroshima, the bishop of Amarillo and a woman at work in a uranium foundry in Ohio, who drills each newly cast slug of the metal to get shavings for routine sampling. "The woman wears a flower in her hair because it is the week before Christmas."

A final section of the book records the interviews the artist did with the people he met. These are sometimes hot, sometimes cool. A work of art, the book is also a diagnostic cross section of the tissue of life today; we all await judgment by history, the ultimate pathologist. Is this condition within the nations malignant or benign? The final citation is from that most influential of physicists, Dr. Edward Teller of Palo Alto. Once again he patiently explains that even "if you explode all the nuclear weapons in all the arsenals of the world, a large fraction of humanity will not be injured." Reassurance may require more than that.

FLANAGAN'S VERSION: A SPECTATOR'S GUIDE TO SCIENCE ON THE EVE OF THE 21ST CENTURY, by Dennis Flanagan. Alfred A. Knopf ($18.95).

Consider one evocative recollection from Flanagan's guide: "In the late 1940s I visited ... a machine at the Aberdeen Proving Ground in Maryland. The remarkable thing about this early computer is that you could *hear* it working. Over in the corner of a large room filled with racks of relays I would hear the sharp 'bip' of one relay closing. Then ... there would be a fusillade of answering bips. It was about as close as one can get to ... the sensation of being inside a working brain.... In ENIAC the switches were not relays but 18,000 glass vacuum tubes ... one heard the hum of large cooling fans ... as if ENIAC was some kind of heat engine! To be sure, one should not forget that computers, like all other tools, are energy transformers."

So the Version goes, half a dozen witty, deft and anecdotal chapters of explanatory reports on physics, astronomy, geology, biology, technology and—just a quick therapeutic administration—on "the dangerous charm of cranks." Dennis Flanagan has read it all thoughtfully, been everywhere and put good questions to everyone relevant (and remembers the best answers he has heard). He tells, often quizzically, just *wat Wetenschappers doen.* (That phrase has always sounded more impressive to him in Dutch.) The most significant thing scientists do "*is to try to find out things that are not known*" (the italics are his). He is as skilled a professional sharpener of arguments as of blue pencils, cheerfully impatient with cant, vagueness, bathos and pose, a serious seeker for coherence of the basics. From 1948 until his retirement four years ago he was the sophisticated editor of *Scientific American* (including some 250 monthly filings by this reviewer).

Take one biological result that impressed Flanagan. An Oak Ridge ecologist, Jerry Payne, put the carcasses of a few small pigs out in a Tennessee woods. He counted the visiting species (he was one among them) daily for eight days, after which all was dry skin, teeth and bones, plainly a slower phase of degradability. He could identify 522 species of visitors, spread among 359 genera, mostly insects. It is such "sheer pullulating richness" that offers reason to believe "life is inevitable," given the right conditions.

"*At bottom there is nothing in the universe but fields.*" The italics are Flanagan's. It was Columbia physicist Robert Serber who first pointed out to him that each diffracting electron approaching a double slit goes through both holes at the same time, to interfere only with itself. "A light went on in my head," but he kept hoping that someday we would uncover the deeper level underneath the paradoxical properties of quantum fields. More mature now, he recognizes that any preconceived opinion is worth only so much. Maybe such philosophical angst will bring a dire fate to any who ask, "How can it possibly be like that?" Richard Feynman once foretold that they would go "down the drain into a blind alley!"

There is plenty here for thought, the matter not as elementary as one might judge from its brevity and liveliness. A better example of cultivated reading, out of the spirit of the Enlightenment by a master craftsman among the wordwrights of Manhattan, has not been offered to the general scientific reader for a long time.

ESSAY

To sequence or not to sequence

by Robert A. Weinberg

For the past two billion years and more, the organisms inhabiting this planet have been developing extraordinarily complex genetic texts. Our own text—the human genome—now encompasses three billion base pairs of DNA, encoding the information for from 50,000 to 100,000 distinct genes. This genome represents the results of millions of redrafts of a text carried long ago by our single-cell ancestors. Parts of our modern text are recognizably similar to genes carried by yeast and bacteria; other parts, in contrast, seem quite novel.

Given enough hands and money, it is now within reach to determine the entire base sequence of the human genome. Within the sequence lie many clues to our origins and our complex function. The genomic text also contains vestiges of very early drafts, rewrites and rejected paragraphs, as well as lines of the highest poetry of life on this planet.

That sequence can now be deciphered because of two great advances molecular biology has made over the past decade. The first advance involves a cloning technology whereby one can excise discrete DNA segments from the genome and amplify each into an enormous number of copies, which in turn make possible the detailed characterization of each segment. The second advance consists of a group of techniques that enable an experimenter to determine with absolute precision the sequences of the bases composing each segment. The base sequences of the segments, strung end to end, constitute the information content of our genome, in sum a three-billion-character text.

A great debate now surrounds that text and its complete decipherment. There are those who fear the text carries secrets that should remain clouded in obscurity, information not intended for us to see or understand, clues to the essence of our humanity. The debate also involves issues of how biological research is done and perhaps of how it should not be done.

Certainly this public debate involves concerns about money. The torrent of data that will emerge from a human sequencing effort will not come cheaply. Cost estimates ranging up to three billion dollars have been made; some people believe improvements in technology may reduce the figure by as much as a factor of 10. Even so, genome sequencing represents a megaproject that, together with the effort to combat AIDS, may drain substantial funds from the limited pool available for basic biomedical research. Will a headlong rush to sequence the genome solve many of the basic conceptual problems of biology, or will it merely delay their resolution by diverting precious resources?

A genome-sequencing project flies in the face of traditions that are deeply embedded in the culture of modern biological research. Biology has thrived on the fruits of relatively small, highly focused efforts devoted to narrow, well-defined problems. The most successful projects have been executed by specialized, smart and very ambitious individuals coalesced into small, cohesive groups.

The tradition of meticulous inquiry into a narrow, well-defined problem enjoys a simple justification. Biological organisms are so complex that by trying to understand everything, one ends up learning little that is solid. Perhaps it is better to understand the workings of a small number of genes clearly and unambiguously than it is to provide vague, qualified and superficial generalizations about the genome as a whole.

The genome megaproject recalls a moonshot or a supercollider. It conjures up a consortium of dozens of research groups working together in harness, each straining to finish the analysis of its own allotted patch of the genome. Much of the work will be repetitive and routine, to the point where armies of technicians and robotized machines will be required.

Plans call for the project to proceed in two very distinct phases, termed mapping and sequencing. The mappers will determine the order of defined sets of DNA fragments as they are arrayed along the lengths of the human chromosomes. Landmarks the mappers plant will define discrete segments of DNA from hundreds to thousands of bases long. Then, in the project's second phase, the sequencers will move in and determine the order of the bases within these defined segments.

There is little quarrel about the value of mapping; even people averse to megaprojects conclude that it will be useful. The results of mapping would help them to link distinct regions of DNA (and the associated genes) with well-studied biological phenomena. The availability of a complete map of the human genome will enormously facilitate the identification of genes responsible for a thousand diseases such as cystic fibrosis and Huntington's disease as well as for susceptibility to cancer.

But what about the subsequent sequencing? Do we want to know every word and letter in our encyclopedia? Here the debate flares up. It is already apparent that most of the human genome, indeed as much as 95 percent of it, encompasses DNA sequences that have little if any role in templating biological function. Most of the genome seems to represent evolutionary detritus—the discarded drafts of essays that lost any meaning 100 or 1,000 million years ago. The residue, the precious 5 percent that defines us and what we are, encodes the amino acid sequences of the tens of thousands of proteins that create biological structure and function.

Even when the specific sequence of amino acids making up a protein is known, a laboratory can still struggle for a decade or more just trying to learn what that protein does. Even then any resulting understanding of cell and organ physiology is usually limited, since a single protein functions in the context of a complex, interacting network of other proteins. The 50,000 to 100,000 separate amino acid sequences thrust into our lap by this project will only define for us the size of the problem that confronts the next generation of biologists.

In short, the mountain of sequence data will help us a bit but not a whole lot. Ultimately all this information will not really tell us how we function. Insights of that kind will still be gained only by very slow slogging through very rough experimental terrain. The really big problems will remain ahead of us. After all of this, we shall still only have scratched the surface of life.

ROBERT A. WEINBERG is professor of biology at the Massachusetts Institute of Technology and a member of the Whitehead Institute for Biomedical Research. He has written three articles for SCIENTIFIC AMERICAN, including "Finding the Anti-Oncogene," in September.

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.