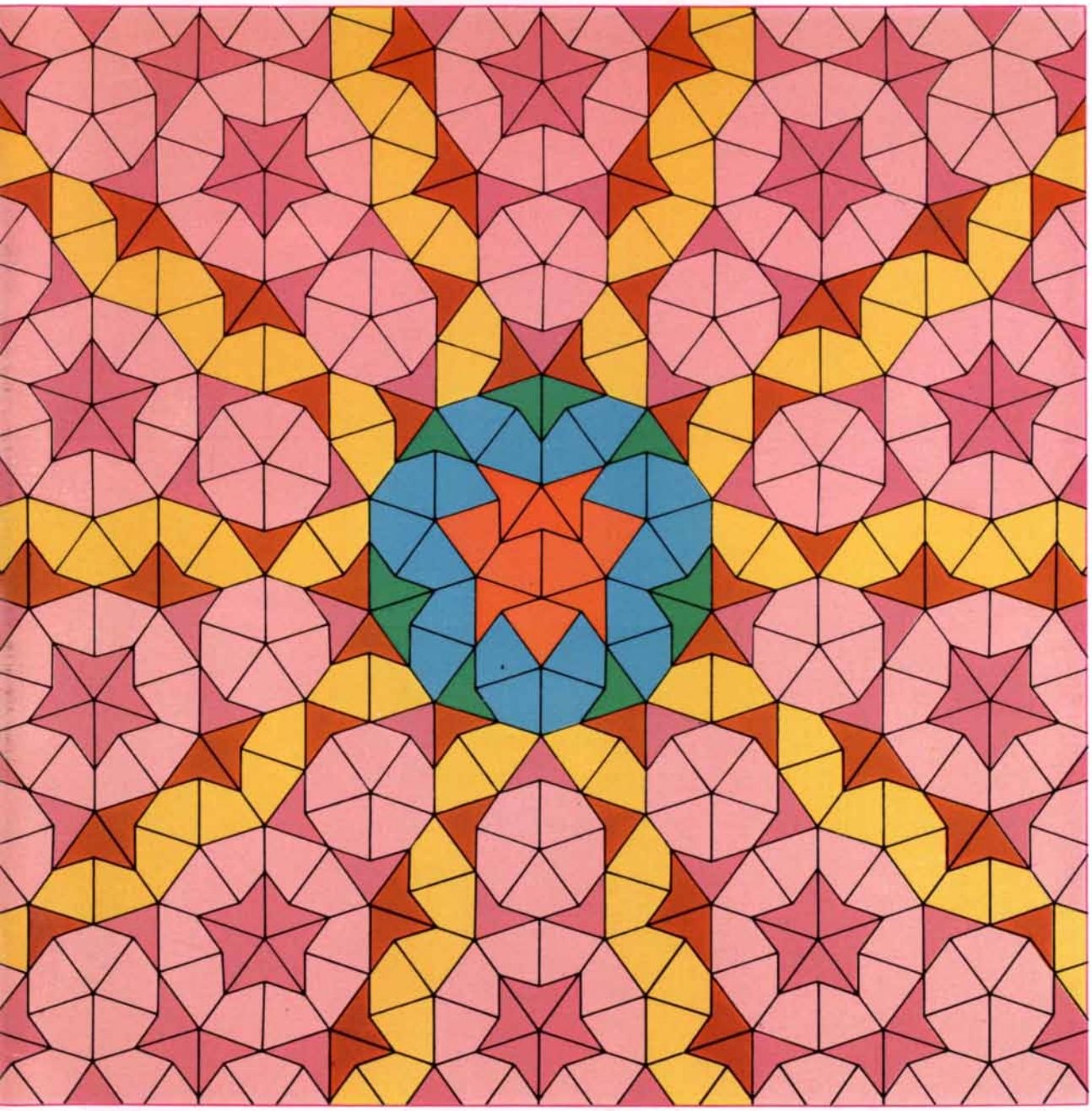
SCIENTIFIC AMERICAN



THEORY OF TILES

$1.50

January 1977

© 1976 SCIENTIFIC AMERICAN, INC

COME IN AND DRIVE

"Legend." That's a pretty strong word, isn't it?

Though we sincerely doubt anyone would consider it too strong when referring to Porsche. After having become possibly the foremost name in racing. After almost thirty years of major automotive innovations not only on the track but on the road. Culminating in the Porsche Turbo Carrera, a car built with such painstaking pride and to such exacting standards that it takes twelve days to build one.

And Audi's credentials are equally impressive.

You probably didn't know it, but Audi has a history that dates all the way

The 1977 AUDI

Our $5445* Audi Fox Sedan with new aerodynamic front-end design and fuel-injected engine which propels the car from 0 to 50 in 8 seconds yet delivers EPA estimated 36 mpg on the highway and 24 in the city for standard shift model (mileage may vary with driver, car's condition and equipment). *Suggested 1977 retail price. P.O.E. Transportation, local taxes, and dealer delivery charges, additional. Whitewalls optional.

© 1976 SCIENTIFIC AMERICAN, INC

ONE OF OUR LEGENDS.

back to 1904 (that's before the Model T). A history that includes a succession of accomplishments on the track and on the road. (Would you believe back in the 30s we had such advanced features as front-wheel drive and dual braking systems?)

But more important than any specific feature of either car is a dedication to engineering excellence and innovation that has been a hallmark of Porsches and Audis through the years. That has separated them from other cars. That has made them more than just...cars.

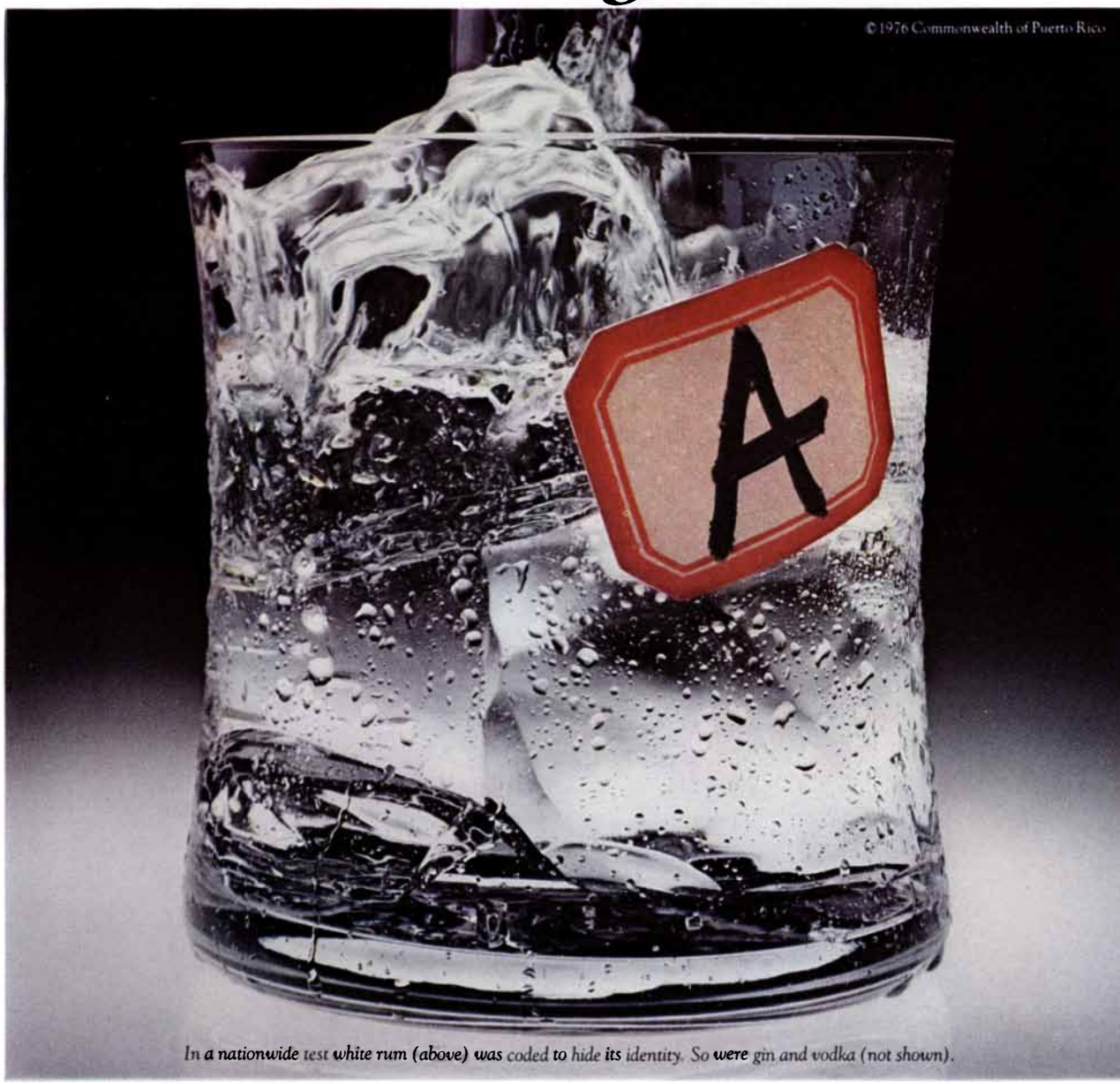
PORSCHE
+AUDI

The 1977 PORSCHE

Our $28,000* Porsche Turbo Carrera. Not only the fastest production car Porsche ever built (0 to 60 in less than 6 seconds) but the most luxurious (supple leather seats, deep cut pile carpeting, automatic temperature control, even 4-speaker stereo). *Suggested 1977 retail price. P.O.E. Transportation, local taxes, and dealer delivery charges, additional.

© 1976 SCIENTIFIC AMERICAN, INC

We spent thousands finding out what we already knew. White rum is smoother than gin or vodka.



In a nationwide test white rum (above) was coded to hide its identity. So were gin and vodka (not shown).

We could have told you white rum was smoother before we spent a mint on a nationwide taste test. But we couldn't advertise it.

Government regulations protect you from idle claims by requiring quantitative substantiation. That's as it should be.

So we went to 21 major cities and asked 550 drinkers to compare white rum with the leading brands of gin and vodka. 24.2% preferred gin. 34.4% preferred vodka. And 41.4% preferred white rum.

It should be noted that the white rum came from Puerto Rico — the only place where the law requires that rum be aged. And since smoothness is a direct result of aging, it's not surprising that more people liked the taste of white rum than gin or vodka.

Enjoy white rum in your next drink calling for gin or vodka. Before long you'll be telling your friends. Fortunately, government regulations don't prohibit you from doing that.

PUERTO RICAN RUMS

For free "White Rum Classics" recipes, write: Puerto Rican Rums, Dept. SC-2, 1290 Avenue of the Americas, N.Y., N.Y. 10019

© 1976 SCIENTIFIC AMERICAN, INC

Established 1845 **SCIENTIFIC AMERICAN** January 1977 Volume 236 Number 1

ARTICLES

DEPARTMENTS

BOARD OF EDITORS Gerard Piel (Publisher), Dennis Flanagan (Editor), Francis Bello (Associate Editor), Philip Morrison (Book Editor), Trudy E. Bell, Brian P. Hayes, Jonathan B. Piel, John Purcell, James T. Rogers, Armand Schwab, Jr., Jonathan B. Tucker, Joseph Wisnovsky

ART DEPARTMENT Samuel L. Howard (Art Director), Ilil Arbel, Edward Bell

PRODUCTION DEPARTMENT Richard Sasso (Production Manager), Carol Hansen, Michelle Lynn and Leo J. Petruzzi (Assistant Production Managers), Susan Caputo, Zelda Gilbert, Annette Rosa, Julio E. Xavier

COPY DEPARTMENT Sally Porter Jenks (Copy Chief), Judith Friedman, Dorothy Patterson, Linda Sternberg

GENERAL MANAGER Donald H. Miller, Jr.

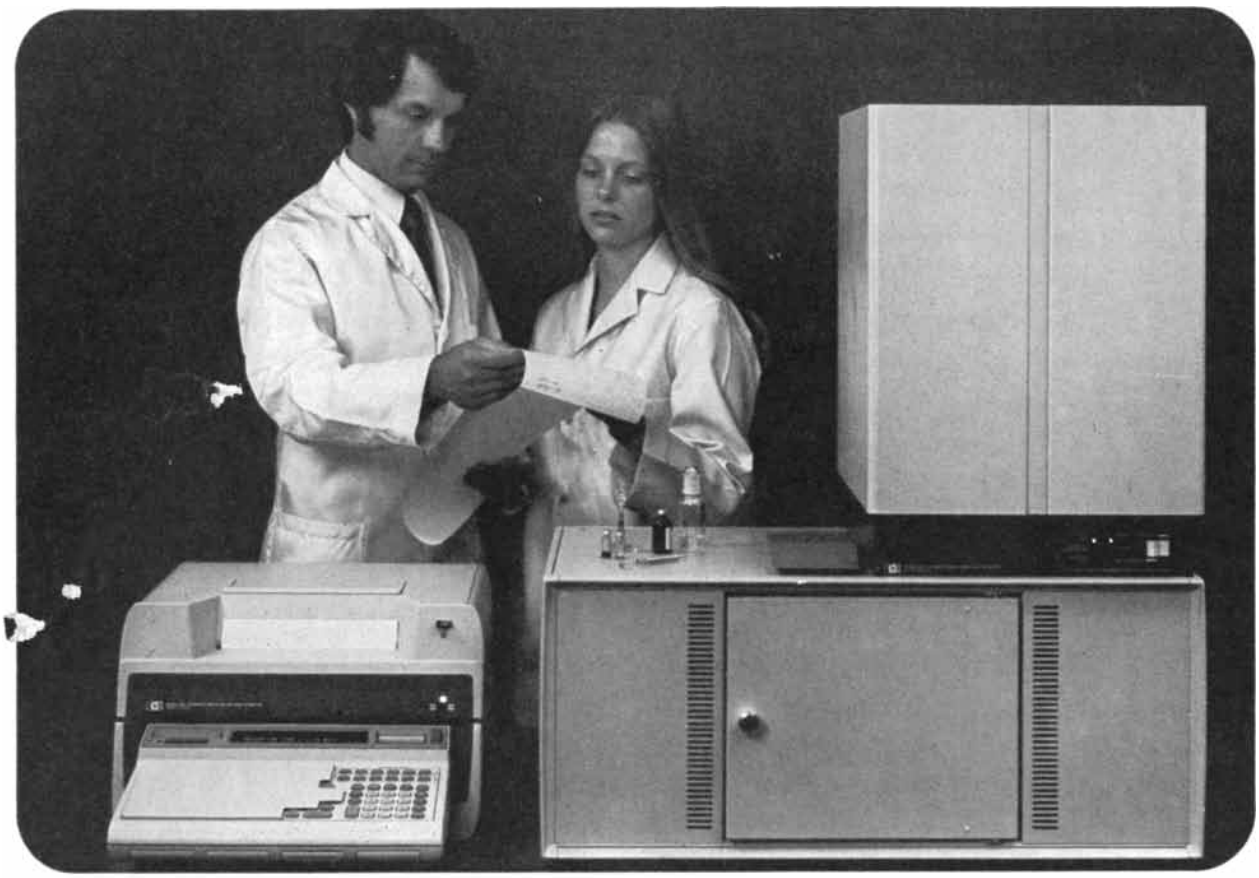
ADVERTISING DIRECTOR C. John Kirby

ASSISTANT TO THE PUBLISHER George S. Conn

CIRCULATION MANAGER William H. Yokel

PUBLISHED MONTHLY BY SCIENTIFIC AMERICAN, INC., 415 MADISON AVENUE, NEW YORK, N.Y. 10017. COPYRIGHT © 1976 BY SCIENTIFIC AMERICAN, INC. ALL RIGHTS RESERVED. PRINTED IN THE U.S.A. NO PART OF THIS ISSUE MAY BE REPRODUCED BY ANY MECHANICAL, PHOTOGRAPHIC OR ELECTRONIC PROCESS, OR IN THE FORM OF A PHONOGRAPHIC RECORDING, NOR MAY IT BE STORED IN A RETRIEVAL SYSTEM, TRANSMITTED OR OTHERWISE COPIED FOR PUBLIC OR PRIVATE USE WITHOUT WRITTEN PERMISSION OF THE PUBLISHER. 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 RATE: $15 PER YEAR, U.S., ITS POSSESSIONS AND CANADA; $18 PER YEAR, ALL OTHER COUNTRIES.

© 1976 SCIENTIFIC AMERICAN, INC



The first high-performance GC/MS system small enough to fit on your lab bench.

Thanks to some striking new technology, HP's new 5992A GC/MS system with computing controller, printer/plotter, and full set of software costs just $47,500*. Analytical laboratories will find that this new-generation GC/MS (gas chromatograph/mass spectrometer) has the powerful identification and quantitation capabilities usually associated with systems costing twice as much.

When a manufacturer introduces a new instrument that is half the cost and size of comparable systems, it is reasonable to wonder if performance has been compromised. On the contrary, the 5992A, with a sensitivity, mass range, and scan speed equal to those of much more expensive systems, produces spectra that are in excellent agreement with, and frequently superior to, accepted reference spectra.

This remarkable development is the result of some innovative technology, most notably in the 5992A's new hyperbolic quadrupole mass filter. While quadrupole mass filters have won wide acceptance as the heart of modern GC/MS systems, nearly all of them use round rods which adversely affect sensitivity, resolution, and line shape. Hyperbolic-shaped rods eliminate these problems, but it has been too difficult to fabricate such rods with the high degree of precision required. With the 5992A, this difficulty has been successfully overcome.

The 5992A is far easier to operate than any previous high-performance GC/MS system. With all operations managed by its fully integrated desk-top computing controller, the 5992A is virtually free of the usual knobs, switches, and buttons in the GC and MS modules. No longer does the analyst have to manipulate numerous interactive adjustments while observing the result on an oscillograph, hoping in this way to reach optimum performance. With the 5992A, the operator simply enters essential data on the keyboard, and the controller does the job automatically.

The new system's automatic tuning capability illustrates the point well. The analyst injects a calibration compound, loads the *Autotune* program, and pushes a button. The controller automatically adjusts all the parameters that control the ion source, mass filter, and detector—always determining the proper values to produce high-sensitivity spectra of consistent quality, day in and day out, no matter who is operating the system.

We'd like to tell you more.

© 1976 SCIENTIFIC AMERICAN, INC

New high-performance alphanumeric display increases realiability through reduced parts, circuitry, and connections.

To get the kind of performance and reliability we demand in HP products that use LED (light-emitting diode) displays, we develop and manufacture them ourselves—along with other performance-critical semiconductor components. This new four-character LED alphanumeric display is a case in point. Usually the larger the range of characters and symbols a display generates, the greater the number of parts, circuits, and connections needed to operate it. The new HDSP-2000 display actually reverses this process, reducing system parts by a factor of 36 to 1 and the number of connections by an order of magnitude or better. It is available to help you simplify your display systems.

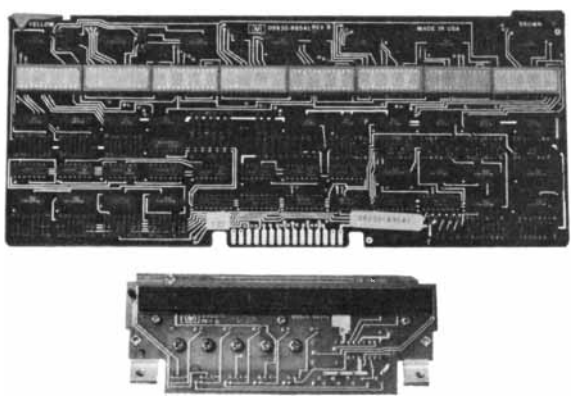
Because a significant portion of the operating electronics are incorporated in the display unit itself, the HDSP-2000 is considerably easier and less expensive to integrate into a system than other current alphanumeric units. And because it requires only 12 pin connections, the display is easier to design into an instrument, easier to troubleshoot, quicker to assemble, and can often contribute to a reduction in the size of the instrument in which it is used. By comparison, existing alphanumeric displays require as many as 35 additional circuit components, and a total of 176 connections to do the same job.

Each 5- by 7-dot character matrix is capable of generating upper and lower case letters, numerals, punctuation marks, and mathematical symbols. The 3.8mm character height permits accurate viewing at up to 3 meters.

The new display is directly compatible with most common integrated circuit logic families, such as TTL, and can stack end to end to any desired line length. It is physically and environmentally rugged—well suited to portable and mobile applications—and requires only 5 volts for operation.

All together, this represents a new performance for alphanumeric displays.

HDSP-2000 displays cost $47* per four-character unit in quantities of 125, and are available from any HP franchised distributor.



Above: a row of eight 4-character displays and the circuitry needed to drive them—state of the art previous to the HDSP-2000. **Below:** a similar row of eight HDSP-2000 displays. Because their shift registers and LED drivers are self-contained, the associated circuitry for the equivalent function is greatly reduced and simplified.

HEWLETT hp PACKARD

1504 Page Mill Road, Palo Alto, California 94304

For assistance call: Washington (301) 948-6370, Chicago (312) 255-9800
Atlanta (404) 434-4000, Los Angeles (213) 877-1282.

Mail to: Hewlett-Packard, 1504 Page Mill Road, Palo Alto, CA 94304.
Please send me further information on

( ) HP 5592 GC/MS system
( ) HDSP-2000 LED alphanumeric display

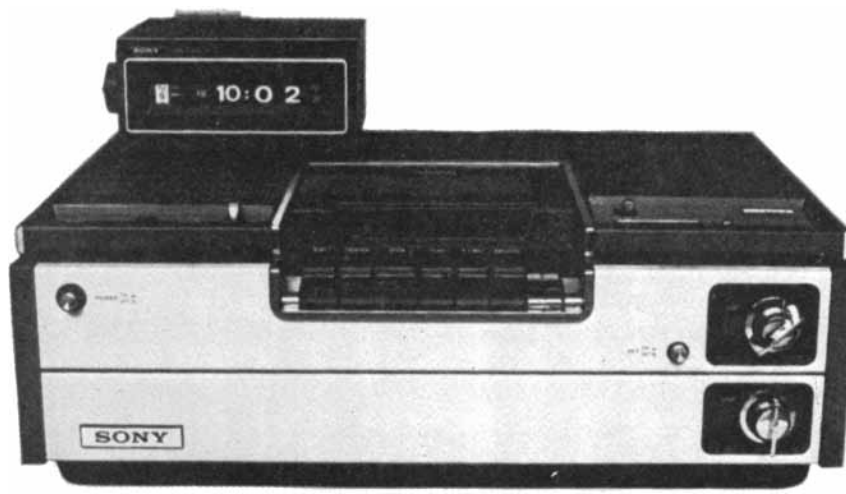
Name _____
Company _____
Address _____
City _____ State _____ Zip _____

*Domestic USA prices only.

00740

© 1976 SCIENTIFIC AMERICAN, INC

SONY PRESENTS THE NEXT THING.



Sony's Betamax Videotape deck: It lets you see shows you would have missed.

You're looking at Betamax, a revolutionary new product from Sony that plugs into any TV and enables you to see programs you would have otherwise missed.

We'll explain.

First off, let's take a situation where there are two shows on opposite each other and you'd like to watch both of them. Well, believe it or not, now you can. Because Sony's Betamax deck can actually videotape something off one channel while you're watching another channel. Then, when you're finished watching one show, all you do is push some buttons and you can play back a tape of the show that you would have missed.

TV Picture Simulated

Pretty incredible, huh? Well, listen to something else Betamax does that's equally incredible.

Let's say you have to go somewhere, or do something, at a time when there's something on TV you want to see. Well, Sony's Betamax is equipped with a timer that can be set to automatically videotape that program while you're not there. Then, whenever you want, you just play back the tape—and again you see what you would have missed.

Our one-hour tapes are reusable—just record over them and use them over and over again.

Imagine. With Sony's Betamax, you'll never again miss a program you want to see.

Ah, progress!

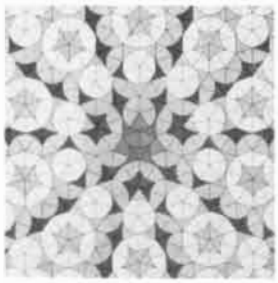
Betamax plugs into any TV, even if it's not a Sony (though you'll be missing a lot if it's not).

BETAMAX®
"IT'S A SONY."®

Model SL-7200 Videotape Player/Recorder

© 1976 Sony Corporation of America
SONY and Betamax are trademarks of Sony Corporation

© 1976 SCIENTIFIC AMERICAN, INC



THE COVER

The design on the cover is a nonperiodic tiling discovered by the British mathematician Roger Penrose (see "Mathematical Games," page 110). A periodic tiling is one in which a transparency of a tile pattern could be laid on top of the pattern and could be moved to an infinity of new positions (without rotation or reflection) and still fit the pattern. Typical periodic tilings are made by the hexagonal tiles of a bathroom floor and the squares of a checkerboard. Freer forms, such as those that appear in some of the repeating designs of the Dutch artist M. C. Escher, can also lend themselves to periodic tilings. Nonperiodic tilings also cover the plane with patterns that may have various kinds of symmetry, but they lack a "fundamental region" that tiles by translation. Some nonperiodic tilings, such as the bilaterally symmetric one on the cover, have a unique center from which features such as the "spokes" radiate to infinity.

THE ILLUSTRATIONS

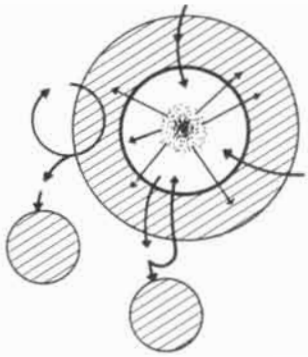
Cover illustration by Andrew Christie

Page	Source
22–27	Jerome Kuhl
29	Allis-Chalmers
30	Alan D. Iselin
31	Ohio Agricultural Research and Development Center
32–33	Alan D. Iselin
35–40	Dan Todd
50	June Almeida, Bernhard Cinader and Allan F. Howatson
51	Louis W. Labaw and David R. Davies
52–58	George V. Kelvin
61	Anne d'Harnoncourt
62–63	Robert Sekuler and Eugene Levinson
64–70	Jerome Kuhl
72	Robert Sekuler and Eugene Levinson (*top*), Jerome Kuhl (*bottom*)
73	Jerome Kuhl
74–75	C. C. Veerman
76–80	Gabor Kiss
81	Gabor Kiss (*top and bottom*), Ernest Rabinowicz (*middle*)
82	William J. Baxter and Stanley R. Rouse
85	National Aeronautics and Space Administration (*top*), U.S. Geological Survey (*bottom*)
88	National Aeronautics and Space Administration (*top*), Hale Observatories (*bottom*)
89	National Aeronautics and Space Administration
90	Richard M. Goldstein
91	Allen Beechel
92	William K. Hartmann, Planetary Science Institute (*top*); Allen Beechel (*bottom*)
93–98	Allen Beechel
99	National Aeronautics and Space Administration
101–102	Hebrew University
104–106	Alan D. Iselin
108	Robert J. Bull
110	M. C. Escher
111–121	Andrew Christie
124	*The Hungry Fly.* © 1976, Harvard University Press
126	Jerome Kuhl

"Designing a nuclear fusion reactor is like planning to reach heaven."

Imagine. Thermonuclear fusion, the energy source of the stars, an ultimate source of power that could last as long as the human race.

Many of the scientific and some of the technological questions of a decade ago now appear obsolete, and we are forging ahead with new enthusiasm. But like reaching heaven, building a fusion reactor represents an unknowable goal. We do know that the technological difficulties far surpass original estimates.



"The Prospect for Fusion" in December's Technology Review is a landmark article by David J. Rose, Professor of Nuclear Engineering at M.I.T., and Michael Feirtag of our Board of Editors. Use the coupon to begin your subscription with this significant article.

Realistic. Insightful. Important.

Technology Review

Please begin my subscription with the December issue for the term indicated below:

Domestic	Foreign
☐ 1 year $15	☐ 1 year $23
☐ 2 years $24	☐ 2 years $38
☐ Check enclosed	☐ Bill me

Name _____

Address _____

_____

To Technology Review, Room E19-430, M.I.T., Cambridge, Mass. 02139

© 1976 SCIENTIFIC AMERICAN, INC

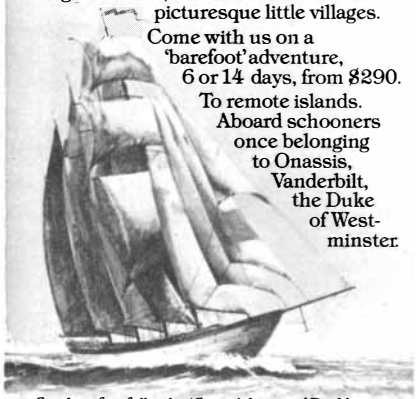
Authors...
LOOKING FOR A PUBLISHER?
Learn how to have your book published.

You are invited to send for a free illustrated guidebook which explains how your book can be published, promoted and marketed. Whether your subject is fiction, non-fiction or poetry, scientific, scholarly, specialized, (even controversial) this handsome 52-page brochure will show you how to arrange for prompt publication.

Unpublished authors, especially, will find this booklet valuable and informative. **For your free copy, write to:**
VANTAGE PRESS, Inc. Dept. F-53
516 W. 34 St., New York, N.Y. 10001

Sail a tall ship to a small island.

Saba. Carriacou. Nevis. Grenadines.
Bligh's breadfruit isle.
Stevenson's Treasure Island.
They are all here in the Caribbean waiting to be discovered again.
Along with unforgettable beaches, forgotten forts, hushed volcanoes and picturesque little villages.
Come with us on a 'barefoot' adventure, 6 or 14 days, from $290.
To remote islands.
Aboard schooners once belonging to Onassis, Vanderbilt, the Duke of Westminster.



Send my free full-color 'Great Adventure' Booklet.
Cap'n Mike, Windjammer 'Barefoot' Cruises
P.O. Box 120, Dept. 187, Miami Beach, Florida 33139

Name _____
Address _____
City/State/Zip _____

Windjammer Barefoot Cruises

LETTERS

Sirs:

In the September issue of *Scientific American,* devoted to food and agriculture, the lead article by Sterling Wortman is somewhat distressing. It says that attention and money should be focused on on-farm processes and problems if we are to win the battle of hunger throughout the world. Indeed, he seems to go much further and say that attention should be focused only on these problems if the battle is to be won.

I have no quarrel with any of the information presented so well by Dr. Wortman, particularly his identification of "two components to the solution of the food problem: increased production of food... and widespread increases in family incomes, particularly among the poor." There is, however, a third component that he and many other agriculturists persistently ignore. I always wonder why. He nicely describes three "nonsolutions" to the problem: continued export of food from developed nations, Western-style agricultural technology and synthetic foods. He offers a number of "significant and hopeful developments" that augur well for the two components he has identified, mentioning eight in all. These group conveniently under a smaller number of headings: a potential for raising on-farm yields, widespread increased understanding of the problem, the worldwide availability of chemical fertilizers and the existence of appropriate local and international institutions.

The big catch arises from the combination of (1) the absence in many places of means (systems) for the delivery of fertilizer and water and seed and expertise and other things in required quantities and qualities at the required times from the place where they are available to the place where they are needed, (2) the absence in many regions of suitable means (systems) for the delivery of farm produce to people and (3) the inefficiencies of both prefarm and postfarm systems that make losses on both sides a major problem. These three statements are only simplified descriptions of a massive systems problem that lies outside the farm.

I agree that it is necessary and desirable to achieve increased productivity on the farm and that it is necessary and desirable to increase farm-family incomes. Neither, however, is likely to put significantly more food on the plate of significantly more people throughout much of the developing world unless some attention is given to problems arising in the activities that come before the involvement of the farm and to those that come after the farm has yielded its produce, namely to the total interlocked and interacting chain of activities that can properly be called the food problem....

H. E. HOELSCHER

Professor
School of Engineering
University of Pittsburgh

Sirs:

Professor Hoelscher's points are good ones, particularly his observations about weaknesses of (1) delivery systems for supplies and information and (2) marketing systems broadly defined to include transport, storage, processing and delivery of produce to consumers. In short, the shift from subsistence agriculture to a highly productive agriculture requires that the entire system—before and after the farm, as Professor Hoelscher puts it—must function. No farmer will or can increase yields unless four sets of requirements are met in their entirety, as I suggested in testimony to committees of the House of Representatives in 1975 (available in "The World Food Situation: A New Initiative." Working Paper, The Rockefeller Foundation, 1133 Avenue of the Americas, New York, N.Y.):

"Farmers, even those who are uneducated and have small landholdings, will

Scientific American, January, 1977; Vol. 236, No. 1. Published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Gerard Piel, president; Dennis Flanagan, vice-president; Donald H. Miller, Jr., vice-president and secretary; George S. Conn, treasurer; Arlene Wright, assistant treasurer.

Editorial correspondence should be addressed to The Editors, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. Manuscripts are submitted at the author's risk and will not be returned unless accompanied by postage.

Advertising correspondence should be addressed to C. John Kirby, Advertising Director, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017.

Offprint correspondence and orders should be addressed to W. H. Freeman and Company, 660 Market Street, San Francisco, Calif. 94104. For each offprint ordered please enclose 35 cents.

Subscription correspondence should be addressed to Subscription Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. For change of address, notify us at least four weeks in advance. Send both old and new addresses and enclose an address imprint from a recent issue. (Date of last issue on your subscription is shown at upper right-hand corner of each month's mailing label.)

Name _____

New Address _____

_____

_____

Old Address _____

_____

_____

© 1976 SCIENTIFIC AMERICAN, INC

Audition free for 10 days Karl Böhm's superb Deutsche Grammophon recording of

Mozart's Six Greatest Symphonies

SPECIAL BONUS: Mozart's Piano Concertos Nos. 12 and 26

All 4 International Award-Winning Records For Less Than The Price Of 1!

Your Introduction To The Unique Money-Saving Program That Offers Free Auditions Of Prize-Winning Recordings By The Leading Concert Artists Of Our Time!

Wolfgang Amadeus Mozart—music's greatest natural genius—divinely gifted beyond any other musician who ever lived! And into his six greatest symphonies he poured a multitude of his most astonishingly beautiful, incredibly moving inspirations!

Symphony No. 41, "Jupiter"
Symphony No. 40 in G-minor
Symphony No. 39 in E-flat
Symphony No. 38, "Prague"
Symphony No. 36, "Linz"
Symphony No. 35, "Haffner"
Extra! Symphony No. 32

Now you are invited to hear these miraculous works in their finest recording . . . interpreted with extraordinary empathy by Karl Böhm . . . played to perfection by the Berlin Philharmonic Orchestra . . . captured in unsurpassed stereo realism by Deutsche Grammophon. So outstanding is this recording that it has won *three* of the music world's most eagerly sought honors: the Grand Prix International du Disque, Edison Award and Deutsche Schallplatten Prize! In addition, you will also receive Mozart's Piano Concertos Nos. 12 and 26, "Coronation," superbly performed by soloist Geza Anda with the Salzburg Camerata Academica — winner of the coveted Grand Prix des Discophiles.

Now enjoy and keep these 4 superb albums for less than the price you'd pay for 1!

Because these magnificent recordings have met with almost unprecedented acclaim, they have been chosen to introduce you to The Great Awards Collection, a totally new concept in home listening. You may enjoy all eight masterpieces, on four superb-quality, imported records, for 10 days absolutely free. Then keep all four, if you wish, for only $6.98 *(that's less than the price you'd pay for just one record)!* Simply mail the coupon at right today.

Here's how the program works . . .

As a member of The Great Awards Collection you will receive only the finest recordings of prize-winning concert performances by preeminent orchestras, conductors and soloists. There is no minimum number of recordings you must buy. And you can cancel your membership at any time. What's more, every superb recording you *choose* to examine comes to you on a free 10-day trial basis. So you'll never waste a penny on disappointing purchases!

Approximately 14 times a year you will receive the current issue of **our Great** Awards Collection Bulletin highlighting an exceptional Selection of the Month and other brilliant classical releases recommended by a distinguished jury of musicians, scholars and musicologists. If you want to examine the Selection of the Month you need do nothing. It will be shipped to you automatically. If you want any other selection, just mark your choices on the order card you will receive and return it by the date specified. If you do not want any selection, just check the proper box on your order card and return it by the date specified. It's that simple!

Special half-price bonus offer saves you money!

In addition to great music, you'll enjoy great savings with our half-price bonus plan. For every record you buy at our low members' price (always well below suggested retail), you may choose another one from a list of award-winners and other critically-acclaimed LPs and pay just half that price! A truly economical way to build your library of superb classical recordings.

How many records are you committing yourself to buy when you return the attached coupon? *None at all!* Even your introductory set of *Mozart's Six Greatest Symphonies* and *Piano Concertos 12 and 26* comes to you *on approval.* Audition it free for ten days, then either return it — or keep it for only $6.98 for all four records (plus a small postage/handling charge).

Here at last is the ideal way to acquire a connoisseur's record library in easy stages. You listen at home to every award-winning selection you elect to examine before deciding whether to buy it! Please act today! Begin your money-saving, no-obligation membership in The Great Awards Collection by mailing the coupon for your free trial, with four-for-less-than-the-price-of-one purchase option of this essential collection of Mozart's Six Greatest Symphonies and exquisite Piano Concertos Nos. 12 and 26.

Eight ways The Great Awards Collection helps you enjoy fine music more than ever!

1. The greatest music by the world's immortal composers.
2. A major, award-winning recording yours to audition each month.
3. Superior imported LPs — silent surfaces.
4. Free ten-day trials of all selections.
5. All selections always far less than retail price.
6. Extra savings with Half-Price Bonus Plan.
7. No obligation to buy — no minimum purchase.
8. Money-saving introductory offer.

Send no money now—Mail coupon TODAY!

Listen for 10 days free. Keep all 4 for less than the price of 1!

The Great Awards Collection • 175 Community Drive • Great Neck, N.Y. 11025

Please send me, for my free audition, the triple-award-winning recording of Mozart's Six Greatest Symphonies by Karl Böhm, together with Geza Anda's award-winning performance of Mozart's Piano Concertos Nos. 12 and 26. I may return the set after 10 days and owe nothing, or keep it and pay only $6.98* for all four Deutsche Grammophon records, plus a small postage/handling charge. This is all four records for less than the regular price of one!

Each month I will also be notified of major award-winning recordings available to Great Award Collection members free for 10 days. For every one of these records I buy at your members' price (always below suggested retail), I may also choose one record at half that price from your special list. I am not obligated to buy any minimum number of records. I may cancel my membership at any time.

NAME _____
(please print)

ADDRESS _____ APT NO _____

CITY _____ STATE _____ ZIP _____

Offer limited to the 48 connecting states. Offer expires 6/30/77. Limit: one membership per household. Only new members eligible. 59203
NOTE: All applications are subject to review, and we reserve the right to reject any application. *Sales tax added for New York residents.

© 1976 SCIENTIFIC AMERICAN, INC

Physics...concrete math...astronomy and astrophysics...chemistry... probability...structural geology... undoubtedly the best books in every scientific area that interests you.

Take any 3 books, values to $60.85, all for only $3.95 as a member of

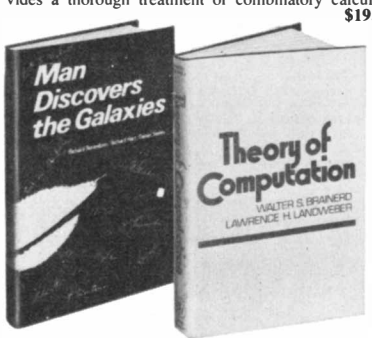
The Library of Science

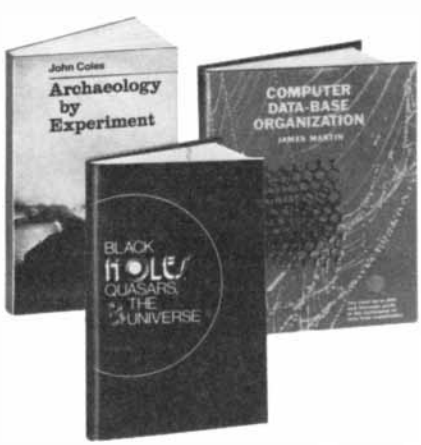
if you will join now for a trial period and agree to accept only 3 more books—at members' discount prices—over the next 12 months.

(Publishers' Prices shown)

59800. MAN DISCOVERS THE GALAXIES. *Richard E. Berendzen, Richard C. Hart & Daniel Seeley.* Topical, current look at modern astronomy dealing mostly with the discovery and arrangement of our Home Galaxy, the Milky Way System and the discovery of the basic properties of the universe of galaxies in which we find ourselves. **$15.95**

83760. THE THEORY OF COMPUTATION. *Walter S. Brainerd and Lawrence H. Landweber.* Demonstrates the equivalence of various algorithmic languages, treats unsolvable problems emphasizing methods for obtaining undecidability results, studies problems known to be effectively computable and provides a thorough treatment of combinatory calculus. **$19.95**





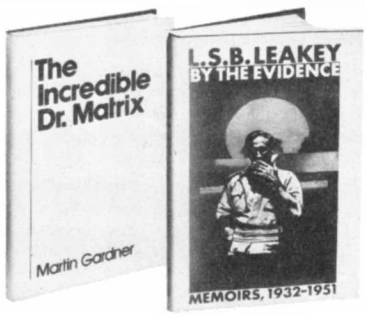
36395. BLACK HOLES, QUASARS AND THE UNIVERSE. *Henry L. Shipman.* Black holes, quasars, white dwarfs, neutron stars, supernovae, Seyfert Galaxies — explores the frontiers of astronomy. **$12.95**

39865. COMPUTER DATE-BASE ORGANIZATION. *James Martin.* How you can use the latest design and implementation techniques to improve your data system's real-time accessibility and improve flexibility — and do it at the lowest possible cost. *Counts as 2 of your 3 books.* **$26.50**

34400. ARCHAEOLOGY BY EXPERIMENT. *John Coles.* Scientific overview of an exciting new branch of archaeological research. Spotlights such archaeological experiments as the highly sophisticated simulated Iron Age settlements that have been established today. **$7.95**

55190. THE INCREDIBLE DR. MATRIX. *Martin Gardner.* The delightful incredibly brilliant Dr. Matrix returns with more amazing feats and adventures in number theory, sleight of word, numerological analysis and puzzles. **$8.95**

37195. BY THE EVIDENCE. *L.S.B. Leakey.* Fascinating reading of Leakey's early years in Africa and his development into one of the most respected scientists of the twentieth century. **$9.95**



© 1976 SCIENTIFIC AMERICAN, INC

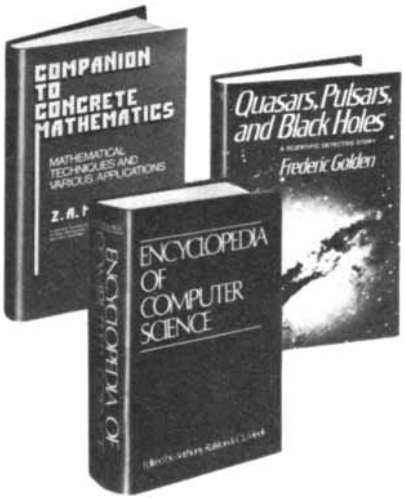
34140. APPLIED ANALYTICAL MATHEMATICS FOR PHYSICAL SCIENTISTS. *James T. Cushing.* Deals with topics ranging from spectral analysis of linear operations to complex variables, to group theory. **$19.95**

61530. MATHEMATICS DICTIONARY. Fourth Edition. *Edited by Robert C. James and Edwin F. Beckenbach.* Revised and expanded edition of this leading mathematics dictionary offers clear, accurate definitions of more than 8,000 terms. **$17.95**

70190. PRINCIPLES OF OPERATIONS RESEARCH. Second Edition. *Harvey M. Wagner.* This 1039-page book offers a complete overview of all the fundamental concepts. **$19.95**

49340. FUNDAMENTAL ASTRONOMY: SOLAR SYSTEM AND BEYOND. *Franklyn W. Cole.* Traces our assault on the mysteries of the universe from the earliest sensory-oriented glimmerings to modern astronomical thought. **$16.50**

81780. SYMMETRY DISCOVERED. *Joe Rosen.* Fascinating look into one of the most beautiful and practical of the sciences. **$11.95**



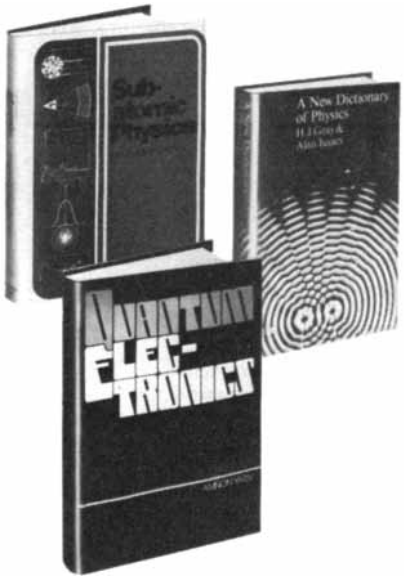
44900. THE ENCYCLOPEDIA OF COMPUTER SCIENCE. *Edited by Anthony Ralston and Chester L. Meek.* A monumental 1,550 page volume, offering condensed, accurate information on essential topics ranging from theory to applications, from program structures to the logical design of central processing units, from Boolean algebra to microprogramming. Bibliographies, cross references, complete indexing and 700 charts and tables. *Counts as 3 of your 3 books.* **$60.00**

46500. THE EVOLUTION OF RADIO ASTRONOMY. *J.S. Hey.* One of the pioneers of astronomy offers an absorbing history of this powerful technique's remarkable development. "...an important work in the history of astronomy." — M.L. Meeks, in *Sky and Telescope.* **$10.00**

73240. QUASARS, PULSARS, AND BLACK HOLES. *Frederic Golden.* A thorough review of the spectacular discoveries in astronomy, as told by the science editor of *Time* magazine. Illustrated with twelve drawings and twenty-four color photographs. **$7.95**

39745. COMPANION TO CONCRETE MATHEMATICS. *Z.A. Melzak.* "A delightful collection of problems organized into over sixty sections, together with historical notes, solutions, observations, and remarks. Mathematicians at all levels will enjoy sampling this work." — *American Mathematical Monthly.* **$16.50**

IMPORTANT GUARANTEE! Examine your 3 introductory books for 10 days. If not delighted, simply return them and your membership will be canceled. You pay nothing, owe nothing!



32300. ADVENTURES OF A MATHEMATICIAN. *S.M. Ulam.* First-hand account of the birth and childhood of the Atomic Age. **$14.95**

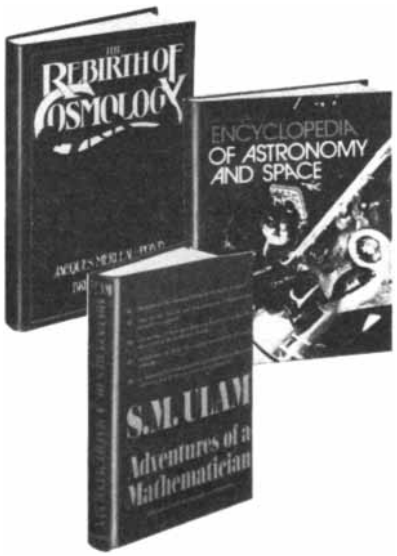
81150. SUBATOMIC PHYSICS. *Hans Frauenfelder and Ernest M. Henley.* Covers everything from how accelerators and detectors work to the structure of subatomic particles. **$17.50**

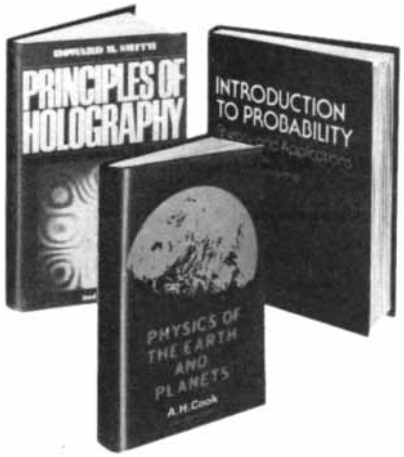
64145. THE NEW DICTIONARY OF PHYSICS. *Edited by H.J. Gay and Alan Isaacs.* Over 9,000 entries in 600 pages. A monumental reference that is intermediate between a strict dictionary and an encyclopedia. About 25% greater than the 1958 edition, it reflects all major changes characteristic of the field. *Counts as 2 of your 3 books.* **$35.00**

82660. TECHNOLOGY MATHEMATICS HANDBOOK. *Jan J. Tuma.* Instant access to concise summaries of all major definitions, formulas, graphs, tables of elementary and intermediate mathematics — from algebra to determinants and vectors and much, much more. **$15.95**

73165. QUANTUM ELECTRONICS. *Amnon Yariv.* Classic work incorporates all the latest advances in quantum electronics. Containing 70% new material, it deals with quantum mechanics, electromagnetism, laser oscillation, nonlinear optics, and much more. **$20.95**

73530. THE REBIRTH OF COSMOLOGY. *Jacques Merleau-Ponty and Bruno Morando.* A thorough account of the Greek, classical and astrophysical ages of the fascinating study of the universe as a whole physical system. **$12.50**





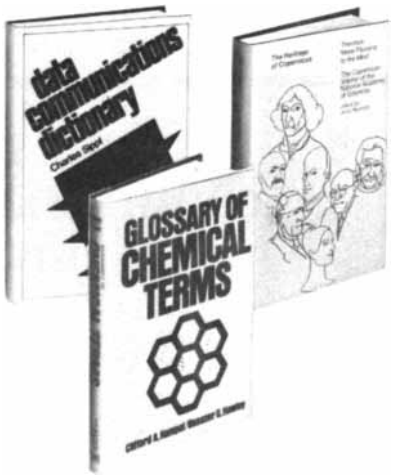
70170. PRINCIPLES OF HOLOGRAPHY. Second Edition. *Howard M. Smith.* From fundamentals to rigorous theory, from chemical formulas to applications. **$16.50**

56380. INTRODUCTION TO PROBABILITY. *Richard L. Scheaffer and William Mendenhall.* Covers everything from Markov chains to the arrival times for the Poisson process. **$13.50**

68735. PHYSICS OF THE EARTH AND PLANETS. *Alan H. Cook.* Brilliant investigation by one of the world's leading authorities on physical meteorology. *Counts as 2 of your 3 books.* **$24.75**

49267. FRONTIERS OF ASTROPHYSICS. *Edited by Eugene Avrett.* A contemporary source on astrophysical topics from recent solar research, to neutron stars, supernovae and intergalactic matter. *Counts as 2 of your 3 books.* **$20.00**

50380. GLOSSARY OF CHEMICAL TERMS. *Clifford A. Hampel and Gessner G. Hawley.* Over 2,000 concise definitions: all major chemical groups, important functional terms, basic phenomena and processes, all the chemical elements, and more. **$14.95**



52650. THE HERITAGE OF COPERNICUS. *Edited by Jerzy Neyman.* 25 highly readable essays on the quasi-Copernican revolutions that have occurred in the past five centuries. *Counts as 2 of your 3 books.* **$30.00**

66570. AN OUTLINE OF STRUCTURAL GEOLOGY. *Bruce E. Hobbs, Winthrop D. Meansand Paul F. Williams.* Up-to-date account of every aspect of structural geology — from the nature of stress and strain to the latest information on plate tectonics. **$16.95**

56590. INTRODUCTORY GROUP THEORY AND ITS APPLICATION TO MOLECULAR STRUCTURE. *John R. Ferraro.* Includes derivation of selection rules, potential force fields, and added emphasis on space symmetry and point symmetry. **$19.95**

41640. DATA COMMUNICATIONS DICTIONARY. *Charles J. Sippl.* Over 14,500 definitions and explanations of data processing and communications terms, acronyms, concepts and abbreviations. **$19.95**

55000. THE ILLUSTRATED ENCYCLOPEDIA OF ASTRONOMY AND SPACE. *Ian Roxburgh.* Covers the full range of astronomical knowledge and space exploration, from ancient times to now. **$16.95**

If the reply card has been removed, please write to
The Library of Science
Dept. 2-4AJ, Riverside, N.J. 08075
to obtain membership information and an application.

© 1976 SCIENTIFIC AMERICAN, INC

"The Best Mystery With Each Clue Unfolding"

BIBLICAL Archaeology Review

"I await each of your issues with added enthusiasm. It is like reading the best mystery with each clue unfolding."
Shirley Woldar, Manhasset, New York

"Thank you for making **BAR** so interesting that even a professional archaeologist can enjoy it."
Prof. Emanuel Levine, Rider College

"Your magazine is marvelous. As a church librarian, I can certainly recommend it."
Mrs. Fran Rowe, Tahlequah, Oklahoma

"If it were possible, I think I could do nothing but read from your magazine for the rest of my life."
Blaine Huebner, Sheboygan, Wisconsin

. . . In language written to be understood.

Enclosed is my check for $8.50 for a 1-year's subscription. Full refund if I'm not delighted with your magazine.

Name _____

Address _____

_____ Zip _____

Send to: **The Biblical Archaeology Review**
1819 H St. N.W., Wash., D.C. 20006 AS

Plant a THINK TANK anywhere and watch the minds grow!

home-office-school-park-club-churches-laboratory

Unique instructional games designed by university professors to make learning fun through brain-to-brain action. Beginning games can be mastered by young children— final games will challenge intelligent adults. These are the famous GAMES FOR THINKERS from WFF 'N PROOF Publishers.

WFF 'N PROOF (logic)	13.00*
QUERIES 'N THEORIES (science)	13.00*
EQUATIONS (mathematics)	10.00*
ON-SETS (set theory)	10.00*
PROPAGANDA (social studies)	11.00*
ON-WORDS (word structures)	10.00*
CONFIGURATIONS (geometry)	6.75*
TRI-NIM (problem solving)	5.75*
REAL NUMBERS (arithmetic)	2.25*
WFF (beginner's logic)	2.25*
QWIK-SANE (puzzle)	2.25*
TAC-TICKLE (pure strategy)	1.75*
Teachers Manual	1.25*
THINKERS Bookends	16.00*
12-Kit THINK TANK & Teachers Manual With THINKERS Bookends (save $10.75)	94.50*
Without Bookends (save $4.75)	84.50*

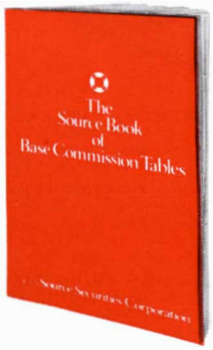
*postage & handling included

Order from: **WFF 'N PROOF**
1490-EF, South Boulevard, Ann Arbor, Mi. 48104
Phone Free 800-528-6050 Ext. 702

Gifts that are a COMPLIMENT to receive!

A useful guide for the commission-conscious investor.

This free, pocket-sized book enables you to quickly figure and compare practically any broker's stock and option commissions. Its 90 pages of tables display the old fixed commission* and commissions at a wide range of discounts from it, in cents-per-share and dollars-per-contract rates for almost any whole-priced trade.



Get this book simply by mailing this ad or asking for one at **800-221-5338** toll-free or in N.Y. State at **212-269-9127** collect.

From the people who started commission discounts rolling for the individual investor.

*Base Commissions. Pre-May 1, 1975 NYSE and CBOE-fixed commissions.

Name _____

Address _____

City _____ State _____ Zip _____

Source Securities Corporation
70 Pine Street, New York, N.Y. 10005

AUG SCA 1/77

adopt new systems provided that four conditions are met:

"1. There must be available to them more highly productive and more highly profitable farming systems. These systems must be complete. For crops, this means suitable varieties, plus fertilizer-use practices, plus means of disease and insect control, plus associated cropping practices. Development of such systems involves sophisticated work which properly trained scientists can do, but which is beyond the capabilities of the usually poorly trained 'extension specialist.'

"2. The necessary inputs—fertilizers, seed, pesticides, and credit—must be available to the farmer when and where he needs them, and at a reasonable price. Systems for distribution of these products in the rural areas must be in place, and must function.

"3. The farmer must be shown, usually on his own land or that of a neighbor, how to utilize effectively the new technological system.

"4. The farmer must know before he invests in new plantings or other operations that there will be at harvest a market for his product at a price upon which he can depend, for the small farmer cannot take great risks. A 'market' requires roads, transport, effective demand for products, favorable prices. . . .

"Well-organized campaigns are needed now to force the pace of agricultural development, moving it at a speed with which few nations have had experience. There must be fast-moving scientific efforts to develop the technology including the identification of the complete new systems for use at the farm level. Provisions must be made for the supply of the inputs and for marketing. Road networks need to be extended, both for supply of inputs and for marketing of produce. Power grids must be elaborated, means of using the media must be devised, systems of providing credit to small farmers, requiring additional innovation, must be set up—in short, the full range of institutions required for a market-oriented agriculture must function. Great numbers of people will need to be trained, mostly on the job, as they participate in direct efforts to develop agriculture and rural areas."

In concurring with Professor Hoelscher, I do not wish to imply that on-farm improvements are any less important than indicated in my article in the September issue of *Scientific American*. In agrarian countries on-farm gains in productivity coupled with increased incomes are prerequisites to rural and general economic development.

STERLING WORTMAN

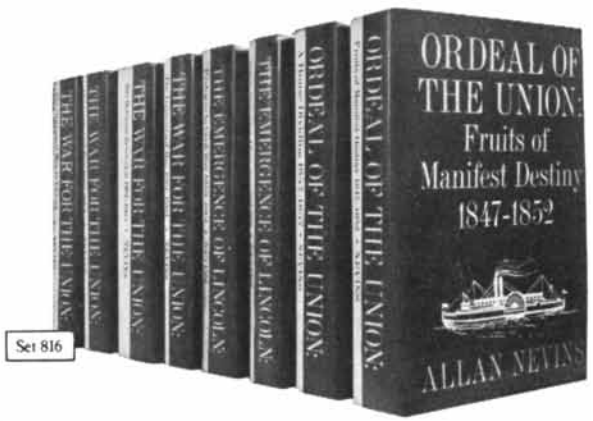
Vice-President
The Rockefeller Foundation
New York

© 1976 SCIENTIFIC AMERICAN, INC

Drink deeply from the wellspring of history...

Choose Any Set For A Dollar A Volume

A rare opportunity to enrich your library with one of these enduring works, at a fraction of its usual cost. Savings up to $112, with trial membership.



Allan Nevins'

THE ORDEAL OF THE UNION

8 Volumes, Publisher's Price $120.00/ *Yours For Only $8*

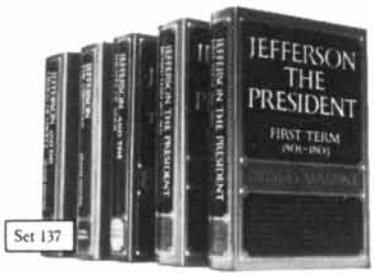
The masterwork of this great historian. Monumental account of our national history, 1847 to 1865. Told with a sweep and passion that is almost Tolstoian. *"We are not likely to see more history of this character and scope for some time to come."* —N.Y. Times, National Book Award.

Fernand Braudel's

THE MEDITERRANEAN: And the Mediterranean World in the Age of Philip II

2 Volumes, Publisher's Price $35.00/ *Yours For Only $2*

A landmark social history, these two volumes are classics on the 16th century Mediterranean world. "Probably the most significant historical work to appear since World War II." —*N. Y. Times*



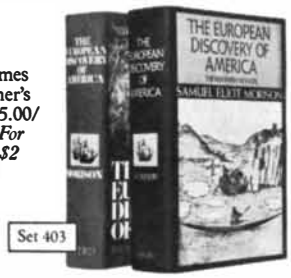
Dumas Malone's Pulitzer Prize-Winning

JEFFERSON AND HIS TIME

5 Vols. Pub. Price $72.50/ *Yours For Only $5*

From his Virginia youth through his troubled second term as president. Biography surpassing anything you've ever read on Jefferson and his world. 1975 Pulitzer Prize.

2 Volumes Publisher's Price $35.00/ *Yours For Only $2*

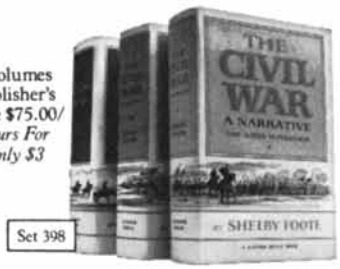


Samuel Eliot Morison's

THE EUROPEAN DISCOVERY OF AMERICA

The crowning achievement of one of America's foremost historians. A brilliant narrative account of the Atlantic voyages, from 500 A.D., through 1616. Over 450 illustrations, 70 maps.

3 Volumes Publisher's Price $75.00/ *Yours For Only $3*



Shelby Foote's

THE CIVIL WAR: A Narrative

By a historian who is also a gifted novelist—the opening shot at Fort Sumter through the surrender at Appomattox. "A remarkable achievement, prodigiously researched, vigorous, detailed, absorbing." —N.Y. Times

The story of history, and the men who made it, makes for reading that is both lively and valuable.

And the best of written history is the province of The History Book Club.

Now in its 30th year, the Club provides its members with the finest new books of history and world affairs—always at substantial savings. If you've never tried us, you've missed some great reading and unusual buys. But none like this:

For a mere dollar a volume, we'll send you any one of the magnificent sets pictured—with your first selection at the reduced member's price (any book listed at right).

If you're less than delighted, return all the books (we'll pay postage) and that's that. Or keep them and take just four more selections during the coming year from the 150-200 offered each month.

They're easy to take at $3, $4, $5 (and more) off bookstore prices. No stinting on physical quality either; all our books are the equal of the publishers' editions. And this Club has a Bonus Plan that lets you pick *any* book we offer, not just certain ones. Use the coupon to join today.

PICK ONE OF THESE SELECTIONS AT THE REDUCED CLUB PRICE:

(First price is publisher's list. **Boldface** shows member's price.)

138. **7 Who Shaped Our Destiny: The Founding Fathers as Revolutionaries** *By Richard B. Morris.* **$10.00/$7.50**

464. **The Twelve Caesars** *By Michael Grant.* From Julius in 49 B.C. through Domitian in 96 A.D. **$12.50/$8.75**

104. **The Indian Heritage of America** *By Alvin M. Josephy, Jr.* **$10.00/$7.50**

671. **Medieval History** *By Norman F. Cantor.* **$11.50/$8.50**

116. **A History of the Vikings** *By Gwyn Jones.* **$12.50/$8.50**

247. **A History of American Law** *By Lawrence M. Friedman.* **$15.95/$9.75**

457. **A True History of the Assassination of Abraham Lincoln** *By Louis Weichmann.* **$15.00/$9.95**

139. **Disease and History** *By Frederick F. Cartwright.* **$9.95/$7.50**

655. **Roosevelt and Churchill 1939-1941: The Partnership That Saved the West** *By Joseph P. Lash.* **$12.95/$8.95**

625. **The Black Family in Slavery & Freedom, 1750-1925** *By Herbert G. Gutman.* **$15.95/$10.75**

WH77A

THE HISTORY BOOK CLUB • Stamford, Conn. 06904

Please enroll me as a trial member and send me the introductory books whose numbers I'm entering below. Bill my set at $1 a volume and my first selection at the reduced Club price (plus modest shipping charges).

Set # _____ Selection # _____

I may return all the books within three weeks at your expense and owe nothing. Or I will take four more selections within a year *(not counting my introductory selections)*—at reduced Club prices plus shipping. Thereafter, I may resign at any time.

SA-09W

Print Name _____

Address _____

City _____ State _____ Zip _____

In Canada: Mail to the History Book Club, 16 Overlea Blvd., Toronto M4H 1A6, Ont.

A note on how the Club works: Every four weeks, you receive our *Review* and a dated reply card. If you want the Editors' Choice, do nothing: the book will come automatically. To order a different book or no book, return the card by the specified date. You will have at least 10 days to decide—but if you get an unwanted book because of late mail, we pay return postage.

© 1976 SCIENTIFIC AMERICAN, INC

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

JANUARY, 1927: "The Raymond Orteig prize of $25,000 for the first nonstop flight between New York and Paris still remains to be won. Many individuals and organizations are considering ways and means of achievement. R. C. Wood describes in *U.S. Air Services* a long-distance record flight of 2,920 miles made in a standard type Bréguet plane by Captain Girier and Lieutenant Dordilly of the French Air Service. Leaving Paris early on a Wednesday morning, the aviators arrived over Cologne in two hours, flew over Berlin and Dantzig and then across the plains of Russia. They landed in Omsk in Siberia after only 29 hours of flight."

"Control of sources of rubber by the British government has led to much discussion of the possibilities of producing synthetic rubber from sources other than the sap of *Hevea brasiliensis*. During this past summer Prof. Roger Adams of the University of Illinois made the following statement about the theoretical possibilities of synthesis: 'Rubber has been made synthetically and the types of raw material necessary are well known. But the serious problem is to find a source of this raw material which is cheap enough to make possible competition of synthetic rubber with the natural. Petroleum offers a possibility. When the high-boiling petroleum is cracked in order to obtain low-boiling fractions which can be used as gasoline in internal-combustion engines, there are contained in these low fractions butadiene and its derivatives, the type of compounds which have been shown to be convertible into synthetic rubber. It remains for the scientists to find out how the yield of these butadienes may be increased and how they may be economically removed from the other closely related products which accompany them.'"

"The class of aluminum alloys known collectively as 'duralumin' came into prominence in America about five years ago. It is a metal combining the lightness of aluminum with the strength of mild steel, and having great freedom from deterioration through rust and corrosion. Duralumin is an especially suitable metal for aircraft construction, and its development in this country was largely due to the efforts of the Navy Department in connection with the building of the great airship *Shenandoah*. The use of duralumin has since been greatly extended, and it is likely to come into increasing favor wherever lightness combined with strength and durability is required. Duralumin is an alloy containing about 3.5 to 4.5 per cent copper, 0.2 to 0.75 per cent magnesium, 0.4 to 1.0 per cent manganese and not less than 92 per cent aluminum. When it is not heat-treated, its physical properties differ but little from those of ordinary aluminum. Its great merit lies in the fact that it is susceptible to heat treatment, by which its strength, ductility and resistance to corrosion may be greatly increased."

"With the recent death of Harry Houdini there passed out the most famous 'magician' of his time, and as some would have it, of all time. Over and above the reputation due to his amazing cleverness there was an air of honesty and sincere purpose about the man that added greatly to his popularity. So versatile was he that he covered both the field of what we might call table magic, which required not much apparatus but a marvelous dexterity of hand, and that which won him his greatest reputation, where he performed his feats with the aid of special paraphernalia. It was probably owing to his inborn character that he seemed to find more satisfaction in his successful exposure of imposture and fraud than in those brilliant feats of legerdemain and escape from handcuffs, prison cells and metal caskets. It was this fact that led to his inclusion on the committee of SCIENTIFIC AMERICAN for the investigation of so-called spiritualistic phenomena, in whose investigations and conclusions he rendered valuable assistance."

SCIENTIFIC AMERICAN

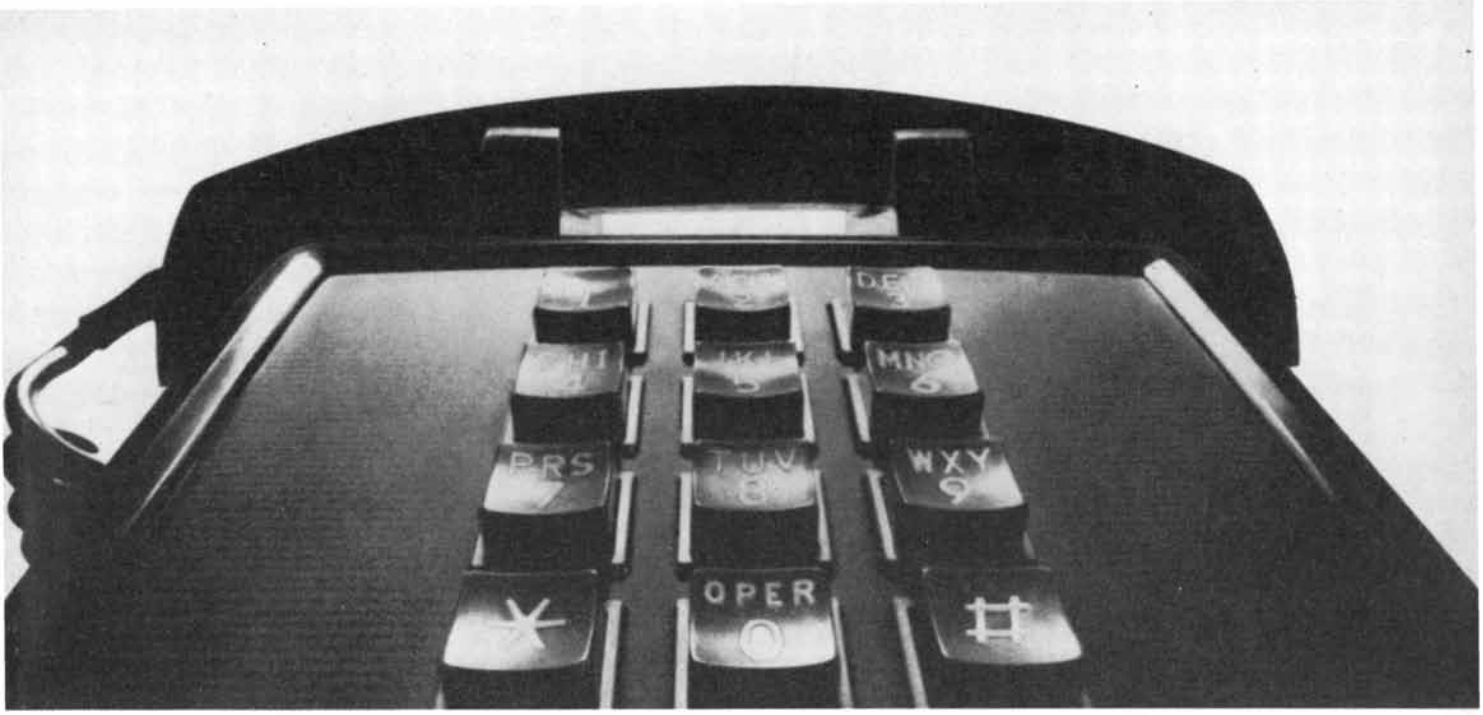
JANUARY, 1877: "Dr. Huggins has made an important advance by his successful application of photography to the spectra of stars, of which he gives an account in a communication to the Royal Society. Although he has up to the present confined his attention to the bright star Vega, which has a well-marked spectrum and is therefore especially adapted to the purpose, there can be little question that he will before long obtain good photographs of the spectra of other bright stars, which will add greatly to our knowledge of their constitution by enabling us to examine the invisible part of their spectra in the ultraviolet, besides giving means for more accurate determination of the position of lines than is ordinarily possible. There is one great advantage photography has over the human eye, namely the length of time during which the effect produced by the luminous body accumulates, a circumstance that makes up for the inferior sensitiveness of the photographic film, and it is by taking advantage of this and giving an exposure of several hours that Dr. Huggins has obtained such good results."

"A despatch from M. Tellier to the French Academy of Sciences announces the arrival of the refrigerator ship *Frigorific* at Pernambuco, Brazil, after a voyage of 70 days from France. The meat transported and kept cool by cold air generated by the Tellier ice machine (which works on the principle of evaporating methylated spirit) was perfectly preserved. It is proposed to load the vessel with meat at La Plata, Argentina, and if the return voyage is successfully accomplished, to establish the regular exportation of South American beef to French markets."

"The President's message, which last year contained about 7,100 words, was transmitted by the Western Union Telegraph Company from Washington to New York, dropping copies at Baltimore and Philadelphia in twenty-eight minutes, without a stop or an error. Ten wires were used. In less than two hours from the time it was filed at Washington the message in full was in the hands of the press in every important town in the country, without any disarrangement to the business of the company."

"In these times it has become fashionable to talk of modern degeneracy and to look back regretfully to the 'good old times.' It is therefore pleasant, for a change, to hear such sentiments as are embodied in the following extract from the annual message of Governor Hartranft of Pennsylvania: 'A hundred years have brought a wonderful change. The population has increased tenfold, the area under cultivation a hundredfold and wealth almost beyond comparison. Thousands of miles of canals and railroads intersect the Commonwealth. Immense mining, manufacturing, agricultural and carrying enterprizes give employment to the toiling millions of the State. All the products of the earth are within our reach; fuel and provisions are brought to our doors; gas and water are in our houses, and the news of the world of yesterday is laid on our breakfast tables in the morning. Thousands of schools and colleges are scattered over the State, and the post is burdened daily with millions of letters, attesting the general diffusion of knowledge. The people are more intelligent, freer and happier; more cheerful, tolerant and liberal. The charges of modern degeneracy are refuted by the clear testimony of a hundred years.'"

© 1976 SCIENTIFIC AMERICAN, INC



Behind the buttons, there's a "brain."

Electronic intelligence has come to the phone system.

It's part of our new electronic switching machines, working round-the-clock to give you better phone service.

And it's another innovation in telecommunications from Bell Labs and Western Electric.

Intelligent Systems.

Conventional electromechanical switching couldn't have kept up with the enormous increase in telephone calling. Only electronic switching can do that. So we've developed a family of electronic machines that are handling calls in rural, suburban and metropolitan areas.

And our new long-distance electronic system handles 550,000 calls an hour—four times as many as before.

So your Bell telephone company can continue to respond to your needs better and, at the same time, keep your costs reasonable.

Since electronic switching machines run on instructions that can be easily changed, their operations and services can be changed without rewiring or adding equipment.

This flexibility makes many new services practical. Like having your calls forwarded to another number or being signalled the instant a busy line becomes free. The instructions are just programmed into the system's memory.

Electronic switching systems can do more because they're much faster than electromechanical systems. So even though they're handling many more calls, electronic systems have time left over for other tasks.

They continuously check themselves for trouble, diagnose any they find, signal for help, and route calls around a trouble spot.

But trouble is rare with electronic switching.

Today's systems are working round-the-clock with less than 12 minutes "downtime" a year on the average. That makes them 99.9998% reliable.

Our Biggest Job Ever.

The development of electronic switching was the biggest project ever undertaken by Bell Labs and Western Electric in our efforts to keep your phone bill down.

It required an enormous investment in research and development.

Building it took even more work and money.

But now there are 1000 electronic switching offices serving communities of every size as well as handling long distance calls. And a new office is being installed about every working day.

The System Works.

It took the Bell System to develop electronic switching.

Because only the Bell System has Bell Laboratories and Western Electric working together to make your telephone system the most reliable and least expensive in the world.

Which is to say, we had the brawn to put a "brain" behind your phone.

One of a series of messages to keep you informed of how telecommunications technology is changing our world—and the part Bell Labs, Western Electric and your Bell Telephone Company are playing in it.

Western Electric / Bell Laboratories

© 1976 SCIENTIFIC AMERICAN, INC

THE AUTHORS

CHRISTOPHER TIETZE and SARAH LEWIT ("Legal Abortion"), husband and wife, have separately and jointly written more than 200 articles, mostly in the field of human fertility and its regulation. Tietze, a native of Austria with an M.D. from the University of Vienna, came to the U.S. in 1938. He has worked for the Johns Hopkins University School of Hygiene and Public Health, the Department of State and the National Committee on Maternal Health; he is currently senior fellow at the Population Council's Institute of Policy Research. Lewit, a graduate of Brooklyn College, was employed by the National Vital Statistics Center and the National Committee on Maternal Health until 1971, when she joined the Population Council as a research associate. Since 1973 she has worked as a free-lance writer. Tietze and Lewit were corecipients of the Margaret Sanger Award in 1973.

GLOVER B. TRIPLETT, JR., and DAVID M. VAN DOREN, JR. ("Agriculture without Tillage"), are agronomists at the Ohio Agricultural Research and Development Center. Their association goes back to the late 1950's, when they were graduate students at Michigan State University. Triplett, who had obtained a B.S. and an M.S. at Mississippi State University, went on to acquire his Ph.D. in farm crops from Michigan State in 1959, a year after Van Doren, who had done his undergraduate work at the University of Illinois, received his Ph.D. in soil science. Both joined the Ohio agricultural center soon after graduation, and they have worked there together ever since. Their research on no-tillage systems utilizes their complementary skills: Triplett, a specialist in crop management, has worked on weed control; Van Doren, a soil physicist, has studied the influence of soil characteristics on plant performance.

S. W. HAWKING ("The Quantum Mechanics of Black Holes") is a theoretical physicist at the University of Cambridge. He was born in Oxford in 1942 and was graduated from the University of Oxford in 1962. He did his graduate work at Cambridge on general relativity, working under the direction of D. W. Sciama. He is currently a fellow of Gonville and Caius College at Cambridge and reader in gravitational physics in the university's department of applied mathematics and theoretical physics. In 1974–1975 he was Sherman Fairchild Distinguished Scholar at the California Institute of Technology. A Fellow of the Royal Society, he has received a number of honors in the past two years, including the Eddington Medal of the Royal Astronomical Society and the Dannie Heineman Prize for Mathematical Physics of the American Physical Society and the American Institute of Physics. Since 1962, when he began his graduate work at Cambridge, Hawking has suffered from a progressive nervous disease that has confined him to a wheelchair for the past seven years. "Fortunately," he writes, "theoretical physics is one of the few fields in which this is not a serious handicap."

J. DONALD CAPRA and ALLEN B. EDMUNDSON ("The Antibody Combining Site") are immunologists at two medical institutions in the Southwest. Capra, who is professor of microbiology at the University of Texas Health Science Center in Dallas, has an M.D. from the University of Vermont. He did research at the National Institutes of Health, Rockefeller University and the Mount Sinai School of Medicine in New York before he moved to Texas in 1974. Edmundson, who is professor of biochemistry at the University of Utah Medical Center in Salt Lake City, received his Ph.D. in biochemistry from Rockefeller University in 1961. Before he took up his present job he worked as a U.S. Public Health Service Fellow in the molecular biology unit of the British Medical Research Council and as a biochemist at the Argonne National Laboratory.

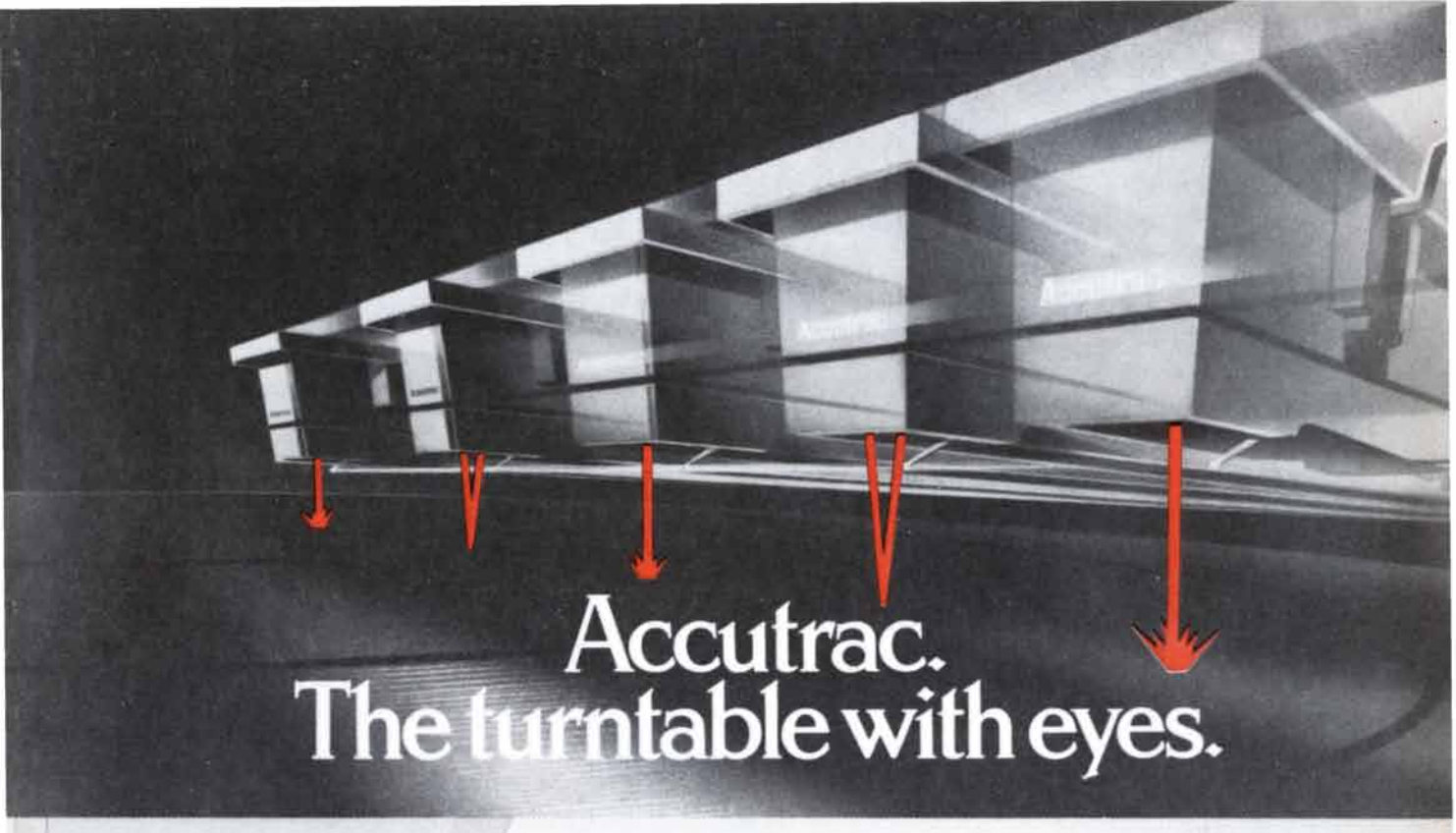
ROBERT SEKULER and EUGENE LEVINSON ("The Perception of Moving Targets") are at Northwestern University, where Sekuler is professor and chairman of the department of psychology and Levinson is a postdoctoral fellow and visiting scholar in the university's Biomedical Engineering Center. Sekuler is a graduate of Brandeis University and holds a Ph.D. from Brown University. He joined the faculty at Northwestern in 1965, after having spent a postdoctoral year at the Massachusetts Institute of Technology. He describes himself as a "fanatical sailor" who is "also active in the Greater Chicago branch of the American Begonia Society." His collaboration with Levinson on a number of projects in visual perception began in 1971, when Levinson took an undergraduate course on perception taught by Sekuler. Levinson went on to do graduate work at Dalhousie University, returning to earn his Ph.D. from Northwestern in 1976. His current research interests are in the area of retinal physiology and pharmacology.

ERNEST RABINOWICZ ("Exoelectrons") is professor of mechanical engineering at the Massachusetts Institute of Technology. His degrees are from the University of Cambridge: a B.A. in physics in 1947 and a Ph.D. in physical chemistry in 1950. He has been at M.I.T. since 1950, except for a semester as visiting professor at the Technion in Haifa in 1969. An expert in the fields of friction and wear, he writes: "Hardly a technical event occurs in the world, be it a new space probe, a new nuclear reactor, a new computing system or a new engine, without a new friction or wear problem being created, and every once in a while I get the opportunity to contribute to a solution. The students in my mechanics courses are divided on whether I teach them much mechanics or whether they have to read the textbook to learn the material themselves, but they agree that I tell the best stories about my past consulting experiences."

WILLIAM K. HARTMANN ("Cratering in the Solar System") works in Tucson as senior scientist for the Planetary Science Institute, a division of Science Applications, Inc. He studied physics as an undergraduate at Pennsylvania State University, then went on to obtain an M.S. in geology and a Ph.D. in astronomy from the University of Arizona. He has investigated a variety of problems related to the origin of planetary systems and the evolution of planetary surfaces. He is the author of *Moons and Planets,* an introductory textbook on planetary science, and is in the process of preparing a new textbook on general astronomy. This is his second article for SCIENTIFIC AMERICAN; his first, "The Smaller Bodies of the Solar System," appeared in the September 1975 issue titled "The Solar System."

SHEMARYAHU TALMON ("The Samaritans") is professor of Bible studies at the Hebrew University of Jerusalem, where he also serves as dean of the humanities faculty. A native of Germany, he emigrated to Israel in 1939. He obtained his master's degree in 1945 and his Ph.D. in 1958, both from the Hebrew University. He taught there for a number of years before he became head of Haifa University College in 1968. In recent years he has spent much of his time lecturing in Europe and the U.S. He was a visiting professor at Harvard University in the academic year 1970–1971 and at Brandeis University in 1971–1972. "Since my return from the U.S. in 1972," he reports, "I have lectured at several universities in Europe, particularly in Germany, Switzerland and England. Early this year I was awarded the Romano Guardini medal and prize by the Catholic Academy of Bavaria. As chairman of the Israel Interfaith Committee, I represent this committee on the International Jewish Committee for Interreligious Consultations, and I have become very actively involved in the Christian-Jewish dialogue on various levels."

© 1976 SCIENTIFIC AMERICAN, INC



Accutrac.
The turntable with eyes.

Introducing Accutrac.
The only turntable in the world that lets you tell an LP which selections you want to hear, the order you want to hear them in, even how many times you want to hear each one.

Sounds like something out of the 21st century, doesn't it? Well, as a result of Accutrac's electro-optics, computer programming and direct drive capabilities, you can have it today.

Just imagine you want to hear cuts 5, 3 and 7 in that order. Maybe you even want to hear cut 3 twice, because it's an old favorite. Simply press buttons 5, 3, 3 again, then 7. Accutrac's unique infra-red beam, located in the tonearm head, scans the record surface. Over the recorded portion the beam scatters but over the smooth surface between selections the infra-red light is reflected back to the tonearm, directing it to follow your instructions.

What's more, it can do this by cordless remote control, even from across the room.

The arm your fingers never have to touch.

Since Accutrac's tonearm is electronically directed to the record, you never risk dropping the tonearm accidently and scratching a record, or damaging a stylus.

And, since it cues electronically, too, you can interrupt your listening and then pick it up again in the same groove, within a fraction of a revolution. Even the best damped cue lever can't provide such accuracy. Or safety.

What you hear is as incredible as what you see.

Because the Accutrac servo-motor which drives the tonearm is decoupled the instant the stylus goes into play, both horizontal and vertical friction are virtually eliminated. That means you get the most accurate tracking possible and the most faithful reproduction.

You also get wow and flutter at a completely inaudible 0.03% WRMS. Rumble at −70 dB (DIN B). A tracking force of a mere 3/4 gram. And tonearm resonance at the ideal 8-10 Hz.

The Accutrac 4000 system. When you see and hear what it can do, you'll never be satisified owning anything else.

Its father was a turntable.
Its mother was a computer.

The Accutrac™ 4000

ADC

ADC Professional Products Group. A division of BSR (USA) Ltd. Route 303, Blauvelt, N.Y. 10913

© 1976 SCIENTIFIC AMERICAN, INC

If you need help with that difficult reaction, DMSO may be the answer. Send for a free CZ chemist.

As a reaction solvent, DMSO (Dimethyl Sulfoxide) is in a class by itself. It improves yields, increases rates, lowers required reaction temperatures and reduces side reactions (not necessarily all at the same time). DMSO has an extremely broad range of applications because of its unique chemical and physical characteristics, and has demonstrated utility in over 1,000 separate reactions that have been reported in the literature.

Properties of DMSO

A powerful aprotic solvent
High dielectric constant–48.9 @ 20°C
Very stable to bases
Good thermal stability
Very low order of toxicity

Will DMSO be of help to you? To answer this question, the Chemical Products Division of Crown Zellerbach Corporation has assembled and catalogued over 8,000 articles on DMSO–more than half of these pertain to the use of "DMSO As A Reaction Solvent."

We realize you do not have the time to spend the days in the library necessary to check 400, let alone 4,000, references in order to find the information specific to your area of interest. But now this information is available to you with essentially no effort.

The full DMSO library has been placed in a computer information retrieval program. And to you, the Research Chemist, we are offering the free services of our Information Scientist to set up a computer literature search that will answer the question, "Can DMSO be of help to you?"

As a guide to the tremendous volume of references in this program, our Information Scientists have developed a literature search manual on DMSO as a reaction solvent.

As the first step in finding out if "DMSO Can Be of Help To You," please attach the following coupon to your company letterhead and send it to us.

CrownZellerbach
Chemical Products Division
Camas 114, Washington 98607
(206) 834-4444

Crown Zellerbach, Chemical Products Division, Camas 114, Washington 98607

I have interest in DMSO as a reaction solvent. Please send me a copy of the literature search manual on "DMSO AS A REACTION SOLVENT."

Name_____ Title_____

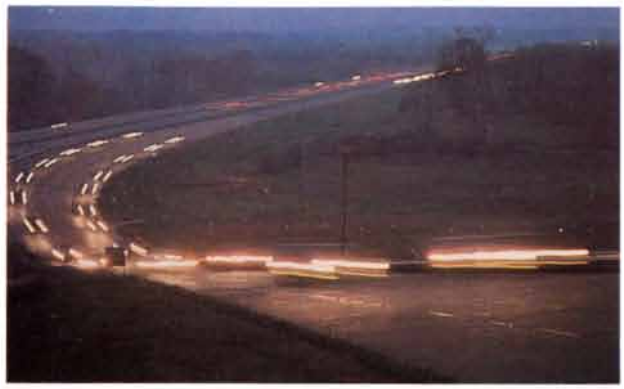
Company_____ Phone_____

Address_____

City_____ State_____ Zip_____

© 1976 SCIENTIFIC AMERICAN, INC

© 1976 SCIENTIFIC AMERICAN, INC



The great sulfate experiment.

Large-scale experimental measurements by the Research Laboratories at General Motors have helped answer an important question in atmospheric science.

Predictions by some scientists had suggested that sulfate emissions from catalyst-equipped cars might reach dangerous levels by 1985. An unlikely occurrence—only under rare atmospheric conditions and if most cars have catalytic converters—but still a possibility, they said.

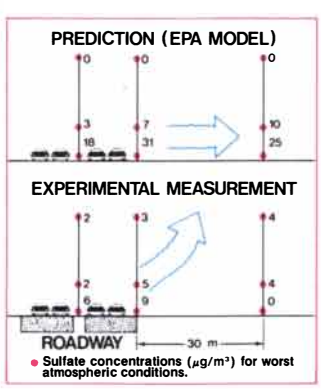
To get hard data, the Research Labs set up "The Great Sulfate Experiment." It was a public test—with the Environmental Protection Agency and other government and university researchers joining in.

Scientific evidence was gathered at the GM Proving Ground by simulating 1985 traffic. During 17 days, a test fleet of 352 catalyst-equipped cars rolled past our six testing towers, creating rush-hour traffic densities of 5,462 vehicles per hour.

During this time, sampling for sulfates and meteorological data took place at 20 points above and at ground level.

The finding: Roadway sulfate emissions are sent aloft by the turbulence and heat generated by vehicles.

It was a key discovery. Atmospheric scientists are now scrapping some old models of dispersion from roadways and drafting new ones based upon this unique data.



• Sulfate concentrations ($\mu g/m^3$) for worst atmospheric conditions.

What's more, today's catalytic converter system—the best solution we now know for changing some other exhaust pollutants into harmless gases—is still free to do this important job.

If you would like more information on the experiment and its findings, we welcome your inquiry.

GM

General Motors Research Laboratories
Warren, Michigan 48090

© 1976 SCIENTIFIC AMERICAN, INC

Established 1845 SCIENTIFIC AMERICAN January 1977 Volume 236 Number 1

Legal Abortion

The 10-year trend in many parts of the world toward the legalization of induced abortion has led to new studies of the practice, more open release of data, better medical procedures and lower mortality rates

by Christopher Tietze and Sarah Lewit

Eight years ago we reported in *Scientific American* that several nations had legalized induced abortion and that as a result the practice "is emerging from the shadows and has become a topic of worldwide discussion and controversy." It is now clear that what we were noting was the beginning of a widespread trend. Over the past 10 years in many countries abortion has changed from a largely disreputable practice into an accepted medical one, from a subject of gossip into an openly debated public issue. The issue is far from being settled. The debate is marked by honest and deeply felt differences of opinion, and public policy with regard to abortion is still subject to shifts in either direction. By and large, however, the decade's developments have enlarged public consciousness of a matter of immense human significance, have improved medical practice and have reduced the number of unnecessary deaths.

The status of abortion has been transformed dramatically by legislative action. Since 1966 many countries have changed formerly restrictive laws either to make abortion legal at the request of the pregnant woman or to specify a broad range of medical and socioeconomic indications for the legal termination of an unwanted pregnancy. Legal changes have in turn helped to modify professional attitudes. Partly in response to these changed attitudes physicians have developed new and safer procedures that make abortion in early pregnancy feasible and safe on an outpatient basis and that take the place of major surgery and reduce the length of hospitalization even later in pregnancy. The increasing experience with legal abortion and the growing awareness of the impact of illegal abortion on public health have encouraged health agencies to collect and publish official statistics that were not previously available. The need to set up guidelines for physicians and hospitals performing large numbers of abortions has stimulated studies on the possible complications of abortion. The results of the studies, widely disseminated in the open arena of public information, have in turn encouraged women who seek abortion to do so earlier in pregnancy and under safe medical conditions.

The fact remains that official and public interest in elective abortions in much of the world is more than offset by indifference in vast areas. The result is that worldwide information on abortion remains highly unsatisfactory. In many countries, particularly those with restrictive legislation, illegal abortions continue to be performed in large, untabulated and generally unacknowledged numbers. By aggregating individual-country data of varying accuracy one can estimate that between 30 and 55 million abortions are induced worldwide every year. The lower estimate implies an abortions-to-live-births ratio of about one to four, the higher estimate a ratio of almost one to two. Perhaps half of those abortions are performed legally; it is impossible to be sure because even some countries with very permissive laws, including the U.S.S.R. and China, either lack accurate figures or do not choose to make them public.

A review of the legislation over the past decade shows that more than a score of countries have made significant changes in the direction of less restriction and in many instances have chosen full legalization [*see illustration on next two pages*]. Four countries in eastern Europe that had very permissive abortion laws moved in the opposite direction. Other countries that had either nonrestrictive or only moderately restrictive laws made no changes in the course of the decade. In the remaining countries of the world, most of them in Latin America and in Africa, no significant changes were made in laws that either prohibit abortion without exception or allow it only to avert a serious threat to the life or health of a pregnant woman.

About a third of the world's people now live in countries with nonrestrictive abortion laws, that is, countries where pregnancies can be terminated on request at least up to a specified stage of pregnancy (from 10 weeks to 24 weeks, with 12 weeks the usual upper limit) and thereafter for medical indications; in a few countries the granting of permission for abortion on request may depend on the age of the woman and the size of her family.

Another third of the world's people are in countries with moderately restrictive abortion laws, where unwanted pregnancies may be terminated not on simple request but for broadly interpreted medical, psychological and socioeconomic reasons. Housing, income, marital status and other factors affecting a woman's life situation may be considered. Juridical indications generally include criminal acts such as rape and incest. Advances in medical techniques for the early diagnosis of fetal abnormalities have provided new and expanded indications for abortion. A provision of the British Abortion Act of 1967 (and of several other laws modeled on it) allows abortion if the continuation of the pregnancy would constitute a greater risk than its termination.

The remaining third of the world's people live in countries where abortion either is completely illegal or is allowed only if the woman's life or health is severely threatened by continuation of the pregnancy.

© 1976 SCIENTIFIC AMERICAN, INC

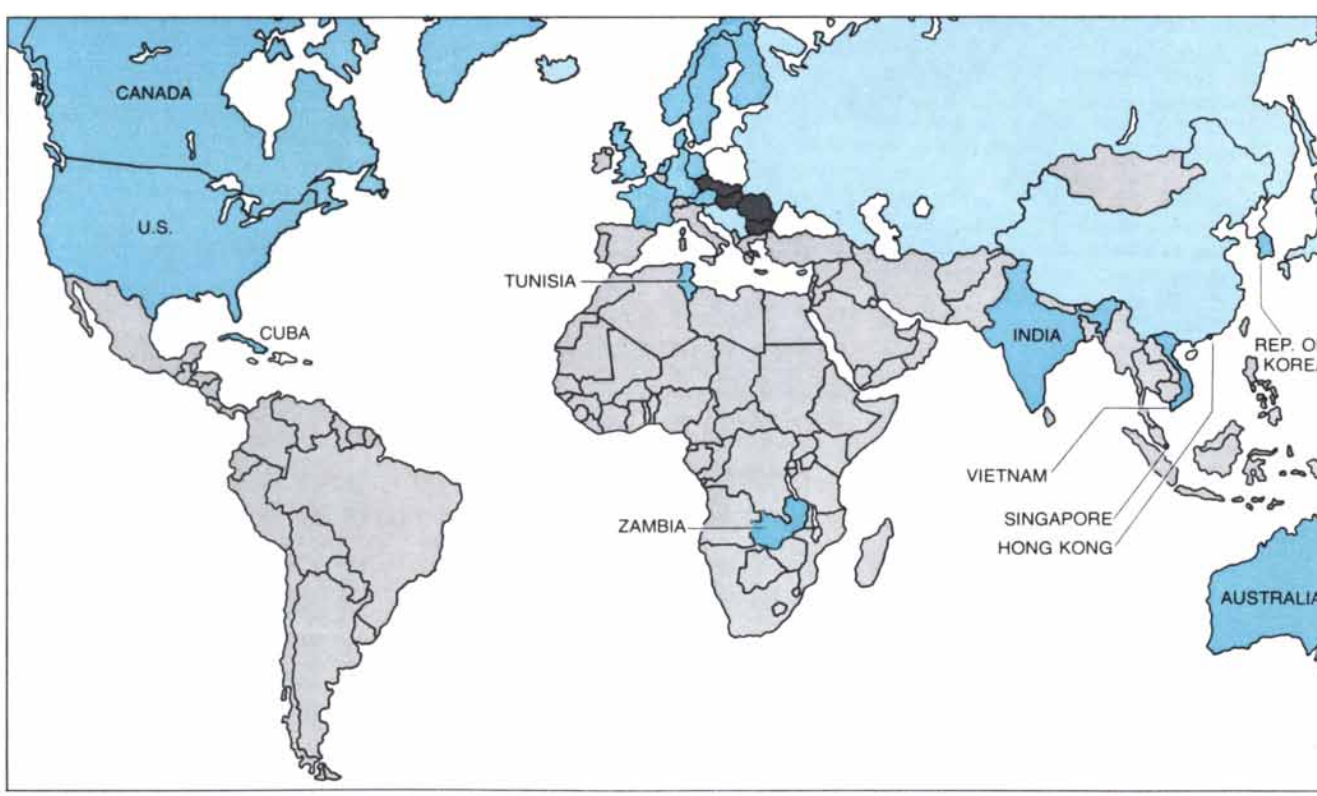
Practice does not always coincide with the law, to be sure. In some countries where abortions are forbidden they are nonetheless available to those who can pay for them. On the other hand, in countries with nonrestrictive laws conservative physicians, hospital administrators or officials may refuse to perform abortions or to allow them to be performed. The unavailability of adequate, convenient or reasonably priced facilities or services can also effectively deny women an abortion they seek.

In many cases women who are unable to have their pregnancies terminated in their own community have traveled to places with less restrictive laws or practices. The record of abortions in England after the enactment of the Abortion Act of 1967 is instructive in this regard [*see illustration on page 24*]. Significant numbers of women from America and the Continent came to London for abortions beginning in 1968, but the numbers from particular countries decreased sharply whenever abortion became more readily available in those or nearby countries. For example, the passage of New York State's nonrestrictive law in 1970 was reflected in an abrupt decline the following year in the number of U.S. and Canadian women who obtained abortions in England. The emergence of officially tolerated abortion clinics in the Netherlands reduced the number of patients from that country in 1972 and the number of patients from neighboring Belgium and West Germany a year later. The enactment of a nonrestrictive law in France reduced the number of Frenchwomen going to England for abortions by 60 percent in one year. On the other hand, increasing numbers of women were still arriving in 1975 from Ireland, Italy and Spain, where abortions continued to be proscribed.

In the U.S. changes in state abortion laws over the past decade have paralleled the general international trend. Beginning in 1967 about a dozen states enacted legislation based on the model penal code of the American Law Institute. The laws provided for abortion in cases where the pregnancy posed a "substantial risk" to a woman's mental or physical health, or if the child would be born with a "grave physical or mental defect," or if the pregnancy resulted from rape or incest. In 1970 four states went further: they passed laws allowing abortions without specifying any reasons, thus in effect authorizing abortions on request.

Two Supreme Court decisions in 1973 invalidated the restrictive abortion laws that were still in effect in most states by ruling, largely on grounds of privacy, that abortion could not be denied to a woman in the first three months (first trimester) of her pregnancy. After three months, the Court said, a state might "regulate the abortion procedure in ways that are reasonably related to maternal health." After the fetus reaches viability (the capability of surviving, given the appropriate life-support facilities), which is ordinarily at about 24 weeks, the state might "proscribe abortion except where necessary... for the preservation of the life or health" of the woman.

Those Supreme Court rulings have never been fully implemented. A vigorous "right to life" movement has successfully exerted pressure on some state legislatures to enact new restrictive laws and on many hospital administrators to prevent the delivery of abortion services in defiance of the decisions. A number of states moved to require parental consent in the case of minors or the husband's consent in the case of married women. In at least one state the preferred method for second-trimester abortions was outlawed. No abortions are performed in Catholic hospitals, and large numbers of non-Catholic hospitals are reluctant to allow abortions or even to set up the necessary facilities. Many of the restrictive state laws have already been challenged in court and found unconstitutional. Attempts to amend the Federal Constitution, either to outlaw abortion or to give states the right to do so, have received some support but have not yet got beyond the Congressional-committee stage. Last fall Congress did, however, attach to an appropriation bill a rider forbidding Federal Government



ABORTION LAWS AND PRACTICES have become significantly less restrictive during the past decade in more than a score of countries (*dark color*). In other countries with nonrestrictive or moderately restrictive laws there was no significant change (*light color*). Four

© 1976 SCIENTIFIC AMERICAN, INC

financial support for abortions conducted under Medicaid, a move that immediately came under attack as an unconstitutional denial of equal rights to poor people.

The relaxation of legal barriers to abortion and the changes in public attitudes and practices have tended to overcome the reluctance of some countries to release information on abortion. At this point national statistics are available on perhaps three million legal abortions a year, with major variations from country to country in completeness of coverage and in amount of detail.

We have worked out the legal-abortion rate (per 1,000 women from 15 to 44 years old) for 12 countries that have nonrestrictive or moderately restrictive laws and for which data are available covering all or most of the period from 1966 to 1975. In most of these countries the abortion rate increased over the decade. One reason for such increases is presumably a shift—the extent of which of course cannot be documented—from illegal and unreported to legal, recorded procedures. In addition some women who would not seek an illegal abortion feel free to choose a legal procedure when it becomes available. The abortion-rate curves trace fluctuations in national abortion policies, but roughly similar policies can have different effects [*see illustration on page 25*]. In Singapore, for example, the replacement of a moderately restrictive law by a nonrestrictive one in 1974 was followed by a 65 percent increase in the abortion rate. In Hungary the passage of restrictive provisions in 1973 sharply accelerated a decline from the 1969 peak; in Bulgaria and Czechoslovakia similarly restrictive changes brought about an initial decline that was followed by a slight rise. Where a plateau in the abortion rate is reached in a country with nonrestrictive laws, as in England and Wales and in Finland, it probably reflects the increasing and more effective practice of contraception. A very low abortion rate such as India's 1.4 in 1975 can be attributed in part to underreporting of even legal abortions, but it is also true that a shortage of physicians and facilities has limited the number of legal abortions in India. In some countries with particularly high rates women apparently rely more on abortion to prevent births, whereas in other countries contraception is more widely and effectively practiced, limiting pregnancies and therefore abortions as well.

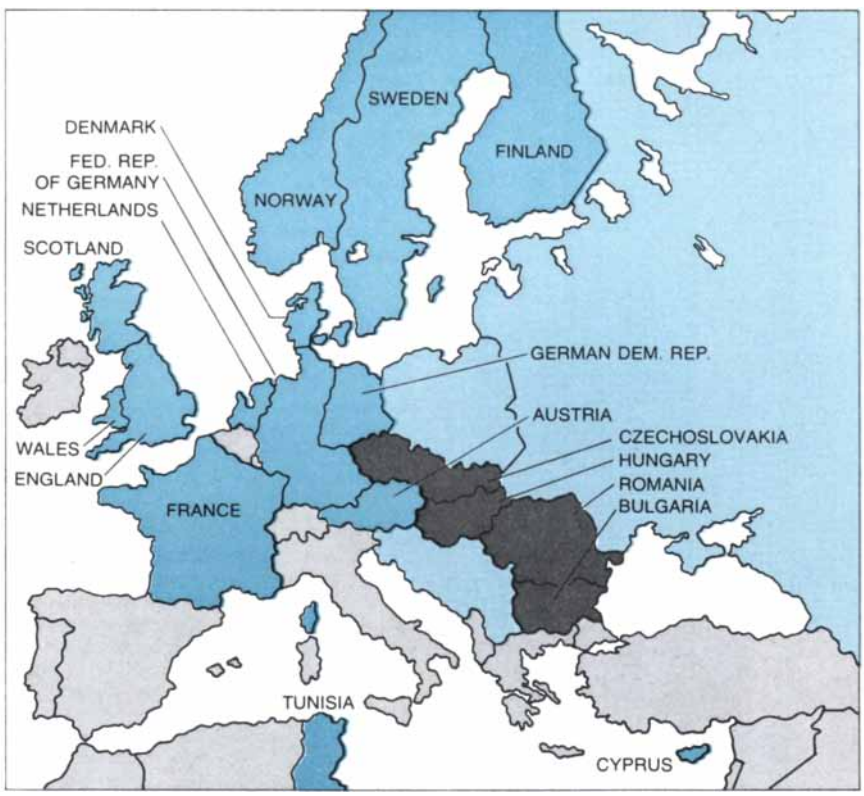
In five countries for which the age distribution of women obtaining legal abortions both early and late in the decade could be compared, more women whose pregnancies were terminated were in their twenties than were in their teens or over 30. In the U.S. teenagers now account for about a third of the abortions, in England and Wales and in Sweden for about a fourth. In Czechoslovakia and Hungary the proportion of very young women is much smaller; analysis of age-specific rates makes it clear that this pattern reflects high rates among adult women in the two countries rather than high rates among teenagers in the other countries. In all five countries more abortions were obtained by women under 20 late in the decade than early in it. That may reflect earlier physical maturation, later marriage (which increases the probability of premarital pregnancies), a widening gap between sexual and social adulthood, growing acceptance of abortion as an alternative to forced marriage or out-of-wedlock birth, liberalization of abortion laws and changes in the attitudes and practices of the medical profession.

There is little reliable information on the number of women who have had more than one induced abortion during their lifetime. Apparently about 60 percent of the Hungarian women who obtained elective abortions in 1973 had had at least one earlier abortion. In New York City the comparable figure in 1974 was about 22 percent. Such proportions will inevitably increase with time, even where contraception is practiced conscientiously, unless sterilization is widely offered and accepted by women (or men) who do not want more children.

The most important contribution of the new developments in abortion is the growing realization that early abortions are safer and less traumatic than late ones. First-trimester abortion is now generally accomplished by evacuating the uterus by suction rather than by curettage (scraping). Later in pregnancy the traditional surgical procedures have been largely replaced by the infusion into the uterus of a salt or hormone solution, which results in expulsion of the fetus. In three countries for which reliable figures are available on the distribution of abortions by period of gestation, the proportion of first-trimester abortions has increased as women and physicians recognize their relative safety and as legal-abortion services make them more readily available. (In Czechoslovakia and Hungary there have always been very few second-trimester legal abortions, which are authorized only for medical reasons.)

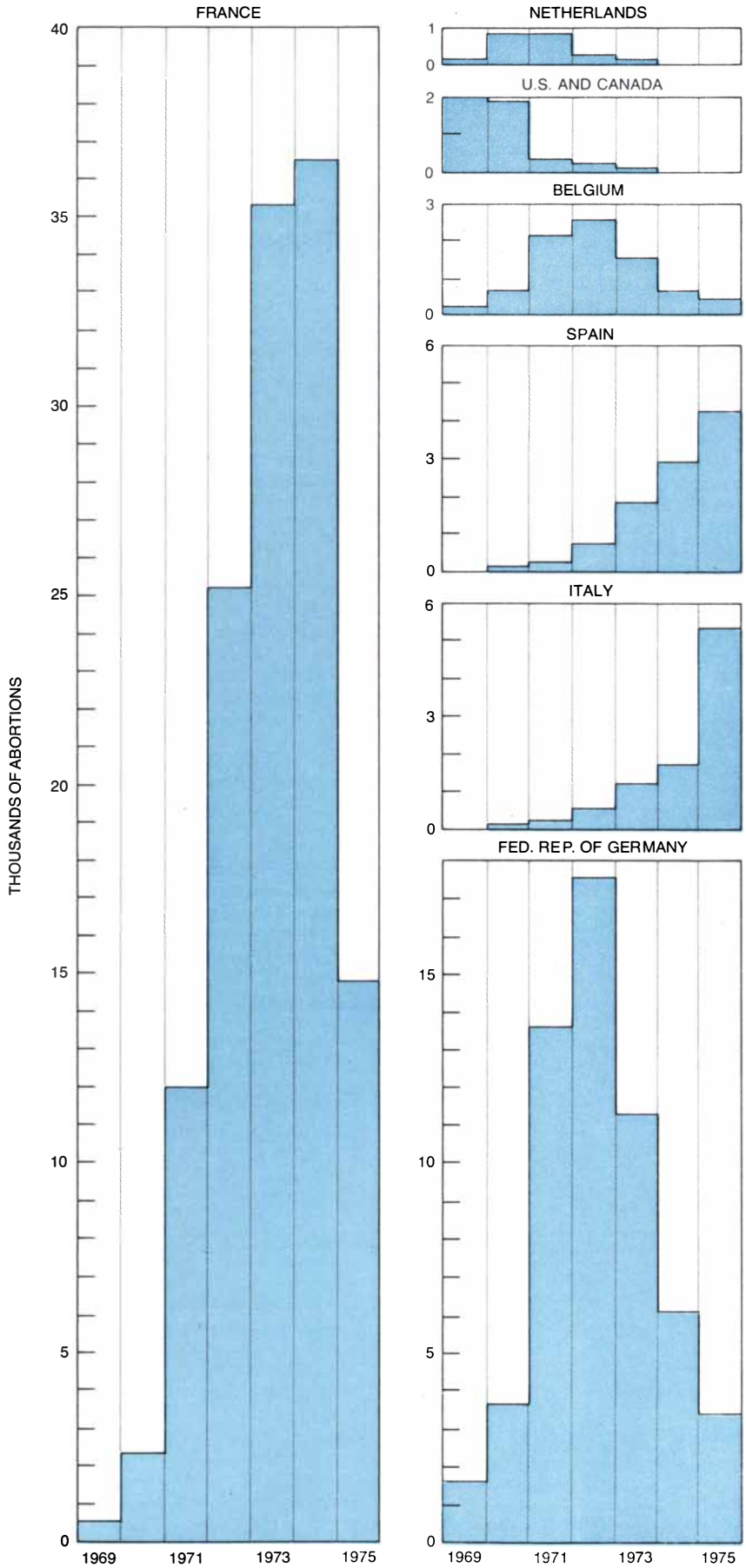
In spite of the well-established greater safety of early abortions some women continue to obtain abortions when their pregnancy is advanced and the procedure is riskier. A study in England and Wales showed that occupational level played a role: women who were (or whose husbands were) unskilled laborers tended to have abortions later in pregnancy than women who had (or whose husbands had) skilled, white-collar or professional occupations. In the U.S. black women tend to have more late abortions than white women.

Very young women also tend to be



countries made significant changes from less to more restrictive regulations (*dark gray*). In most countries the abortion situation remains generally restrictive or is unknown (*light gray*).

© 1976 SCIENTIFIC AMERICAN, INC



aborted late in pregnancy. In New York City the highest proportion of second-trimester abortions was reported for women under 15 and the lowest proportion was for women 35 or older [*see top illustration on page 27*]. The failure of very young women to seek abortion early may be attributable to an inability to recognize pregnancy symptoms, a belief that "it couldn't happen to me" or wishful thinking that "it'll go away," ignorance of where to seek help, reluctance to confide in parents or lack of money; cumbersome administrative procedures may also delay the necessary authorization while the pregnancy advances.

Some women are necessarily aborted only late in pregnancy: those who undergo diagnostic tests and learn that they are carrying a fetus with a serious physical or mental defect. Some 60 handicapping conditions in the fetus can now be recognized during pregnancy. The major diagnostic method is amniocentesis, the withdrawal for analysis of some of the amniotic fluid in which the fetus is immersed. The fetus can also be examined by ultrasound waves or, still as an experimental procedure, by fetoscopy (direct visual inspection) or by sampling the fetal blood. The optimum time for such prenatal tests is about the 16th week of gestation, and additional time may be required for tissue culture, bringing the fetus close to the period of viability. The number of abortions in this category is still very small (fewer than 200 a year in the U.S.) because the diagnostic procedures can be done at only a few medical centers.

Induced abortion at any stage carries some risk of complications, which range from very minor complaints to an occasional fatality, depending on the circumstances and of course on the stage of gestation. Recent studies of abortions in hospitals and nonhospital clinics show that the complications fall into two major categories: early (during the procedure or within about a month after it) and late (more than a month after it). Early complications are primarily minor complaints such as headaches or other discomfort and slight bleeding. The rate of serious early complications is only about .5 percent for abortions in the first trimester and about 2 percent thereafter, according to a large study conducted in the U.S. in 1970 and 1971.

NUMBER OF NONRESIDENTS who obtained abortions in England after it became possible in 1968 tended to increase as women became aware of the option but then decreased sharply if and when abortions became more available in the women's own countries. The data reflect the easing of restrictions in New York State (and thus on abortions for women from the U.S. and Canada), in the Netherlands (a change that also affected women from Belgium and the Federal Republic of Germany) and in France.

© 1976 SCIENTIFIC AMERICAN, INC

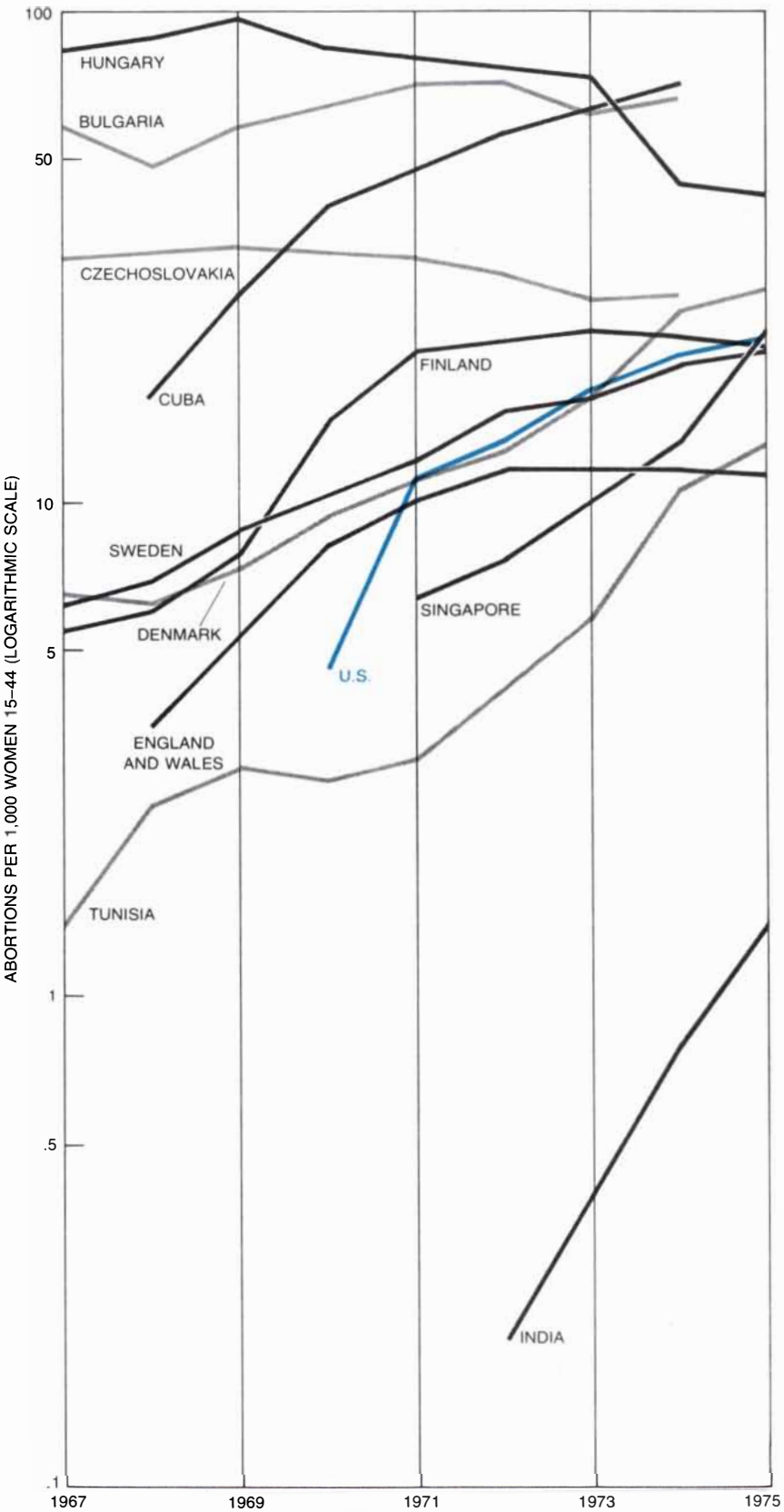
These complications include perforation of the uterus, major hemorrhage, laceration of the cervix (neck of the uterus), severe disturbances of blood coagulation and hypernatremia (excessive sodium in the blood) associated with saline-induced abortion, severe pelvic or generalized infection, venous thrombophlebitis and pulmonary embolism.

Identification of late complications is still far from complete. Rh-negative women can be sensitized by red blood cells from an aborted Rh-positive fetus, but this can be prevented by immunization with Rh globulin. Abortion by incision of the uterus (hysterotomy) can lead to endometriosis, the proliferation of uterine mucous tissue. The risk in subsequent pregnancies of spontaneous abortion in the second trimester, premature birth or early infant mortality has been studied, but with conflicting or inconclusive results. The value of retrospective studies on such effects of previous abortions is questionable, since women who subsequently have a normal birth are less likely to report they once had an abortion than are women whose pregnancies end unsatisfactorily.

The World Health Organization has sponsored prospective studies in seven European countries on the effects on a subsequent pregnancy and delivery of such variables as the duration of the previously terminated pregnancy, the technique (curettage or suction) and the interval between the abortion and the subsequent conception. The results will be available late this year. The role of induced abortion in subsequent cases of secondary sterility, ectopic pregnancy (development of the fetus outside the uterus), placenta previa (development of the placenta near the opening of the uterus) and painful menstruation has not been established; further studies are required in such areas and also on the psychological effects (if any) of induced abortion or, on the other hand, of giving birth to an unwanted child.

The mortality rate for all types of abortion (legal, illegal and spontaneous) has declined dramatically over the past decade even as the number of legal abortions has increased with the easing of restrictions. In the U.S., for example, there were on the average 5.7 abortion-related deaths per year per million women from 15 to 44 years old in the years from 1963 to 1967; there was only one death per million in 1974. A comparable decline from a comparable level was registered in England and Wales in the same period. In Romania, on the other hand, deaths due to abortion have increased almost sixfold since a nonrestrictive abortion law was replaced by the restrictive one of 1966.

The improvement in abortion-related mortality in many countries coincides with the trend toward earlier abortion

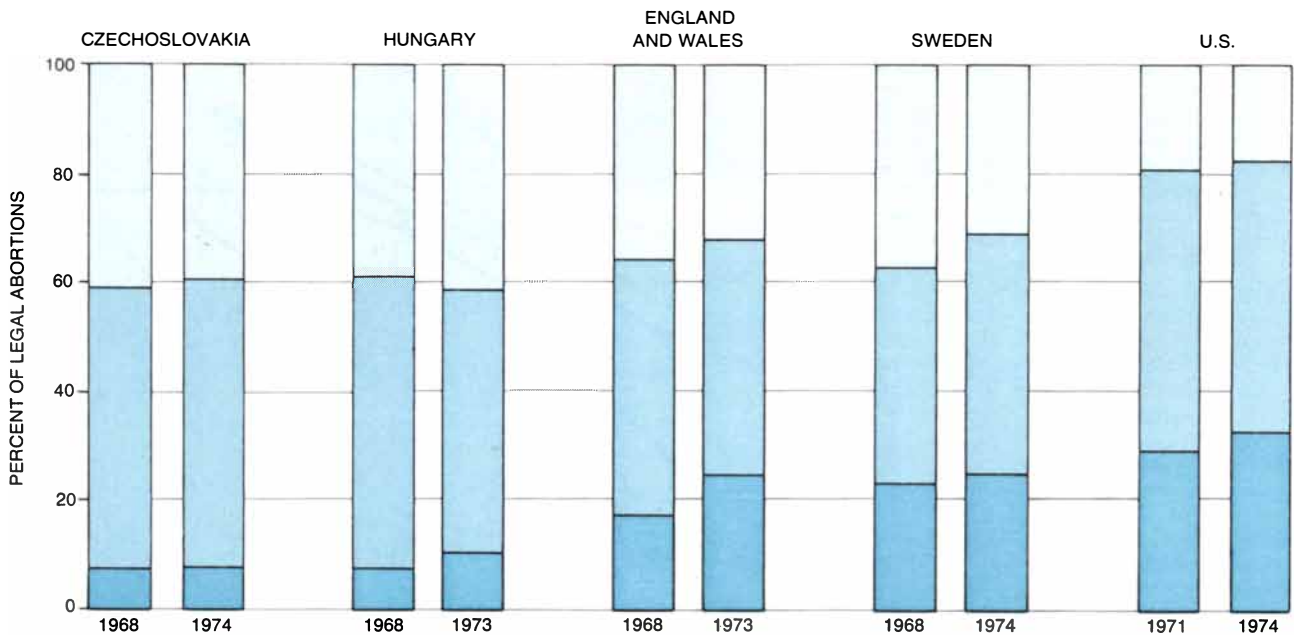


ABORTION RATES reflect changes in abortion laws and practices, the availability of services, the extent and effectiveness of contraception and the degree to which abortion is relied on to regulate fertility. Here the number of legal abortions per 1,000 women from 15 to 44 years old is plotted on a logarithmic scale, which means that parallel lines reflect equal rates of change. The data for England and Wales are for residents of that area only, not for visitors.

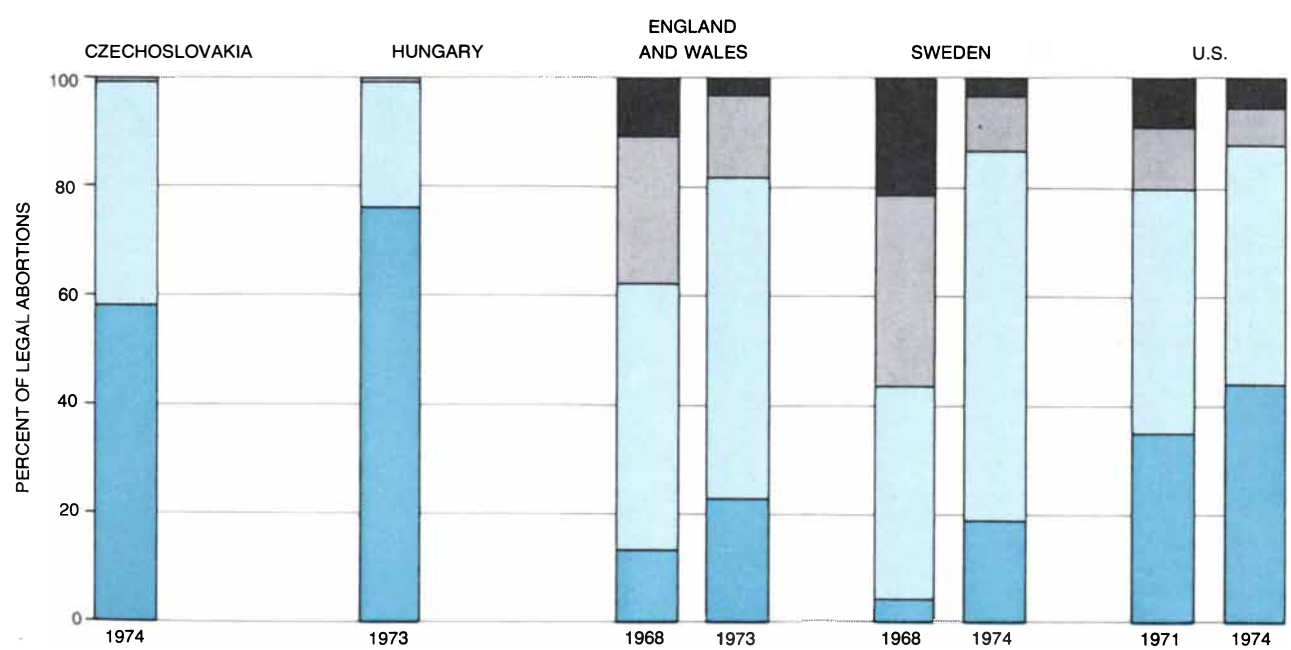
© 1976 SCIENTIFIC AMERICAN, INC

(when the procedure is simpler, safer and less traumatic) and toward abortion involving younger women (since, regardless of the stage of gestation, the risk is greater for older women). Another factor that presumably contributes to the decline in mortality is the increased sharing among physicians of knowledge they acquire in the course of their much-enlarged experience with the procedure. A particularly interesting comparison can be drawn between the risk of mortality associated with abortion and the risk related to pregnancy and childbirth. The death rate is far lower for legal abortions done during the first 12 weeks of pregnancy than it is for the complications of childbirth, according to figures from the U.S. and Britain [*see bottom illustration on opposite page*]. The death rates for abortion between the 13th and the 16th week and for childbirth are comparable, except in the case of older women, for whom childbirth remains a greater risk. Only abortions performed at 17 weeks or more of gestation show a generally higher mortality rate than childbirth.

In countries where contraception is practiced widely and effectively, as it is in the U.S. and Britain, abortion rates are substantially lower than they are in countries where there is little or no con-



WOMEN UNDER 20 obtain a larger proportion of the abortions in England and Wales, Sweden and the U.S. than they do in Hungary and Czechoslovakia. The bars show what percent of the legal abortions were performed in the indicated years on women who were under 20 (*dark color*), from 20 to 29 (*medium color*) and 30 or over (*light color*). The teenagers' share has increased in all these countries.



TREND TO EARLY ABORTION (in particular during the first eight weeks) is seen in England and Wales, Sweden and the U.S. Bars give the percent of abortions performed at eight weeks or less of gestation (*dark color*), nine to 12 weeks (*light color*), 13 to 16 weeks (*light gray*) and 17 weeks or more (*dark gray*). In Czechoslovakia and Hungary few legal abortions are performed after the first trimester.

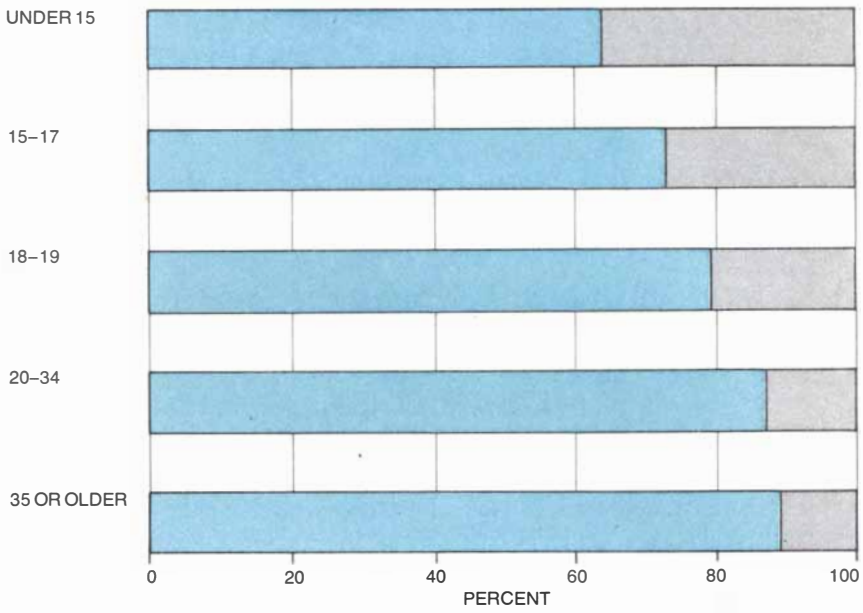
© 1976 SCIENTIFIC AMERICAN, INC

traception and where unplanned pregnancies are frequently terminated by abortion. Modern contraceptive methods such as the "pill" or the intrauterine device (IUD), although they are highly effective in preventing pregnancy, have some recognized side effects, and occasionally those effects can be fatal. It is therefore pertinent to investigate the relative risks of early abortion and of various contraceptive methods as means of fertility control.

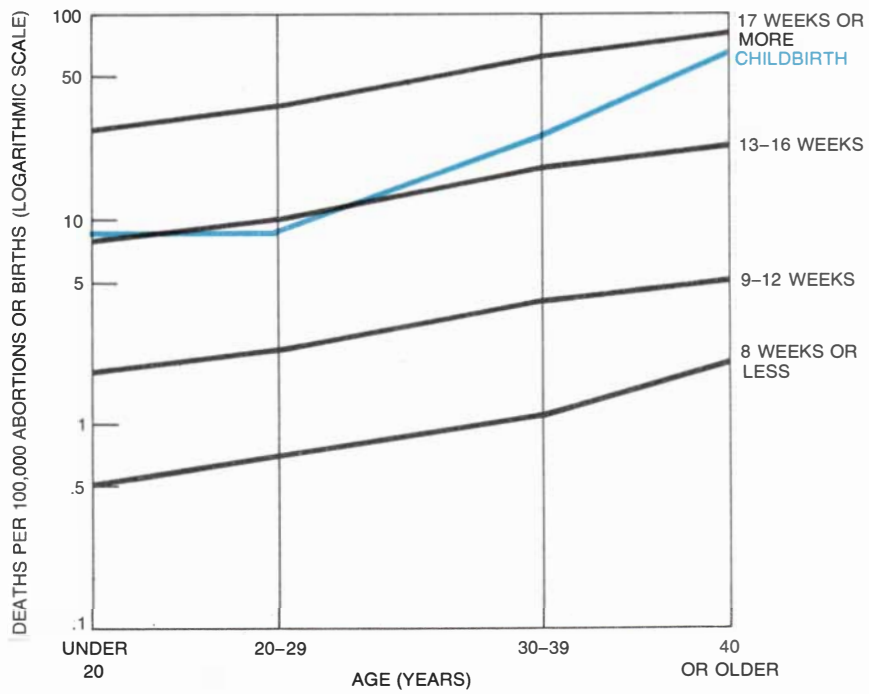
One of us (Tietze) recently collaborated in a study, based on a computer model, of the mortality that is associated in developed countries with each of four major contraceptive methods (pill, IUD, condom and diaphragm) and with first-trimester abortion. The total mortality was the sum of the mortality associated with the method of fertility control and that associated with any accidental pregnancies and eventual births. In the case of the pill most of the mortality was due to the contraceptive method; for the IUD it was due about equally to the contraceptive device and to accidental pregnancies; for the safe but less effective condom and diaphragm the deaths were all due to accidental pregnancies, and for early abortion the mortality was only that associated with the abortion procedure, since pregnancy was terminated.

The results showed that among women under 30 very low mortality rates (between one death and two deaths per 100,000 women per year) were associated with each of the contraceptive methods and with early abortion when any one of these fertility controls was used alone. The comparable risk from pregnancy and childbirth among young women who did not practice contraception or resort to abortion was estimated at six deaths per 100,000. The study showed, moreover, that a 100 percent level of fertility control and the lowest mortality rate (less than .5 per 100,000) could be achieved at all ages by combining the traditional contraceptive methods with early abortion as a backup: by utilizing the safe but somewhat less effective condom or diaphragm and then terminating any accidental pregnancies by means of first-trimester abortion.

In spite of the tens of millions of abortions—the vast majority of them illegal—that are performed annually and in spite of their impact on public health and on fertility, little interest in abortion has been manifested at the international level. Legal abortion did not even appear on the agenda of the World Population Conference of 1974 or that of the World Conference of the International Women's Year in 1975; there is one reference in the plan of action issued by the latter conference to illegal abortion, which is simply listed as a risk contributing to high maternal-mortality rates. Only one semiofficial report, limited to the U.S., deals with the subject unequivocally. *Legalized Abortion and the Public Health,* prepared by the Institute of Medicine of the National Academy of Sciences and issued in 1975, calls for research on "long-term medical complications, particularly after multiple abortions, the effects of abortion and denied abortion on the mental health and social welfare of individuals and families and the factors of motivation, behavior and access associated with contraceptive use and the choice of abortion." The valuable results already achieved by a few broad studies of legalized abortion confirm the institute's judgment that more such research is needed.



LATE ABORTIONS are most frequent among the youngest women and least frequent among women 35 or over. The bars show what proportion of the procedures were done at 12 weeks of gestation or less (*color*) and at 13 weeks or more (*gray*) in New York City in 1974.



MORTALITY associated with legal abortion at various ages and at different periods of gestation is compared with mortality associated with childbirth. The data reflect combined U.S. and British experience; the vertical scale is logarithmic. The abortion death rate increases with the duration of pregnancy and to a lesser degree with the woman's age. For all age groups the death rate for first-trimester abortion is substantially lower than the rate for childbirth.

© 1976 SCIENTIFIC AMERICAN, INC

Agriculture without Tillage

Within a few years much of the cropland in the U.S. will be planted without a moldboard plow. In most conditions planting without tillage (but with herbicides) can save labor, energy, water and soil

by Glover B. Triplett, Jr., and David M. Van Doren, Jr.

For centuries the moldboard plow has been the basic tool of agriculture, breaking and turning the soil as the first step in the series of operations collectively known as tillage. On a large and growing area of farmland, however, the plow is being displaced by a system of farming that involves either no tillage or a greatly reduced amount of it. The seeds for a new crop are simply planted in soil that remains covered with the residue of the old crop. The control of weeds, which is a prime objective of tilling, is achieved mainly by the application of herbicides but partly by the fact that the old crop acts as a mulch, stifling the growth of unwanted plants. The main advantages of the new methods are that they reduce the labor cost of farming and virtually eliminate the erosion of the soil by wind and water.

Tillage as it was practiced 40 or 50 years ago entailed, for a farmer growing corn in the U.S. Middle West, some 10 trips over the field by harvesttime. Plowing came first and was followed immediately by dragging or harrowing operations that broke up clods and made a fine, firm seedbed. Just before the crop emerged the field was lightly tilled again to destroy weeds that were emerging simultaneously. While the crop was growing the field would be cultivated two to four times more, mainly to control weeds. Eventually the corn plants grew high enough to keep the weeds suppressed by cutting them off from sunlight, but if the field was particularly weedy, it might be tilled again with a small cultivator pulled by a horse or a mule, even when the corn was as much as shoulder high. Hand labor was often required to pull out or attack with hoes the weeds that survived the successive cultivations.

For cotton the procedure was much the same. The production of small grains such as wheat, oats and barley required somewhat less tilling, since the plants grow so close together that there is no room for cultivating machinery once the crop has established itself.

These systems of tillage were developed over a long period of time and until quite recently were thought to represent the best and most dependable method of crop husbandry. They prepared an excellent seedbed and were effective in controlling weeds. They produced a loose, aerated soil that would usually absorb rainwater well. Tilling buried the residue of old plants, burying with them certain disease organisms and insects that might attack the new crop. Finally, tilling enhanced the effectiveness of the fertilizer that was applied to the soil.

One naturally wonders, then, why anyone would even think of eliminating or reducing such firmly established practices. The reason in brief is that under the right conditions the elimination or reduction of tillage can make the farmer's task easier and more profitable. Notwithstanding the widespread reliance on the plow, systems involving the intensive cultivation of the soil do have certain disadvantages. The demand that tilling makes on power and labor is high, particularly during the planting season. Moreover, since the optimum planting period for most crops is rather short, the amount of land the workers on a farm can till at the appropriate time is limited. Tilling removes the residue of the previous crop from the surface of the soil, thereby increasing the susceptibility of the soil to erosion, which in turn reduces the agricultural potential of the land. Bare soil is also likely to lose more water by evaporation and runoff than soil covered by plant residue loses, so that the crop has less water for growth.

One of the first efforts to modify the established methods of tilling was the plow-plant system devised some 30 years ago by Ray L. Cook, who was then chairman of the soil-science department at Michigan State University. In this system the entire field was plowed, but further tilling was confined to the rows where the crop was planted. They were strips of soil six to 10 inches wide and spaced 40 inches apart, and they were tilled enough to provide for the satisfactory operation of the planting machine. The initial plowing left the land able to take up water readily, but since the soil between the rows was left undisturbed after the initial plowing, runoff and erosion were reduced.

The system gained only limited acceptance in the corn belt. The soil conditions necessary for employing the system successfully are more demanding than they are for conventional tillage, and each operation must be performed carefully for the system to function properly. It is an axiom of tilling that successive operations correct earlier mistakes, and the plow-plant system offered fewer opportunities for correction. Moreover, the plowing usually had to be done during the optimum planting time rather than earlier, and the amount of land that could be handled by a work force of a given size was therefore severely limited.

Edward H. Faulkner, who was a farmer and a county agricultural agent in Ohio, devised one of the first systems of tillage that omitted plowing. In his system the soil was worked with disks, which do not go as deep as a plow and do not turn the earth as much. The system also involved leaving a certain amount of crop residue on the soil to reduce erosion. Faulkner described his methods in a book, *Plowman's Folly,* that was published in 1943. At the time plowing was regarded as being virtually synonymous with agriculture, and the book therefore generated a good deal of controversy.

A variation of Faulkner's system was "mulch tillage." In this system the soil was tilled in some way before planting, but then a mulch of manure was applied, either at the time of planting or soon afterward. The mulch helped to protect the soil and reduce erosion.

A once-over system called till planting was developed in the early 1950's. A special machine worked the soil in the rows where the crop would be planted, and it planted the seeds at the same time. The area between rows was worked over with sweeps, which cut into the soil enough to kill weeds.

Changes in tillage practices are more often the result of acceptance by farmers than of efforts by research workers

© 1976 SCIENTIFIC AMERICAN, INC

and farm-equipment salesmen. When a new concept or machine is introduced, a few farmers will evaluate its performance. If yields do not suffer and the innovation solves problems or reduces costs, it will be accepted and other farmers will adopt it.

By this criterion none of the early tillage systems that eliminated conventional plowing fared very well. In the development phase the different systems performed satisfactorily enough to offer encouragement about reducing erosion and other problems associated with plowing, but farmers who tried them had trouble controlling weeds, particularly perennial weeds, and their crop yields diminished.

The solution of the weed problem came from another development, the discovery of selective herbicides, that took place during the time when tractors were replacing draft animals and nitrogenous fertilizers became inexpensive enough to be widely available. The first of the selective herbicides was 2,4-D; it was introduced to farmers in the late 1940's and proved to be quite effective for the control of broadleaf weeds in corn. Although tillage effectively controls the weeds growing between rows of corn, the ones growing in the row cannot be reached by tilling machines. Herbicides proved useful in controlling those weeds, thereby reducing the need for hand labor.

Most selective herbicides are organic compounds that achieve their objectives at low rates of application and with low toxicity to human beings and animals. In the three decades since the introduction of 2,4-D well over 100 selective herbicides have been developed for application with different crops. Some of the herbicides are effective when they are applied to the foliage of growing weeds; others are active in the soil, inhibiting the germination of weed seeds or killing the seedlings after they sprout.

The ultimate herbicide, meaning one that controls all vegetation except the planted crop, remains to be developed. Many herbicides are most effective on groups of weeds, such as annual grasses or broadleaf plants. It is therefore a common practice to apply combinations of at least two herbicides to broaden the spectrum of weeds that can be controlled. For the control of weeds in corn alone nearly 20 different herbicides have been approved by the Environmental Protection Agency.

Herbicides work in various ways. Some of them inhibit the growth of the weed plants, others interfere with pho-

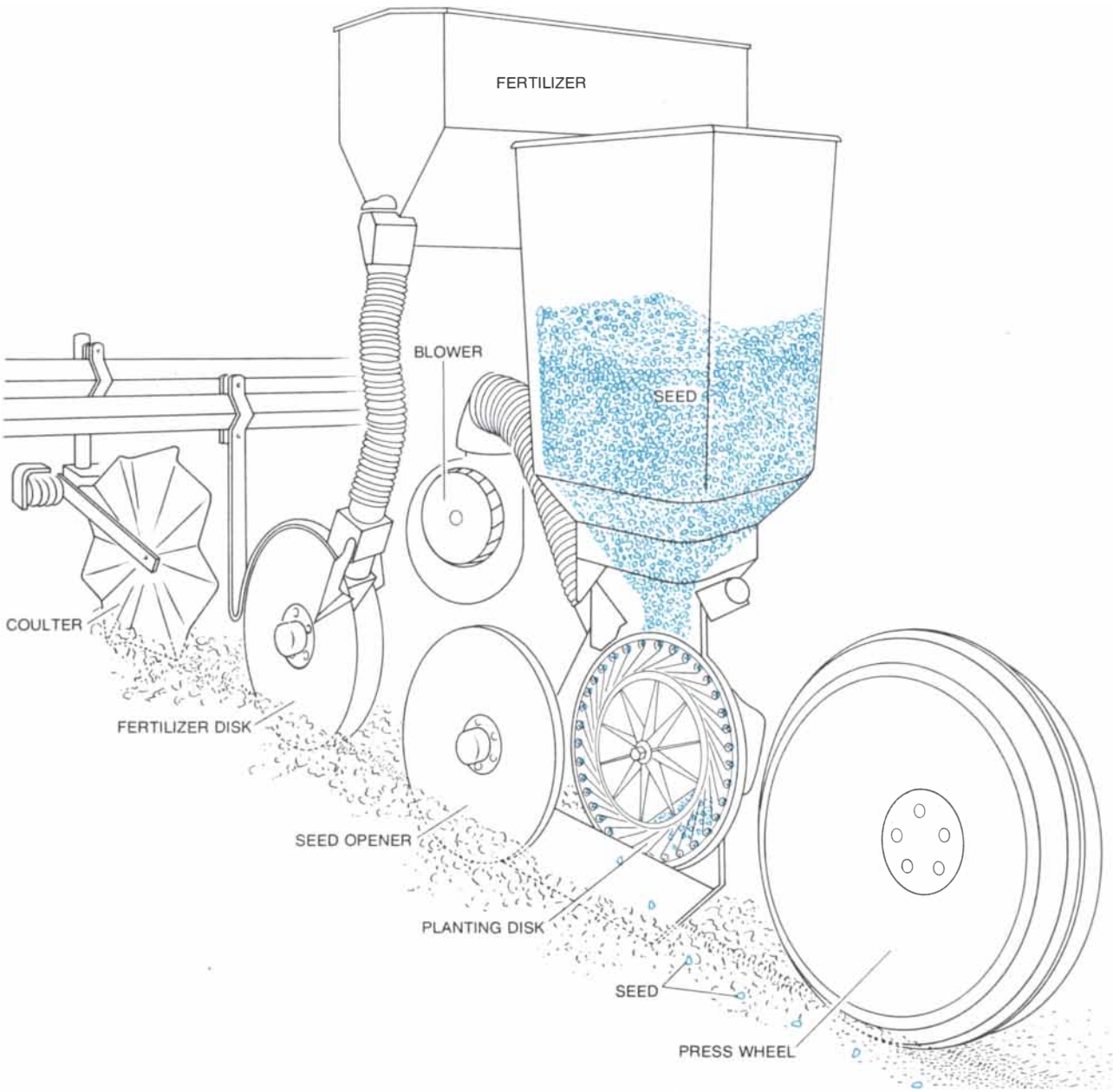
MULTIPLE CROPPING is made possible on a farm in Ohio by planting without tillage. The machine at the right is harvesting a crop of wheat as the machine at the left plants soybean seed in the wheat stubble. The two larger white containers on the planting machine carry dry fertilizer. Ahead of them but not visible in the photograph are implements that cut a furrow in the soil for the soybean seed, which is dropped into the furrow from the four smaller white containers. The four wheels at the rear close the furrow over the seed.

© 1976 SCIENTIFIC AMERICAN, INC

tosynthesis or with enzyme systems and still others desiccate the foliage of the weeds. It is of course crucial that the herbicide spare the cash crop. This objective can be achieved through the ability of the crop to metabolize the herbicide, through applying the herbicide to the weeds but not to the crop or through positioning the seeds of the crop properly in relation to the weed seeds and to the herbicide.

At first herbicides were applied as a supplement to established methods of tilling, which continued largely unchanged. With the passage of time and the introduction of more effective herbicides farmers began to eliminate some of the tilling they had done before planting and before the crop emerged. In addition they reduced the number of cultivations they did while the crop was growing. Herbicides also controlled several of the troublesome perennial weeds, further reducing the need for multiple tilling operations. With effective herbicides mulch tillage, till planting and Faulkner's technique all became workable systems that were adopted widely by farmers and are being employed now as the trend toward less tillage continues.

The ultimate extension of such a trend would be a system that involves no tillage. Some of the first efforts in this direction were programs for improving pastures. The land put to pasture is often unsuitable for the production of row crops because it is steep or stony or has



PLANTING MACHINE employed in a no-tillage system performs several operations at once. The direction of movement of the machine in this drawing is to the left. The first wheel is a fluted coulter that opens a narrow band in the untilled soil, usually cutting also through the residue of the preceding crop, which is left on the ground as a mulch. The coulter is followed by a disk that applies fertilizer. Here the fertilizer is dry, but in some machines it is applied as a liquid. (Herbicide can also be applied at about the time of planting to control weeds.) Next comes a disk, offset from the fertilizer disk by about two inches, that opens the furrow in which the seed is to be planted. It is followed by the planting unit, which receives the seed from the container above it. The seed is forced into the slots in the planting disk by air from a blower; the air also holds each seed in place until the slot approaches the ground and the seed drops into the furrow. The last wheel presses the soil down over the seed. Several units of this kind are normally ganged together in one machine.

© 1976 SCIENTIFIC AMERICAN, INC

other characteristics that make tilling difficult or impossible. Herbicides were applied to pastures to destroy unproductive vegetation so that desirable forage plants could be established.

To our knowledge the first demonstration that a row crop can be grown without tillage was achieved in Michigan some 20 years ago by representatives of a company that manufactures chemicals. A field on which sod had been growing was sprayed with herbicides to kill the sod. The seeds for a corn crop were then planted through the killed sod. At that time, however, herbicides had to be applied repeatedly in order to achieve good control of weeds. Only about 10 years later, after more effective herbicides had become available, was the practice of growing row crops without tillage adopted to any extent by farmers.

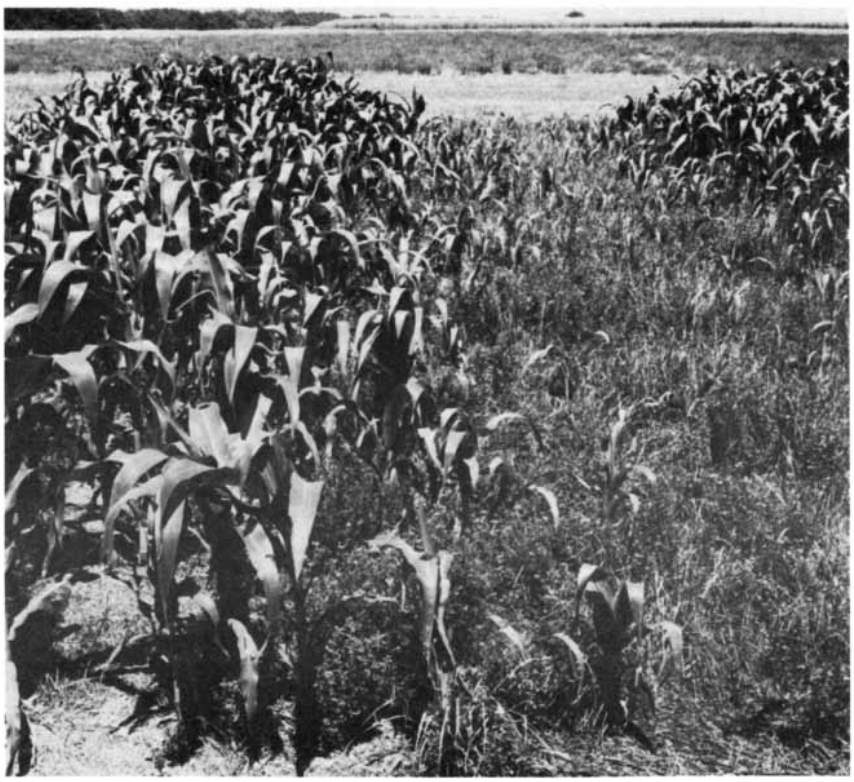
The spray-plant-harvest systems now employed go by such names as no tillage, zero tillage, sod planting, direct drilling and slot planting. For a row crop such as corn the farmer first sprays herbicides on the field to kill any growing vegetation and to prevent the growth of weeds from ungerminated seeds in the soil. The next step is to supply fertilizer, which can be put on in dry form but often is applied as a solution that also carries the herbicide. Seeds are planted by a machine that cuts through the plant residue on the surface of the soil, positions the seeds appropriately in the soil and covers them, all in one operation. The soil is undisturbed except for a band two or three inches wide made by the planter. Often no other operations are needed before harvesttime.

Corn is the row crop most widely grown without tillage. Other crops that have been produced successfully in this way are soybeans, cotton, peanuts, wheat, tobacco, grain sorghum, forage crops and certain vegetables.

The elimination of tillage means that the farmer must rely entirely on herbicides to control weeds. Since no single herbicide will do the job satisfactorily, careful thought must be given to the combinations of herbicides that will work best. They must be tailored to the type of vegetation present before planting and to the crop being grown.

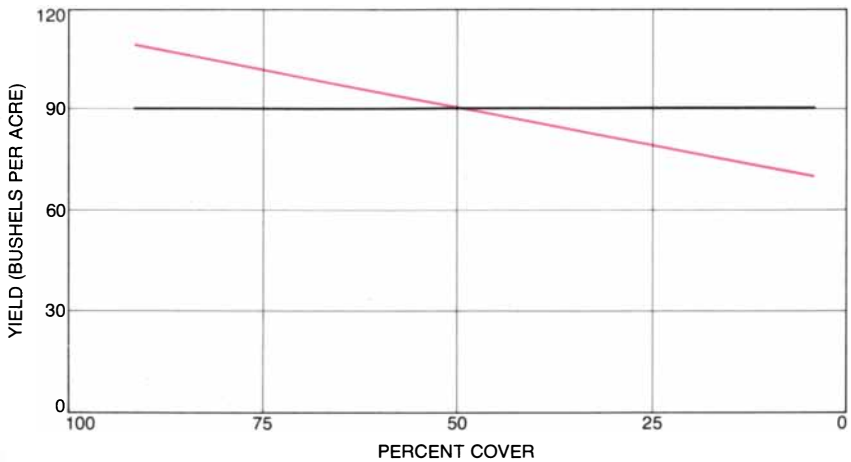
One of the effects of conventional tillage is to mix well into the soil the fertilizer that has been put on the surface of the field. In a no-tillage system such a mixing process is absent. This change has not been a problem as far as the three major nutrients (nitrogen, phosphorus and potassium) commonly applied for crops are concerned. Nitrogen is soluble and therefore mobile; it moves into the soil when rain falls or when the field is irrigated. Phosphorus and potassium move quite slowly, but the cover of mulch that is a feature of every no-tillage system keeps the moisture near the surface of the soil, so that

CROP OF SOYBEANS grows through a mulch of cornstalks on an experimental plot at the Ohio Agricultural Research and Development Center. The cornstalks lie as they were left by the machine that picked the corn. In this plot different combinations of herbicides were applied to the ground at the left and the ground at the right. The combination that was applied at the right was considerably less successful, as is indicated by weeds growing among soybeans.



CROP OF CORN grows at the Ohio center on ground that was previously planted to forage crops. Herbicides were applied in different combinations to kill the sod and to prevent the emergence of new weeds. The combination applied in the center was the least successful one.

© 1976 SCIENTIFIC AMERICAN, INC



IMPORTANCE OF SOIL COVER, or residue left on the ground after the harvesting of the preceding crop, in no-tillage agriculture is shown in the results of experimental plantings of corn in the same Ohio soil (Wooster silt loam). The black line shows the yield under conventional tillage: 91 bushels per acre. The amount of cover has no effect on the yield. The colored line shows yield under no tillage: 110 bushels at 90 percent cover and 71 bushels at 5 percent.

the roots of the crop grow in the zone that contains the immobile nutrients.

With forage crops a further innovation has been to introduce the seeds of other crops into the sward of forage. For example, seeds of legumes are put into a sward of grass. Many desirable legumes do not persist in pastures or meadows and must be reintroduced periodically. The legume is desirable as a source of protein for animals and of nitrogen for the growing grass.

In many areas swards are composed of perennial plants that have a rather narrow range of optimum temperatures, so that they grow best either in the warmer parts of the year or in the cooler parts. Annuals that have different temperature requirements are introduced into these swards to extend the productive time of the land. For example, winter wheat (a cool-season annual) is drilled into swards of Bermuda grass (a warm-season perennial) in the southeastern U.S.

Similar steps have been taken to make better use of the growing season for annual crops. One approach is to establish a second crop either after the harvest of the first crop or before that crop is fully mature. A common multiple-crop arrangement in the U.S. is to plant soybeans after the harvest of a small-grain crop. No-tillage methods of planting have contributed significantly to the practice.

The small-grain crop is harvested several weeks after the optimum time for planting soybeans. If the soybean crop is to succeed, it must be planted quickly and must grow rapidly. The planting machines are equipped with coulters, or cutting wheels, that slice through the stubble of the small-grain crop, so that soybean seeds can be planted immediately after the grain has been harvested. Preemergence herbicides are applied to kill weeds that are in the stubble of the small grain and to prevent the establishment of weeds that emerge after the soybeans have been planted. The northern limits of double cropping in the U.S. have been extended for several hundred miles by no-tillage methods of planting.

We have mentioned the reduction of erosion as a major benefit of reducing or eliminating tillage. The reduction is significant with almost any combination of less tillage and a cover of mulch on the soil, but the most dramatic reductions in erosion have resulted from systems that eliminate tillage. In 1969 a field test was made at the North Appalachian Experimental Watershed Research Station of the Department of Agriculture near Coshocton, Ohio. Corn was planted under conditions of no tillage and conventional tillage on several watersheds ranging in size from one acre to several acres. Instruments measured the loss of water and soil from the fields during rainstorms. A five-inch rainfall that fell during July, a time when conventionally tilled fields are particularly vulnerable to erosion, caused losses of up to 20 tons of soil per acre from the conventionally tilled fields with slopes of 6 to 8 percent. The loss of soil from a watershed with a slope of 20 percent, where corn had been planted in a killed meadow without tillage, was less than 100 pounds per acre.

A slope of 20 percent is too steep for the production of row crops with conventional tillage because of the ravages of erosion. Such a slope is also near the upper limit for the safe operation of farm equipment. Since erosion can be reduced a hundredfold or more with no-tillage planting, the production of row crops on rolling terrain becomes practical. Although highly productive soils are found in many hilly areas, the practice has been to devote them to forage crops as a conservation measure. With no-tillage methods a higher proportion of this land can be planted to more profitable crops.

The runoff from fields where tilling is minimal carries a markedly diminished load of silt, so that the quality of streams is improved. Since herbicides are an essential adjunct of methods that reduce or eliminate tillage, we have monitored the runoff of herbicide chemicals. The transport of herbicides dissolved in runoff water is highest when rain falls within a few hours after the herbicide has been applied, but the total movement of herbicides is almost negligible. We have measured it at a few grams per acre for rains falling within two days after the application of the herbicide. Even then the rainfall must be heavy enough to cause a significant amount of runoff before any movement of herbicides is produced. On a watershed consisting of several thousand acres, planting and the application of herbicides are done in individual fields over a period of four to six weeks. Thus if there was a rainfall heavy enough to cause a significant movement of herbicides, it would affect only the few fields that were most susceptible to the runoff of herbicides.

Another advantage of the systems we have been describing is their saving of time and labor. Conventional methods of tillage require a considerable amount of power and labor for the preparation of the seedbed. Although in some situations part of the tilling can be done before the planting season, the practice is not acceptable in fields where erosion is a hazard. Adding to the problems faced by farmers who have only a short time to till are rainy periods that stop operations in the field. Rain and wet soil at planting time are hazards from Iowa eastward to the Atlantic coast.

Direct planting into untilled soil with suitable equipment is a rapid operation with a relatively low demand for power, so that the need for large tractors is reduced. Whenever the moisture of the soil is favorable for tilling of any kind, planting machines can be operated.

The other essential operation for no-tillage planting, the application of herbicides, can also be accomplished rapidly, often at the same time as planting. In general the adoption of no-tillage methods can be expected to increase the productivity of farm workers as much as threefold. The amount of fuel required to establish the crop can be reduced by as much as two-thirds of what would be consumed in conventional tillage.

After the requirements for weed control and crop establishment have been met, the most important factor in determining the response of a crop to

© 1976 SCIENTIFIC AMERICAN, INC

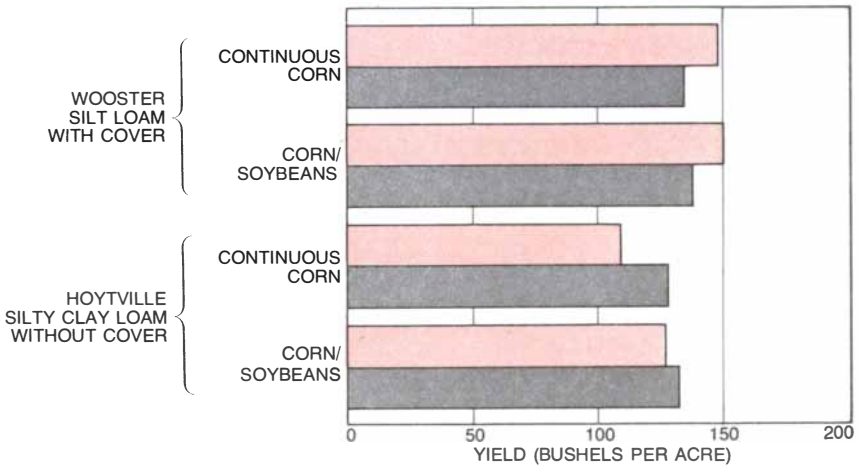
tillage is the drainage of the soil. The better the drainage is, the less tillage is necessary. On well-drained soil the no-tillage system is emerging as the best method of producing a crop, provided that at least 60 percent of the surface is covered with crop residue. This is fortunate, since the soils with the best drainage are often found in rolling terrain where erosion is a problem.

A cover of mulch from the previous crop conserves moisture not only by reducing evaporation but also by increasing the intake of rainfall, which tends to run off bare soil. The mulch also maintains soil's tilth, or good structure, which is preserved by not destroying the structure with tilling. Moreover, earthworms are often abundant under the mulch cover of the untilled field, and the less they are disturbed, the more they aid in the development of desirable soil structure. In these conditions corn responds favorably, yielding 10 to 20 percent more than corn grown on conventionally tilled soil. For a number of years we have measured yields on fields of well-drained soil that have been planted continuously to corn and other fields where corn is planted every third year in rotation with two other crops. The fields planted without tillage have exceeded the yield of the conventionally tilled fields by 18 bushels per acre per year where the corn was grown continuously and by seven bushels where corn was grown every third year in rotation.

No-tillage methods are not suited to every type of soil or soil condition. On poorly drained soils the mulch cover retains more moisture in the soil during the early spring than is desirable. The moisture in turn retards the warming of the soil and the rate at which the crop sprouts and grows. On the other hand, the lower temperature under a cover of mulch can be an advantage in tropical areas, where the temperature of a bare soil surface can reach 50 degrees Celsius (122 degrees Fahrenheit), which is high enough to damage seedlings.

We have found that in poorly drained soil the yield of corn grown continuously can be as much as 30 percent lower with no-tillage methods than it is when the soil is plowed. These differences take two or three years to develop, and they do not appear if the corn follows some other crop. We suspect that disease organisms build up in poorly drained soil if the soil is not plowed. Yields are equivalent for continuous corn with both conventional tillage and no tillage if the land is plowed the preceding fall or winter.

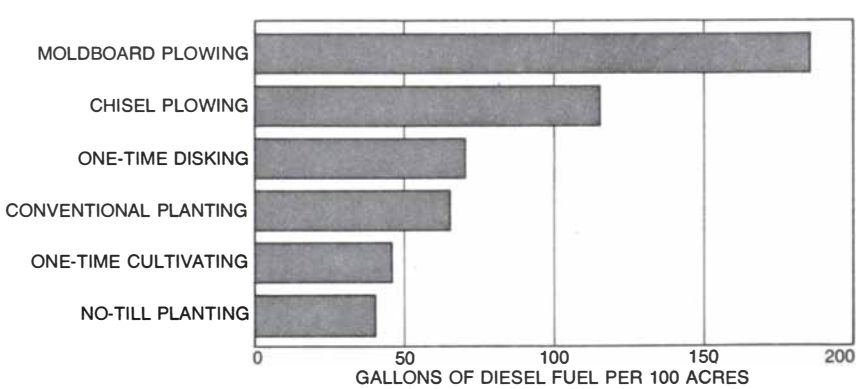
We do not want to leave the impression that reduced-tillage systems are without problems of their own. A farmer moving to such a system must take pains to acquire certain skills if he is to exploit the system successfully. The skills relate to the methods of planting, the use of chemicals and the choice of a method of tilling that is appropriate to the condition of the soil.



IMPORTANCE OF SOIL TYPE, together with cover conditions, in no-tillage agriculture is shown in the results of experimental plantings in two different Ohio soils: the well-drained Wooster silt loam and the less well-drained Hoytville silty clay loam. For both the continuous cultivation of corn and the alternating cultivation of corn and soybeans the yield on the Wooster silt loam was higher with no tillage (*colored bars*) than with conventional tillage (*gray bars*). For same plantings yield on Hoytville silty clay loam was higher with conventional tillage.

Pests, weeds and insects can be more of a problem with reduced-tillage systems than they are with conventional tillage. To the extent that the amount of tilling is reduced, the farmer's reliance on the proper employment of herbicides is increased. Tilling helps in the control of insects by disturbing their habitats, whereas a layer of mulch provides cover and a convenient place for various insect pests to deposit eggs. Armyworms, which are only an incidental problem in tilled fields, have destroyed entire fields of corn planted without tillage in killed sod. Slugs are more prevalent on fields covered with mulch. The population of field mice is often higher in fields with heavy cover, and mice have eaten corn seeds before the crop emerged. These various pests can be controlled successfully, but the farmer must take care to diagnose problems of this nature and deal with them before they become serious.

Taking the tillage systems of 40 or 50 years ago as a point of reference, one finds that a strong movement toward the adoption of reduced-tillage systems has already occurred. A survey made recently by the Soil Conservation Service indicated that in the U.S. in 1976 some 7.3 million acres were planted without tillage and that on 52.5 million additional acres tillage was reduced from the conventional level. The amount of land conventionally tilled was 218 million acres. In an assessment of minimum tillage published in 1975 the Department of Agriculture predicted that by the year 2010 more than 90 percent of the acreage of crops will be grown with reduced-tillage systems and that on more than half of the acreage some form of no-tillage farming will be the practice.



PREPARATIVE ENERGY COSTS of no-tillage agriculture are compared with those of conventional tillage in terms of the fuel required by a 100-horsepower diesel tractor. The energy costs of conventional tillage include not only the costs of plowing but also those of disking, planting and cultivating. The data for this chart and others on these two pages were supplied by the authors and their colleagues at Ohio Agricultural Research and Development Center.

© 1976 SCIENTIFIC AMERICAN, INC

The Quantum Mechanics of Black Holes

Black holes are often defined as areas from which nothing, not even light, can escape. There is good reason to believe, however, that particles can get out of them by "tunneling"

by S. W. Hawking

The first 30 years of this century saw the emergence of three theories that radically altered man's view of physics and of reality itself. Physicists are still trying to explore their implications and to fit them together. The three theories were the special theory of relativity (1905), the general theory of relativity (1915) and the theory of quantum mechanics (c. 1926). Albert Einstein was largely responsible for the first, was entirely responsible for the second and played a major role in the development of the third. Yet Einstein never accepted quantum mechanics because of its element of chance and uncertainty. His feelings were summed up in his often-quoted statement "God does not play dice." Most physicists, however, readily accepted both special relativity and quantum mechanics because they described effects that could be directly observed. General relativity, on the other hand, was largely ignored because it seemed too complicated mathematically, was not testable in the laboratory and was a purely classical theory that did not seem compatible with quantum mechanics. Thus general relativity remained in the doldrums for nearly 50 years.

The great extension of astronomical observations that began early in the 1960's brought about a revival of interest in the classical theory of general relativity because it seemed that many of the new phenomena that were being discovered, such as quasars, pulsars and compact X-ray sources, indicated the existence of very strong gravitational fields, fields that could be described only by general relativity. Quasars are starlike objects that must be many times brighter than entire galaxies if they are as distant as the reddening of their spectra indicates; pulsars are the rapidly blinking remnants of supernova explosions, believed to be ultradense neutron stars; compact X-ray sources, revealed by instruments aboard space vehicles, may also be neutron stars or may be hypothetical objects of still higher density, namely black holes.

One of the problems facing physicists who sought to apply general relativity to these newly discovered or hypothetical objects was to make it compatible with quantum mechanics. Within the past few years there have been developments that give rise to the hope that before too long we shall have a fully consistent quantum theory of gravity, one that will agree with general relativity for macroscopic objects and will, one hopes, be free of the mathematical infinities that have long bedeviled other quantum field theories. These developments have to do with certain recently discovered quantum effects associated with black holes, which provide a remarkable connection between black holes and the laws of thermodynamics.

Let me describe briefly how a black hole might be created. Imagine a star with a mass 10 times that of the sun. During most of its lifetime of about a billion years the star will generate heat at its center by converting hydrogen into helium. The energy released will create sufficient pressure to support the star against its own gravity, giving rise to an object with a radius about five times the radius of the sun. The escape velocity from the surface of such a star would be about 1,000 kilometers per second. That is to say, an object fired vertically upward from the surface of the star with a velocity of less than 1,000 kilometers per second would be dragged back by the gravitational field of the star and would return to the surface, whereas an object with a velocity greater than that would escape to infinity.

When the star had exhausted its nuclear fuel, there would be nothing to maintain the outward pressure, and the star would begin to collapse because of its own gravity. As the star shrank, the gravitational field at the surface would become stronger and the escape velocity would increase. By the time the radius had got down to 30 kilometers the escape velocity would have increased to 300,000 kilometers per second, the velocity of light. After that time any light emitted from the star would not be able to escape to infinity but would be dragged back by the gravitational field. According to the special theory of relativity nothing can travel faster than light, so that if light cannot escape, nothing else can either.

The result would be a black hole: a region of space-time from which it is not possible to escape to infinity. The boundary of the black hole is called the event horizon. It corresponds to a wave front of light from the star that just fails to escape to infinity but remains hovering at the Schwarzschild radius: $2GM/c^2$, where G is Newton's constant of gravity, M is the mass of the star and c is the velocity of light. For a star of about 10 solar masses the Schwarzschild radius is about 30 kilometers.

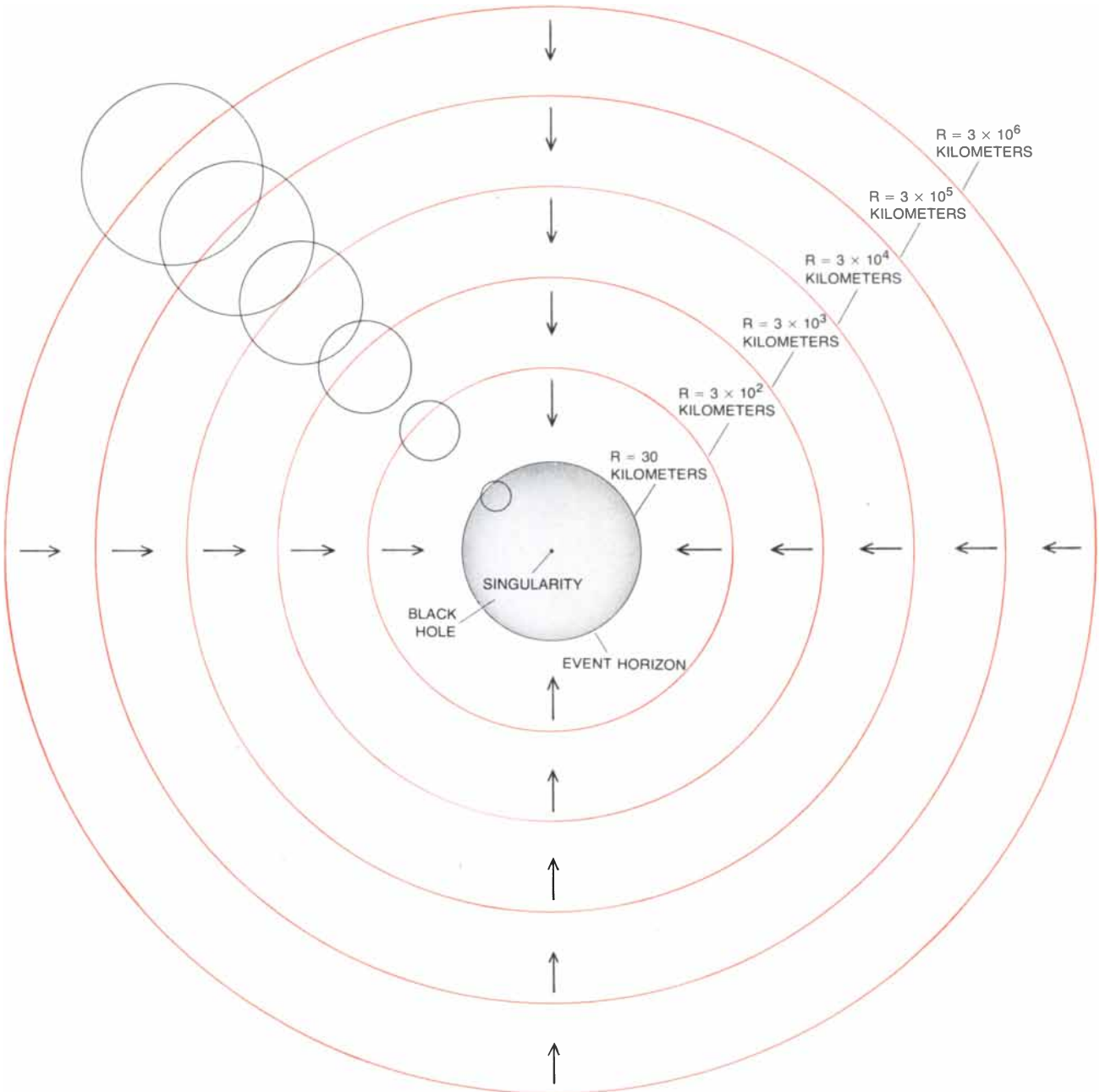
There is now fairly good observational evidence to suggest that black holes of about this size exist in double-star systems such as the X-ray source known as Cygnus X-1 [see "The Search for Black Holes," by Kip S. Thorne; SCIENTIFIC AMERICAN, December, 1974]. There might also be quite a number of very much smaller black holes scattered around the universe, formed not by the collapse of stars but by the collapse of highly compressed regions in the hot, dense medium that is believed to have existed shortly after the "big bang" in which the universe originated. Such "primordial" black holes are of greatest interest for the quantum effects I shall describe here. A black hole weighing a billion tons (about the mass of a mountain) would have a radius of about 10^{-13} centimeter (the size of a neutron or a

© 1976 SCIENTIFIC AMERICAN, INC

proton). It could be in orbit either around the sun or around the center of the galaxy.

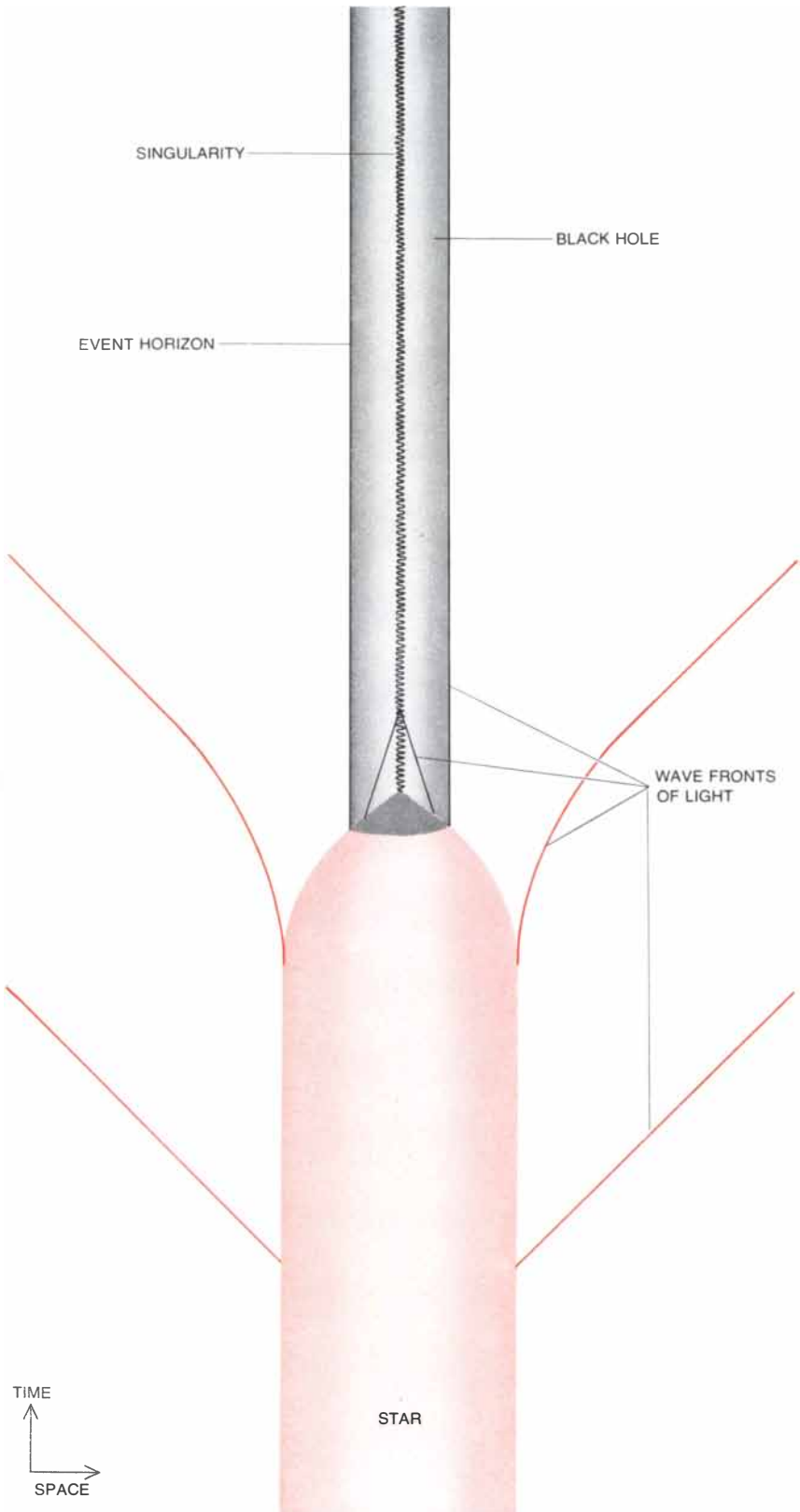
The first hint that there might be a connection between black holes and thermodynamics came with the mathematical discovery in 1970 that the surface area of the event horizon, the boundary of a black hole, has the property that it always increases when additional matter or radiation falls into the black hole. Moreover, if two black holes collide and merge to form a single black hole, the area of the event horizon around the resulting black hole is greater than the sum of the areas of the event horizons around the original black holes. These properties suggest that there is a resemblance between the area of the event horizon of a black hole and the concept of entropy in thermodynamics. Entropy can be regarded as a measure of the disorder of a system or, equivalently, as a lack of knowledge of its precise state. The famous second law of thermodynamics says that entropy always increases with time.

The analogy between the properties of black holes and the laws of thermodynamics has been extended by James M.



COLLAPSE OF STAR of 10 solar masses is depicted schematically from an original radius of three million kilometers (about five times the radius of the sun) to 30 kilometers, when it disappears within the "event horizon" that defines the outer limits of a black hole. The star continues to collapse to what is called a space-time singularity, about which the laws of physics are silent. The series of six small circles represents the wave fronts of light emitted from the successive surfaces an instant before the star had collapsed to the dimensions shown. Radii of the star and of the wave fronts are on a logarithmic scale. At each stage of collapse more of the wave front falls within the volume of the star as the escape velocity increases from 1,000 kilometers per second to 300,000 kilometers per second, the velocity of light. The final velocity is reached as the star disappears within the event horizon. No light emitted after that can ever reach outside observers.

© 1976 SCIENTIFIC AMERICAN, INC



GRAVITATIONAL COLLAPSE OF A STAR is depicted in a space-time diagram in which two of the three dimensions of space have been suppressed. The vertical dimension is time. When the radius of the star reaches a critical value, the Schwarzschild radius, the light emitted by the star can no longer escape but remains hovering at that radius, forming the event horizon, the boundary of the black hole. Inside black hole star continues collapse to a singularity.

Bardeen of the University of Washington, Brandon Carter, who is now at the Meudon Observatory, and me. The first law of thermodynamics says that a small change in the entropy of a system is accompanied by a proportional change in the energy of the system. The fact of proportionality is called the temperature of the system. Bardeen, Carter and I found a similar law relating the change in mass of a black hole to a change in the area of the event horizon. Here the factor of proportionality involves a quantity called the surface gravity, which is a measure of the strength of the gravitational field at the event horizon. If one accepts that the area of the event horizon is analogous to entropy, then it would seem that the surface gravity is analogous to temperature. The resemblance is strengthened by the fact that the surface gravity turns out to be the same at all points on the event horizon, just as the temperature is the same everywhere in a body at thermal equilibrium.

Although there is clearly a similarity between entropy and the area of the event horizon, it was not obvious to us how the area could be identified as the entropy of a black hole. What would be meant by the entropy of a black hole? The crucial suggestion was made in 1972 by Jacob D. Bekenstein, who was then a graduate student at Princeton University and is now at the University of the Negev in Israel. It goes like this. When a black hole is created by gravitational collapse, it rapidly settles down to a stationary state that is characterized by only three parameters: the mass, the angular momentum and the electric charge. Apart from these three properties the black hole preserves no other details of the object that collapsed. This conclusion, known as the theorem "A black hole has no hair," was proved by the combined work of Carter, Werner Israel of the University of Alberta, David C. Robinson of King's College, London, and me.

The no-hair theorem implies that a large amount of information is lost in a gravitational collapse. For example, the final black-hole state is independent of whether the body that collapsed was composed of matter or antimatter and whether it was spherical or highly irregular in shape. In other words, a black hole of a given mass, angular momentum and electric charge could have been formed by the collapse of any one of a large number of different configurations of matter. Indeed, if quantum effects are neglected, the number of configurations would be infinite, since the black hole could have been formed by the collapse of a cloud of an indefinitely large number of particles of indefinitely low mass.

The uncertainty principle of quantum

© 1976 SCIENTIFIC AMERICAN, INC

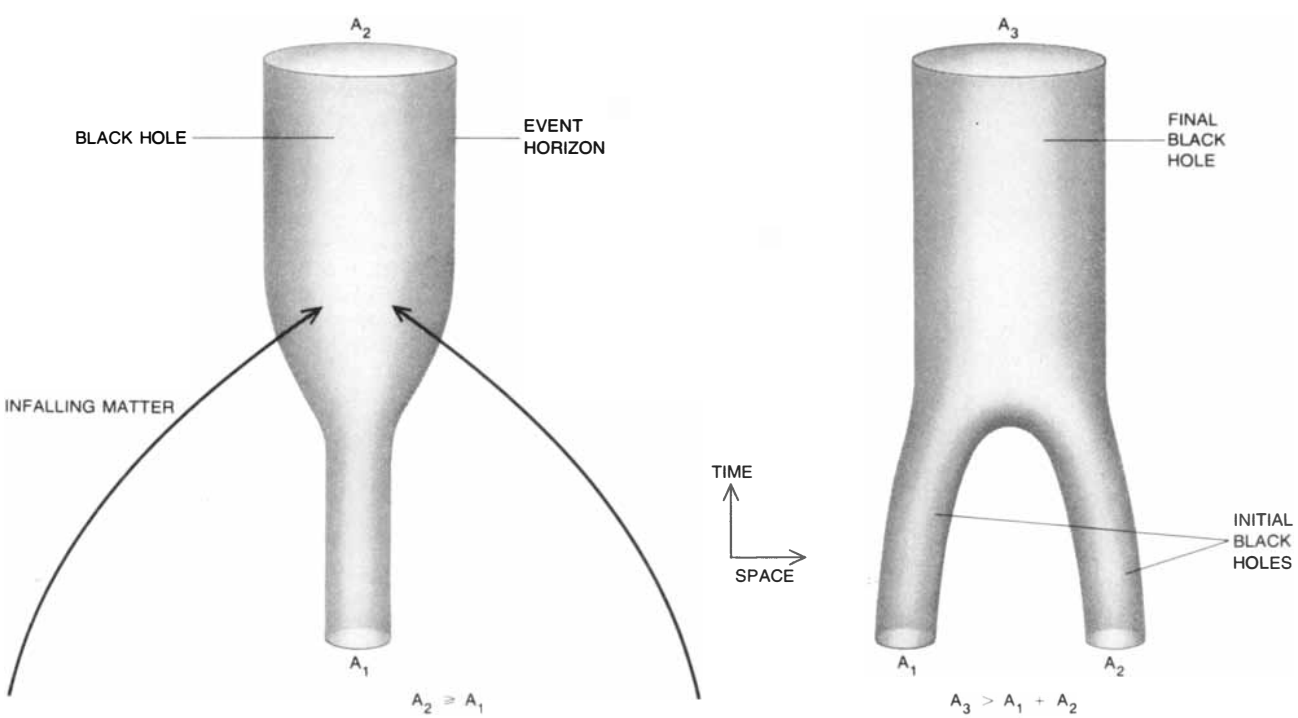
mechanics implies, however, that a particle of mass m behaves like a wave of wavelength h/mc, where h is Planck's constant (the small number 6.62×10^{-27} erg-second) and c is the velocity of light. In order for a cloud of particles to be able to collapse to form a black hole it would seem necessary for this wavelength to be smaller than the size of the black hole that would be formed. It therefore appears that the number of configurations that could form a black hole of a given mass, angular momentum and electric charge, although very large, may be finite. Bekenstein suggested that one could interpret the logarithm of this number as the entropy of a black hole. The logarithm of the number would be a measure of the amount of information that was irretrievably lost during the collapse through the event horizon when a black hole was created.

The apparently fatal flaw in Bekenstein's suggestion was that if a black hole has a finite entropy that is proportional to the area of its event horizon, it also ought to have a finite temperature, which would be proportional to its surface gravity. This would imply that a black hole could be in equilibrium with thermal radiation at some temperature other than zero. Yet according to classical concepts no such equilibrium is possible, since the black hole would absorb any thermal radiation that fell on it but by definition would not be able to emit anything in return.

This paradox remained until early in 1974, when I was investigating what the behavior of matter in the vicinity of a black hole would be according to quantum mechanics. To my great surprise I found that the black hole seemed to emit particles at a steady rate. Like everyone else at that time, I accepted the dictum that a black hole could not emit anything. I therefore put quite a lot of effort into trying to get rid of this embarrassing effect. It refused to go away, so that in the end I had to accept it. What finally convinced me it was a real physical process was that the outgoing particles have a spectrum that is precisely thermal: the black hole creates and emits particles and radiation just as if it were an ordinary hot body with a temperature that is proportional to the surface gravity and inversely proportional to the mass. This made Bekenstein's suggestion that a black hole had a finite entropy fully consistent, since it implied that a black hole could be in thermal equilibrium at some finite temperature other than zero.

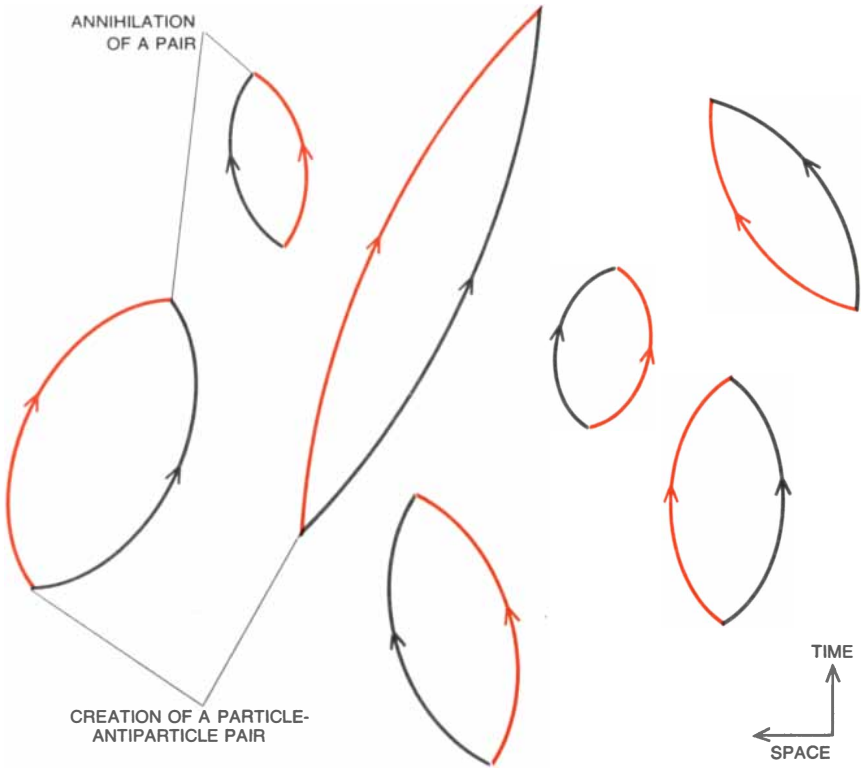
Since that time the mathematical evidence that black holes can emit thermally has been confirmed by a number of other people with various different approaches. One way to understand the emission is as follows. Quantum mechanics implies that the whole of space is filled with pairs of "virtual" particles and antiparticles that are constantly materializing in pairs, separating and then coming together again and annihilating each other. These particles are called virtual because, unlike "real" particles, they cannot be observed directly with a particle detector. Their indirect effects can nonetheless be measured, and their existence has been confirmed by a small shift (the "Lamb shift") they produce in the spectrum of light from excited hydrogen atoms. Now, in the presence of a black hole one member of a pair of virtual particles may fall into the hole, leaving the other member without a partner with which to annihilate. The forsaken particle or antiparticle may fall into the black hole after its partner, but it may also escape to infinity, where it appears to be radiation emitted by the black hole.

Another way of looking at the process is to regard the member of the pair of particles that falls into the black hole—the antiparticle, say—as being really a particle that is traveling backward in time. Thus the antiparticle falling into the black hole can be regarded as a particle coming out of the black hole but traveling backward in time. When the particle reaches the point at which the

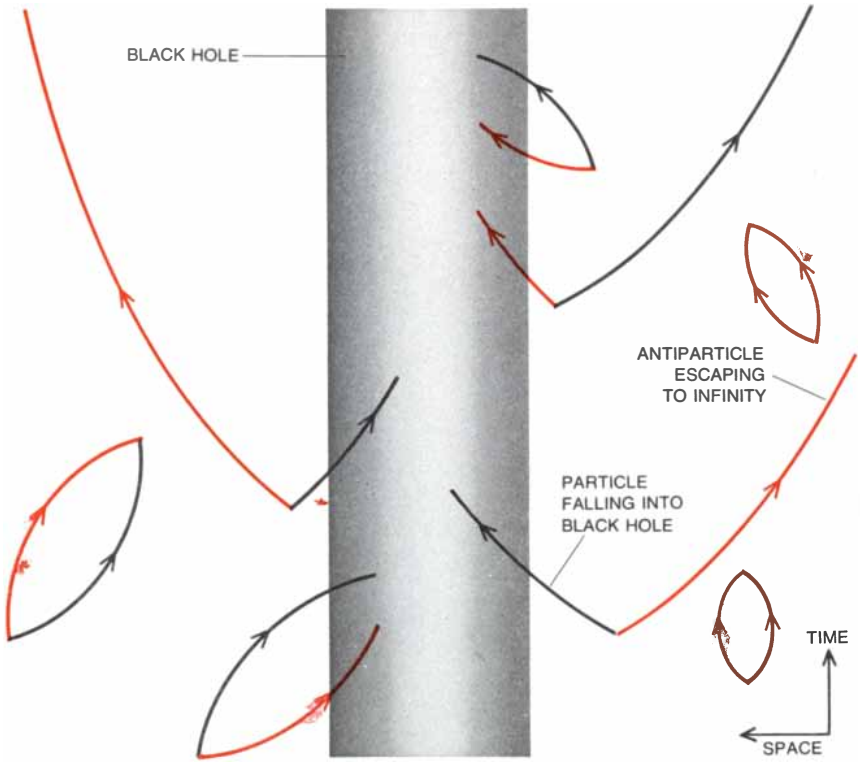


CERTAIN PROPERTIES OF BLACK HOLES suggest that there is a resemblance between the area of the event horizon of a black hole and the concept of entropy in thermodynamics. As matter and radiation continue to fall into a black hole (*space-time configuration at left*) the area of the cross section of the event horizon steadily increases. If two black holes collide and merge (*configuration at right*), the area of the cross section of the event horizon of the resulting black hole is greater than the sum of the areas of the event horizons of the initial black holes. The second law of thermodynamics says that the entropy of an isolated system always increases with passage of time.

© 1976 SCIENTIFIC AMERICAN, INC



"EMPTY" SPACE-TIME is full of "virtual" pairs of particles (*black*) and antiparticles (*color*). Members of a pair come into existence simultaneously at a point in space-time, move apart and come together again, annihilating each other. They are called virtual because unlike "real" particles they cannot be detected directly. Their indirect effects can nonetheless be measured.



IN THE NEIGHBORHOOD OF A BLACK HOLE one member of a particle-antiparticle pair may fall into the black hole, leaving the other member of the pair without a partner with which to annihilate. If surviving member of pair does not follow its partner into black hole, it may escape to infinity. Thus black hole will appear to be emitting particles and antiparticles.

particle-antiparticle pair originally materialized, it is scattered by the gravitational field so that it travels forward in time.

Quantum mechanics has therefore allowed a particle to escape from inside a black hole, something that is not allowed in classical mechanics. There are, however, many other situations in atomic and nuclear physics where there is some kind of barrier that particles should not be able to penetrate on classical principles but that they are able to tunnel through on quantum-mechanical principles.

The thickness of the barrier around a black hole is proportional to the size of the black hole. This means that very few particles can escape from a black hole as large as the one hypothesized to exist in Cygnus X-1 but that particles can leak very rapidly out of smaller black holes. Detailed calculations show that the emitted particles have a thermal spectrum corresponding to a temperature that increases rapidly as the mass of the black hole decreases. For a black hole with the mass of the sun the temperature is only about a ten-millionth of a degree above absolute zero. The thermal radiation leaving a black hole with that temperature would be completely swamped by the general background of radiation in the universe. On the other hand, a black hole with a mass of only a billion tons, that is, a primordial black hole roughly the size of a proton, would have a temperature of some 120 billion degrees Kelvin, which corresponds to an energy of some 10 million electron volts. At such a temperature a black hole would be able to create electron-positron pairs and particles of zero mass, such as photons, neutrinos and gravitons (the presumed carriers of gravitational energy). A primordial black hole would release energy at the rate of 6,000 megawatts, equivalent to the output of six large nuclear power plants.

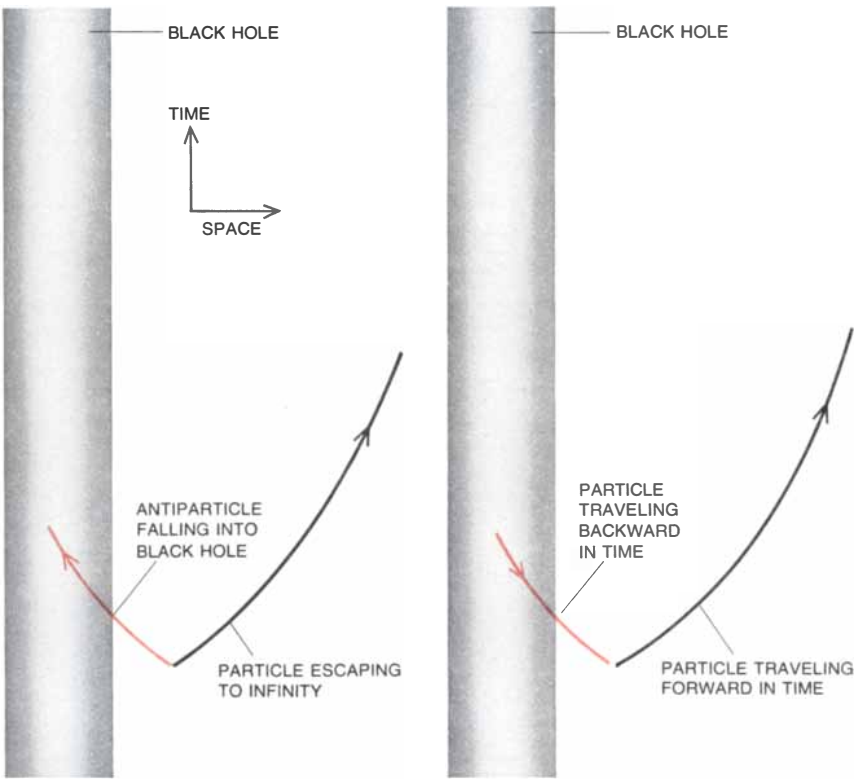
As a black hole emits particles its mass and size steadily decrease. This makes it easier for more particles to tunnel out, and so the emission will continue at an ever increasing rate until eventually the black hole radiates itself out of existence. In the long run every black hole in the universe will evaporate in this way. For large black holes, however, the time it will take is very long indeed: a black hole with the mass of the sun will last for about 10^{66} years. On the other hand, a primordial black hole should have almost completely evaporated in the 10 billion years that have elapsed since the big bang, the beginning of the universe as we know it. Such black holes should now be emitting hard gamma rays with an energy of about 100 million electron volts.

© 1976 SCIENTIFIC AMERICAN, INC

Calculations made by Don N. Page of the California Institute of Technology and me, based on measurements of the cosmic background of gamma radiation made by the satellite SAS-2, show that the average density of primordial black holes in the universe must be less than about 200 per cubic light-year. The local density in our galaxy could be a million times higher than this figure if primordial black holes were concentrated in the "halo" of galaxies—the thin cloud of rapidly moving stars in which each galaxy is embedded—rather than being uniformly distributed throughout the universe. This would imply that the primordial black hole closest to the earth is probably at least as far away as the planet Pluto.

The final stage of the evaporation of a black hole would proceed so rapidly that it would end in a tremendous explosion. How powerful this explosion would be depends on how many different species of elementary particles there are. If, as is now widely believed, all particles are made up of perhaps six different kinds of quarks, the final explosion would have an energy equivalent to about 10 million one-megaton hydrogen bombs. On the other hand, an alternative theory of elementary particles put forward by R. Hagedorn of the European Organization for Nuclear Research argues that there is an infinite number of elementary particles of higher and higher mass. As a black hole got smaller and hotter, it would emit a larger and larger number of different species of particles and would produce an explosion perhaps 100,000 times more powerful than the one calculated on the quark hypothesis. Hence the observation of a black-hole explosion would provide very important information on elementary particle physics, information that might not be available any other way.

A black-hole explosion would produce a massive outpouring of high-energy gamma rays. Although they might be observed by gamma-ray detectors on satellites or balloons, it would be difficult to fly a detector large enough to have a reasonable chance of intercepting a significant number of gamma-ray photons from one explosion. One possibility would be to employ a space shuttle to build a large gamma-ray detector in orbit. An easier and much cheaper alternative would be to let the earth's upper atmosphere serve as a detector. A high-energy gamma ray plunging into the atmosphere will create a shower of electron-positron pairs, which initially will be traveling through the atmosphere faster than light can. (Light is slowed down by interactions with the air molecules.) Thus the electrons and positrons will set up a kind of sonic boom, or shock wave, in the electromagnetic field. Such a shock wave, called Cerenkov radiation, could be detected from the ground as a flash of visible light.



ALTERNATIVE INTERPRETATIONS can explain the emission of particles by a black hole. One explanation (*left*) invokes the formation of a virtual particle-antiparticle pair, one member of which is trapped by the black hole as the other escapes. In another explanation (*right*) one can regard an antiparticle falling into a black hole as being a normal particle that is traveling backward in time out of the black hole. Once outside it is scattered by the gravitational field and converted into a particle traveling forward in time, which escapes to infinity.

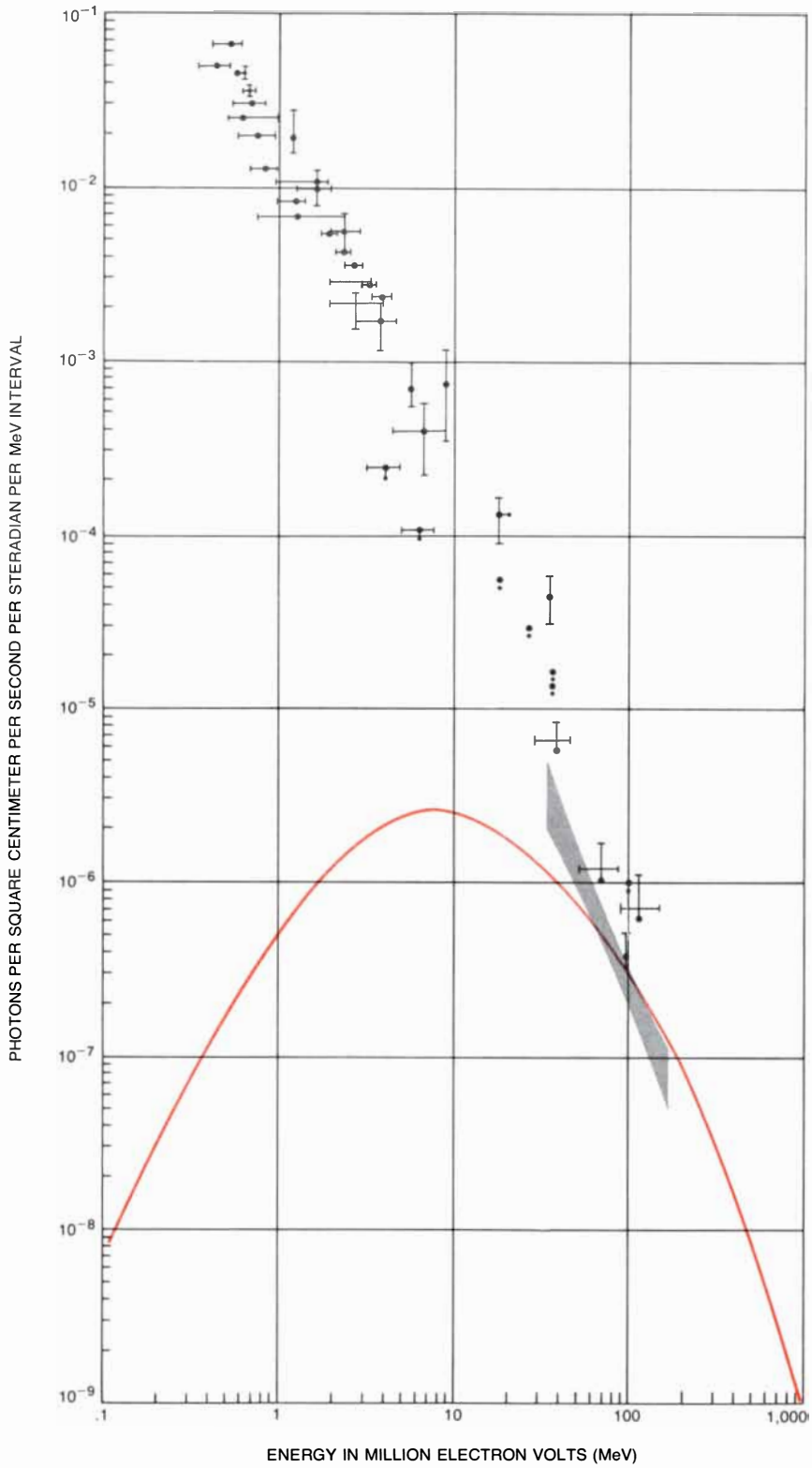
A preliminary experiment by Neil A. Porter and Trevor C. Weekes of University College, Dublin, indicates that if black holes explode the way Hagedorn's theory predicts, there are fewer than two black-hole explosions per cubic light-year per century in our region of the galaxy. This would imply that the density of primordial black holes is less than 100 million per cubic light-year. It should be possible to greatly increase the sensitivity of such observations. Even if they do not yield any positive evidence of primordial black holes, they will be very valuable. By placing a low upper limit on the density of such black holes, the observations will indicate that the early universe must have been very smooth and nonturbulent.

The big bang resembles a black-hole explosion but on a vastly larger scale. One therefore hopes that an understanding of how black holes create particles will lead to a similar understanding of how the big bang created everything in the universe. In a black hole matter collapses and is lost forever but new matter is created in its place. It may therefore be that there was an earlier phase of the universe in which matter collapsed, to be re-created in the big bang.

If the matter that collapses to form a black hole has a net electric charge, the resulting black hole will carry the same charge. This means that the black hole will tend to attract those members of the virtual particle-antiparticle pairs that have the opposite charge and repel those that have a like charge. The black hole will therefore preferentially emit particles with charge of the same sign as itself and so will rapidly lose its charge. Similarly, if the collapsing matter has a net angular momentum, the resulting black hole will be rotating and will preferentially emit particles that carry away its angular momentum. The reason a black hole "remembers" the electric charge, angular momentum and mass of the matter that collapsed and "forgets" everything else is that these three quantities are coupled to long-range fields: in the case of charge the electromagnetic field and in the case of angular momentum and mass the gravitational field.

Experiments by Robert H. Dicke of Princeton University and Vladimir Braginsky of Moscow State University have

© 1976 SCIENTIFIC AMERICAN, INC



PRIMORDIAL BLACK HOLES, each about the size of an elementary particle and weighing about a billion tons, may have been formed in large numbers shortly after the big bang, the beginning of the universe as we know it. Such black holes would have a temperature of about 70 billion degrees Kelvin, corresponding to an energy of 10 million electron volts (MeV). The particles emitted at that energy would produce a diffuse spectrum of gamma rays detectable by satellites. The data points and the shaded region represent actual measurements of the diffuse gamma-ray spectrum in nearby space. The measurements indicate that the average density of such black holes in the universe must be less than about a million per cubic light-year. Solid curve is predicted spectrum from such a density of primordial black holes, based on plausible assumptions about the density of matter in universe and distribution of black holes.

indicated that there is no long-range field associated with the quantum property designated baryon number. (Baryons are the class of particles including the proton and the neutron.) Hence a black hole formed out of the collapse of a collection of baryons would forget its baryon number and radiate equal quantities of baryons and antibaryons. Therefore when the black hole disappeared, it would violate one of the most cherished laws of particle physics, the law of baryon conservation.

Although Bekenstein's hypothesis that black holes have a finite entropy requires for its consistency that black holes should radiate thermally, at first it seems a complete miracle that the detailed quantum-mechanical calculations of particle creation should give rise to emission with a thermal spectrum. The explanation is that the emitted particles tunnel out of the black hole from a region of which an external observer has no knowledge other than its mass, angular momentum and electric charge. This means that all combinations or configurations of emitted particles that have the same energy, angular momentum and electric charge are equally probable. Indeed, it is possible that the black hole could emit a television set or the works of Proust in 10 leather-bound volumes, but the number of configurations of particles that correspond to these exotic possibilities is vanishingly small. By far the largest number of configurations correspond to emission with a spectrum that is nearly thermal.

The emission from black holes has an added degree of uncertainty, or unpredictability, over and above that normally associated with quantum mechanics. In classical mechanics one can predict the results of measuring both the position and the velocity of a particle. In quantum mechanics the uncertainty principle says that only one of these measurements can be predicted; the observer can predict the result of measuring either the position or the velocity but not both. Alternatively he can predict the result of measuring one combination of position and velocity. Thus the observer's ability to make definite predictions is in effect cut in half. With black holes the situation is even worse. Since the particles emitted by a black hole come from a region of which the observer has very limited knowledge, he cannot definitely predict the position or the velocity of a particle or any combination of the two; all he can predict is the probabilities that certain particles will be emitted. It therefore seems that Einstein was doubly wrong when he said, "God does not play dice." Consideration of particle emission from black holes would seem to suggest that God not only plays dice but also sometimes throws them where they cannot be seen.

© 1976 SCIENTIFIC AMERICAN, INC

What a job co-ordinating the people who have lunch on this bridge!

It's a big dining room connecting two wings of the Kodak Research Laboratories.

For some of the troops who dine here, the objective in coming to work each day is ever-closer approximation by dyes to the perceptions of color induced in the human brain by the real world.

For others, the chemical and physical problems that occupy the workday tend to make them equate photographic progress with photographic speed. Photography began with long minutes of exposure in bright sunshine, and it's now up to and beyond the point where something too dark even to see can be photographed without prolonged exposure.*

Other diners up there are a practical bunch. Practicality means processability, as they see it. Certainly when we ease the processor's problems, we give you a better chance for a happy result at an affordable price.

The image structure people eat there too. They are in the business of compressing external reality with maximum detail into minimum space. One viewpoint holds that there photography stands invincible.

The quality that quickly distinguishes antique snapshots and movies with their very black shadows and washed-out faces from today's photography is called latitude. The people who work on that seem to have accomplished a lot, but film still has nowhere near the eye's latitude.

When it's time to amble back to work from the bridge, work for a goodly number concerns stability against fading of colors. Improvements in that field over the years are obvious. Still further improvement has probably been achieved, but how can the achievers prove their achievement before they themselves have aged? There is no real substitute for real time.

It is necessary to keep all those special interests (and others too technical to mention) from conflicting with each other too badly.

Whether used for personal satisfaction or for serious business, film is complicated stuff on which to have to stake a good name. There is more to it than, for instance, the stuff you feed into a typewriter. And we don't even dare stop improving it.

Kodak

*"Prolonged" is a flexible word. For more specific working details on our fastest currently marketed color film, ask Dept. 55W, Kodak, Rochester, N.Y. 14650 for Kodak Publication E-37, "KODAK EKTACHROME Professional Films (Process E-6)."

© 1976 SCIENTIFIC AMERICAN, INC

Questions you should ask and answers you should receive when you are setting up, or re-evaluating your employee savings or profit-sharing plan.

Perhaps we might begin by reminding you that you should look into a number of different plans before committing yourself to any one.

When you do, you will find that most tax-qualified plans offer 1) a common stock fund, 2) a fixed income fund, and 3) sometimes a company stock fund. This way, employees can place their contributions in the area or areas that are best suited to their personal needs and goals.

Among these three kinds of funds, most companies today are showing the greatest interest in fixed income funds. They are turning to insurance companies which can provide guarantees of principal and interest that are not available through bond funds, company stock funds or common stock funds. In addition, only insurance companies can provide guaranteed lifetime payments to employees when they retire.

Now assuming that your interest is in fixed income funds, let's get at some questions you should ask before selecting an insurance company from one of the many in this field.

Q. Do I have a guarantee of principal for my participants at all times and under all circumstances?

A. You should for this reason: Suppose your plan has to pay out to employees more than is contributed in a given year? This could occur if there are an unusual number of early retirements or if a plant closes. *Regardless of what happens, Metropolitan guarantees the principal to your participants.*

Q. If I need to can I cancel my contract at any time? If I do will I be guaranteed book value?

A. You should have the right to cancel. You should also be guaranteed reimbursement at book value, but it may involve some form of installment arrangements. At Metropolitan, you can cancel at any time and for any reason and book value is always guaranteed under some form of installment option.

Q. If interest rates increase over the years my contract is in force, will my employees participate in the higher rates?

A. Yes they should participate. Employees know when interest rates have risen and expect to share in the increased earnings. But some insurance companies give increases only under limited conditions. For example, you might participate in increased interest, but only if the insurance company is still writing this type of contract and if credited interest rates are at least ½% higher than your guaranteed rate. At Metropolitan your employees are guaranteed participation in increased earnings when interest rates go up and are protected by minimum guarantees when interest rates go down.

These are just a few of the questions you should ask before awarding your employee savings or profit sharing plan to any company. But there are many other questions of great importance.

If you have others (and you probably do) call or write:

Mr. Leon D. Moskowitz, Vice-President,
Group Insurance & Pensions Marketing,
Metropolitan Life, One Madison Avenue,
New York, N.Y. 10010 (212) 578-3148.

Metropolitan
Where the future is now

© 1976 SCIENTIFIC AMERICAN, INC

SCIENCE AND THE CITIZEN

Operators

A continuing complaint by the public about health care in the U.S. has been that there are too many specialists and not enough physicians who deliver a broad spectrum of family care. Professional manpower studies have confirmed the public perception, and in recent years the medical schools, prodded and supported by state and Federal funding, have moved to increase the supply of family physicians. Nevertheless, some 30 percent of American physicians consider themselves to some degree specialists in surgery, and about a fourth of those beginning residency training have been going into surgical programs. At the Harvard Medical School, Osler L. Peterson and his colleagues undertook to find out just who are doing surgery in the U.S. and how much they do, and whether fewer surgeons could handle the load. Their conclusions have been published in *The New England Journal of Medicine:* Too many physicians of all kinds are performing operations, and most of them—even many of the best-trained ones, thoroughly committed to surgery—have modest work loads.

Peterson, Rita J. Nickerson, Theodore Colton, Bernard S. Bloom and Walter W. Hauck, Jr., examined hospital records for 1970 in four metropolitan regions and noted the operations performed by each physician. Each operation was weighted, on an index that had been devised by the California Medical Association, according to its complexity and the time required for preoperative and postoperative care; for example, a normal delivery or a tonsillectomy had a weight of 4, an appendectomy a weight of 9.5 and the repair of an abdominal aortic aneurysm a weight of 40. The total weighted work load of each of 2,700 physicians was aggregated for the year. The largest work loads, recorded for surgical specialists who had been certified by their specialty board, had a median "California relative value" equivalent to about 180 "typical" operations a year. The operating load of self-described surgical specialists lacking board certification was only 60 percent as great; that of general practitioners with a secondary specialty in surgery was less than 20 percent of the board-certified surgeons' load. Other general practitioners, osteopathic physicians and medical specialists (internists, pediatricians, cardiologists and so on) did far fewer operations a year.

A large fraction of the physicians did very few operations a year and a small fraction did a lot. There was a wide variation in work loads. Thirty-one percent of the 2,700 physicians who did operations had California-relative-value loads of 50 (equivalent to five appendectomies) or less; 14 percent had loads of 2,000 (equivalent to 210 appendectomies or more). As might be expected, the less specialized physicians tended to do fewer and less complex operations, but the variation in load was large even among board-certified surgical specialists. The busiest 10 percent of them did more than 350 typical operations a year and the least busy 25 percent of them did fewer than about 100 a year. A special analysis revealed that it takes a long time to build up a surgical practice to even a modest work load. A study of 780 surgical specialists who had graduated from medical school no more than 25 years before showed that it had taken them about 18 years to reach their maximum operating load and about 13 years to reach 90 percent of that maximum.

The investigators point out that a work-load study provides better evidence on manpower requirements than such traditional measures as physician-to-population ratios or the number of unfilled surgical residencies in hospitals. They base their conclusion that too many physicians of all kinds are operating largely on the broad variation in operating loads and the long time it takes a surgeon to reach his maximum load.

What if the operating load were redistributed? If only surgical specialists did all operations except the very simplest ones, their work load would be increased by 16 percent, or by about two operations a month. If the operations done by specialists who in 1970 had loads of fewer than 50 operations a year were redistributed among those who did more work, the mean load of those busier specialists would increase 20 percent. If surgical operations were limited to board-certified surgeons, the operating load of those surgeons would be increased by 57 percent to just over 300 operations a year, or about six a week. "There is no need to have surgical operations performed by both nonsurgeons and well-trained surgeons," Peterson and his co-workers maintain. "A single standard of surgery is a desirable goal whose achievement will be favored by reduction of the number of physicians performing surgical operations, particularly those least trained and least experienced."

More Miles per Gallon

If 11 million 1977-model automobiles are sold in the U.S., as is projected, they will burn three billion gallons of fuel less in their first year on the road than they would if they ran as inefficiently as cars built only three years ago. With gasoline now costing about 65 cents a gallon (compared with about 41 cents in 1974) the fuel saving will amount to some $2 billion. Largely in response to Federal legislation the sales-weighted fuel economy of 1977 models has climbed to 18.6 miles per gallon compared with 13.9 m.p.g. in 1974, the historic low point reached after a long downward slide that began at the end of World War II. Coming on top of substantial improvements in 1975 and 1976 models, the combined city-highway fuel economy of 18.6 m.p.g. achieved for the 1977 automobile fleet actually exceeds the 18 m.p.g. mandated for next year's fleet in the Energy Policy and Conservation Act of 1975. Of the 13 automobile manufacturers whose products are evaluated in the most recent Environmental Protection Agency (EPA) report on fuel economy, seven—BMW, Nissan (Datsun), Toyo Kogyo (Mazda), Toyota, Volkswagen, Audi and Fuji (Subaru)—are already exceeding the 1980 standard of 20 m.p.g. In fact, three of the seven (Toyota, Volkswagen and Fuji) have surpassed the standard of 27.5 m.p.g. mandated by Congress for the 1985-model year. The high fuel economy achieved by foreign automobile manufacturers is of course partly due to the lower average weight of their vehicles.

The projected weight of 3,923 pounds for the average 1977 car represents a drop of 4 percent from 1975, when the average weight reached an all-time high (4,088 pounds). Taken by itself this would have yielded an improvement of only about 1.5 percent in fuel economy. The difference between 1.5 percent and the 19 percent improvement actually achieved between 1975 and 1977 represents the increase due to changes or improvements in engines, transmissions, axle ratios and emission-control systems. It is noteworthy that the 1977 improvement in fuel economy was reached the same year that Federal exhaust-emission standards were further tightened both in the U.S. as a whole and in California, where the limits are even more stringent. For 1977 models sold nationwide the maximum permissible emission of compounds of nitrogen and oxygen (referred to as NO_x) was reduced from 3.1 grams per mile to two grams; limits on hydrocarbons and carbon monoxide remained unchanged at 1.5 grams per mile and 15 grams. In California the carbon monoxide limit remained at nine grams per mile (40 percent below the Federal standard); the hydrocarbon limit was lowered to .41 gram per mile and the NO_x limit to 1.5 grams. To meet the stricter emission standards 1977 cars sold in California pay a penalty in the form of poorer fuel economy. If Californians bought the same assortment of automobiles as

© 1976 SCIENTIFIC AMERICAN, INC

Old Faithful.

Good Old Faithful. Always got you where you were going. Always did it economically. A whole generation of Americans grew up with Old Faithful. And now, 27 years and 33 million cars later, it's still a symbol of dependability and economy.

Now there's a car that's just as reliable and economical as Old Faithful ever was. It's New Faithful. The 1977 VW Rabbit. With engineering so advanced that automotive experts have hailed it as the kind

©VOLKSWAGEN OF AMERICA, INC.

© 1976 SCIENTIFIC AMERICAN, INC

New Faithful.

of car Detroit will be building in the 1980's.

The Rabbit has a new fuel injection system, so it starts up quick as a bunny. Springs like one, too. 0 to 50 in just 7.7 seconds. The Rabbit also has advanced engineering features like negative steering roll radius to help maintain directional stability in the event of a front-tire blowout; rack-and-pinion steering for more direct maneuvering and better road feel; and an independent stabilizer rear axle, low in unsprung weight, for better road holding.

New Faithful lives up to Old Faithful's reputation for economy, too. Because it has fuel injection, you can use the most economical grade of gas.* But you won't have to use it very often. Rabbit gets 37 mpg on the highway, 24 in the city. (That's EPA's estimate for manual transmission. Your actual mileage may vary, depending upon your driving habits, your car's condition and optional equipment.)

Dependability and economy. That's what Old Faithful gave a whole generation of Americans. And that's what New Faithful is giving a whole new generation of Americans.

New Faithful. The 1977 VW Rabbit.

VW Rabbit

More Volkswagen from Volkswagen

*California excluded.

© 1976 SCIENTIFIC AMERICAN, INC

Boyne Falls... A Natural Alternative

In todays society of plastic simulations, a Boyne Falls home offers you a natural alternative in housing. Made from hand peeled logs, if you desire a traditional design, or sculptured solid cedar for the contemporary taste. Our pre-engineered, pre-cut homes offer an infinite number of design possibilities. They are naturally durable, maintenance free, fire resistant, insulated (an independent testing lab confirms that our homes consume a minimum of 19% less fuel than conventionally built home), and are warranteed for 10 years.

For full details, send three dollars for our colorful portfolio or request our free folder.

Boyne Falls
LOG HOMES

Boyne Falls Log Homes, Inc.
Dept. 7-G
Boyne Falls, Michigan 49713
(616) 549-2421

those bought elsewhere in the U.S., the penalty would be about 12 percent. Actually Californians buy a larger fraction of lighter imported cars, with the result that the 1977 models sold in California should achieve a fuel economy of 18 m.p.g., only 3.4 percent below the 49-state average. According to EPA projections, cars weighing 3,500 pounds or less will account for nearly half of California sales but for less than a third of sales elsewhere. As a result the 1977 models sold in California already meet the Federal fuel-economy standards for 1978 and the proposed 1978 emission standards for hydrocarbons. The industry has asked Congress to amend 1970 legislation that mandated "ultimate" emission standards in 1978: .41 gram of hydrocarbons, 3.4 grams of carbon monoxide and .4 gram of NO_x per mile.

California Dams Revisited

The earthquake that shook the northern suburbs of Los Angeles in February, 1971, came very close to causing a major disaster when the severely damaged Lower San Fernando Dam barely held together long enough for the water in the reservoir to be lowered to a safe level. During the four-day drainage operation some 80,000 people living downstream from the dam were evacuated. The almost total failure of the dam in the earthquake also shook the confidence of state and local officials in the analytical procedures routinely adopted up to that time for evaluating the safety of dams in seismically active areas. A subsequent investigation revealed that the near disaster had resulted from the liquefaction of the dam's soil embankment, which had been built half a century before by the then common hydraulic-earthfill method. Previous conventional analyses of the dam site, the latest one conducted as recently as 1966, had failed to predict the instability of the embankment in an earthquake as a result of the liquefaction phenomenon.

In the aftermath of the San Fernando experience an alternative method of evaluating the design of dams, called the dynamic-analysis technique, has been widely adopted by the Division of Safety of Dams of the state's Department of Water Resources. The new method, developed principally by H. Bolton Seed and his colleagues at the University of California's College of Engineering, has already established that in seismically active regions, such as metropolitan Los Angeles and the San Francisco Bay area, hydraulic-earthfill dams are generally unsafe. Most of the dams in this category are now in the process of being either relocated, completely rebuilt at the original site or substantially rehabilitated.

The Division of Safety of Dams is

© 1976 SCIENTIFIC AMERICAN, INC

currently extending the dynamic-analysis technique to the study of other types of dams in earthquake-prone areas. One problem has been to develop a priority system for studying the dams; there are more than 1,100 dams in California that are subject to regulation by the state, not counting an additional 114 dams under Federal jurisdiction. According to a recent article in the periodical *California Geology* by Gordon W. Dukleth and E. W. Stroppini, two engineers with the Division of Safety of Dams, "the priority listing is based on the factors of location of the dam with respect to earthquake hazard potential, potential for jeopardy to life and property in the event of dam failure, and susceptibility of the soils in the embankment or foundation of the dam to failure during an earthquake."

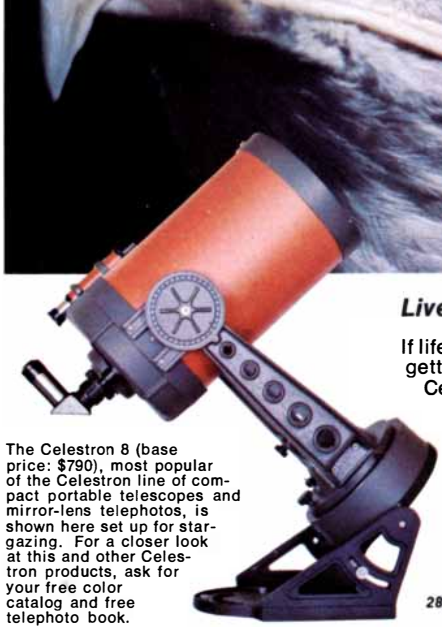
The current focus of attention in the state is the "Palmdale bulge," a region of rising ground northeast of Los Angeles, where a potential threat to public welfare has been declared by the Seismic Safety Commission based on the prediction last year of a possible earthquake in the area. Dukleth and Stroppini report that even before the recent earthquake prediction action had already been taken on the most critical of the 50 or so dams in the vicinity of the bulge. In one case, that of the Bouquet Canyon dam, they point out, a detailed investigation of the dam's ability to withstand an earthquake had been scheduled for a later date, "but it has been expedited because of the Palmdale bulge and the prediction of possible earthquake events in the area." All in all, Dukleth and Stroppini conclude, "the advances in analytical techniques resulting from studies following the San Fernando earthquake have promised a higher degree of safety of dams than ever before possible."

To Clone a Gene

Two powerful techniques of contemporary molecular genetics—recombinant DNA technology and DNA synthesis in the laboratory—have recently been combined for the first time in an effort to understand the DNA sequences called regulatory genes, which control the expression of adjacent structural genes within the functional genetic unit known as the operon. Writing in *Nature,* two independent collaborations, one headed by Ray J. Wu of Cornell University and Saran A. Narang of the National Research Council of Canada and the other by Herbert W. Boyer of the University of California at San Francisco, Richard E. Dickerson of the California Institute of Technology and Arthur D. Riggs of the City of Hope Medical Center in Duarte, Calif., describe their successful efforts to produce large amounts of a synthetic regulatory gene for struc-

Eagle at 30 ft. Photographed by W. Beecher with the Celestron 1250mm, f/10 Multipurpose Telephoto Lens.

Saturn's Rings. Photographed with Celestron 14 Telescope.



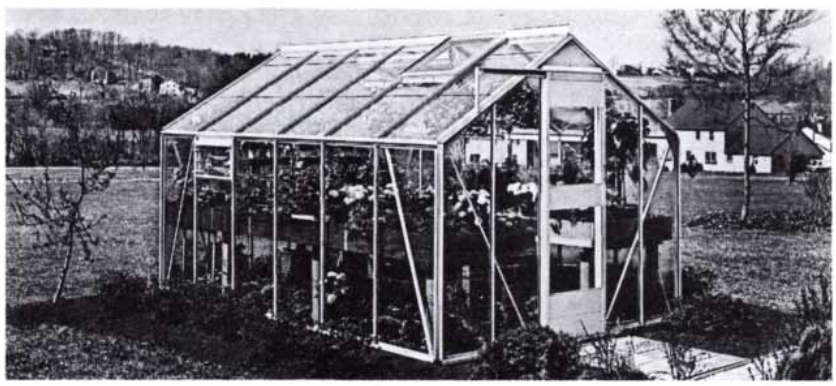
The Celestron 8 (base price: $790), most popular of the Celestron line of compact portable telescopes and mirror-lens telephotos, is shown here set up for stargazing. For a closer look at this and other Celestron products, ask for your free color catalog and free telephoto book.

Live A Little.

If life is about any one single thing, it's about getting closer to the things that interest you. Celestron Telescopes help you get closer. Whether the subject is the muscular structure of an eagle's eye at 30 feet. Or the planet Saturn — one hour, 11 minutes away at the speed of light. If you want to get closer, look into the world's best-selling modern telescopes. Live a little.

Celestron International
2835 Columbia, Box 3578-B
Torrance, Calif. 90503

Celestron



At last! A beautiful, professional-quality greenhouse every home gardener can afford.

The Burpee Greenhouse—priced under $700. After years of searching for the ideal greenhouse, Burpee decided to use our three generations' experience in growing under glass to design our own.

The Burpee Greenhouse is available right now. The materials are first quality throughout. The design is absolutely beautiful, and the price, beginning at $645, makes it an exceptional value for the home gardener.

For pictures, a detailed brochure, and complete price information, without obligation, send in this coupon now.

W. ATLEE BURPEE COMPANY
6117 Burpee Bldg., Warminster, PA 18974

NAME (please print)
ADDRESS
CITY
STATE ZIP

© 1976 SCIENTIFIC AMERICAN, INC

A speaker unlike any other.

Introducing the Bose 901® Series III: the most innovative new speaker since the original Bose 901 was introduced in 1968. The 901 Series III reproduces music with spaciousness and realism unequalled, we believe, by any other speaker. Yet it uses less than ⅓ as much power as the original 901, due to a new, high-efficiency driver. Outstanding bass performance is made possible by the unique Acoustic Matrix™ enclosure (shown in this photo of the 901 III without its grille and walnut veneer cabinet). To fully appreciate its performance, ask a Bose dealer to play the 901 III in comparison to any other speaker, regardless of size or price. For a free brochure, write Bose, Box SA1, The Mountain, Framingham, MA 01701.

BOSE®

Patents issued and pending. Copyright © 1976 Bose Corp.

altair™

Number One in low-cost computing.

Altair, from Mits, is the number one name in microcomputers for home, business, personal and industrial applications. Because the Altair was first, it has set the standard in the industry. More Altair 8800's are now operational than all other microcomputers combined.

Whether you buy a $395 complete computer kit* or a multi-disk system for under $10,000; Mits will provide you with thorough and lasting support. Satisfied Altair users include schools, corporations, small businesses, students, engineers, and hobbyists.

Altair hardware includes three microcomputers; the Altair 8800a, 8800b, and 680b. Mits has a complete selection of Altair plug-compatible memory and interface options, including the new Altair 16K Static board and Altair multi-port serial and parallel I/O boards. Also available is a complete line of Altair peripherals including line printers, CRT's, and multiple disk systems.

Altair software is by far the most complete and best for any microcomputer. Our Extended BASIC and Disk BASIC have received industry wide acclaim for programming power and efficiency. Application packages are available at many Altair Computer Centers.**

The Altair computer is a revolution in low cost computing. Shouldn't you write for more information including our free, color catalogue.

*The Altair 680b turnkey model.
**Retail Altair computer outlets now opened in many large cities.

mits

MITS, Inc. 2450 Alamo S.E. / Albuquerque, New Mexico 87106

tural analysis by generating multiple copies of it within living cells of the intestinal bacterium *Escherichia coli*.

The gene investigated by both collaborations was the *lac* operator, which lies on the *E. coli* chromosome next to three structural genes coding for the synthesis of enzymes involved in the metabolism of the sugar lactose. Like other regulatory genes, the *lac* operator is not itself transcribed into messenger RNA; instead it controls the transcription of the nearby structural genes by serving as a recognition site for a protein called the *lac* repressor. In the absence of lactose the repressor (about 10 copies of which exist per cell) binds to the operator sequence, physically blocking the transcription of the *lac* structural genes by the enzyme RNA polymerase. When lactose is present, however, a modified form of the sugar binds to the repressor, changing its shape so that it is unable to "recognize" the operator DNA. When all the repressor molecules have been similarly inactivated, RNA polymerase is able to attach to the DNA at a site adjacent to the operator called the promoter, and the three *lac* structural genes are sequentially transcribed.

The mechanism by which the *lac* repressor specifically recognizes and associates with the *lac* operator gene out of thousands of other nucleotide sequences on the DNA molecule is of great interest, but its elucidation will require the crystallization and analysis of repressor-operator complexes, a task that presupposes large amounts of purified operator DNA. (Pure *lac* repressor is already available.) To this end the 21-nucleotide-pair sequence of the *lac* operator has been determined and chemically synthesized, but only meager amounts of the pure gene have been obtained in the laboratory. In order to multiply the yield of the synthetic operator gene, the California and Cornell groups devised a novel application of recombinant DNA technology. Making use of the remarkable specificity of restriction enzymes, which recognize a single palindromic nucleotide sequence on the DNA helix and make staggered cuts in both chains of the helix whenever this sequence appears, they synthesized the known recognition sequence for the restriction enzyme designated *Eco*RI and added one copy to each end of the synthetic operator gene. The same restriction enzyme was then used to make complementary single-strand overhangs in a plasmid, a circular piece of nonchromosomal DNA that can be readily transferred from one bacterium to another.

Since the synthetic single-strand overhangs added to the *lac* operator DNA matched those created by the restriction enzyme in the plasmid DNA, these regions paired up spontaneously when the two molecules were incubated together.

© 1976 SCIENTIFIC AMERICAN, INC

The enzyme polynucleotide ligase was then added to covalently join the sugar-phosphate backbones of the two DNA's so that the synthetic operator gene was in effect spliced into the plasmid. When the plasmid-operator hybrids were introduced into *E. coli* cells, they proceeded to reproduce themselves many times over.

Bacterial colonies containing the hybrid DNA (about .3 percent of the total) were identified by two criteria: first, they were resistant to the antibiotic tetracycline by virtue of a gene carried by the plasmid, and second, the 30 or so copies of the *lac* operator produced by the replication of the plasmid vehicle quickly bound up the 10 molecules of *lac* repressor in each cell, thereby preventing the repressor from binding to the single *lac* operator gene on the *E. coli* chromosome. As a result the *lac* structural genes of the host were transcribed, and the affected bacteria began to manufacture the three enzymes required for the metabolism of lactose even though lactose itself was not present in the medium. One of these enzymes, beta-galactosidase, was detected by the California group by feeding the cells a lactose analogue containing a molecule of the dye indigo. The synthetic sugar lacked any inducing effect on the repressor, but it did serve as a substrate for beta-galactosidase. When it was cleaved by the enzyme, it released the dye, coloring those colonies blue.

Replicated operator-plasmid hybrids were then extracted from the blue colonies and digested with restriction enzyme, so that the restriction-sequence "linkers" on each side of the *lac* operator were cleaved. The operator DNA was then liberated from the plasmid DNA, yielding many times the original amount of the synthetic gene. Even so, the increased yield produced by the cloning of single operator sequences was still too low for physicochemical experiments, and both groups of workers are now attempting to produce plasmids containing multiple copies of synthetic *lac* operator. Wu and his collaborators report that they have recently used polynucleotide ligase to link up to 18 operator genes in tandem, with synthetic restriction-sequence linkers between them.

The potential of this technique clearly extends far beyond its present application. In principle restriction-sequence linkers could be added to any piece of synthetic DNA and the desired sequence could then be cloned in plasmids in large quantities. Moreover, Wu and his colleagues propose that synthetic *lac* operator could be introduced in front of a given gene or group of genes, enabling the experimenter to turn these genes on or off with the *lac* repressor. Thus chemical synthesis and cloning technology together enable the molecular geneticist to fabricate any DNA sequence he wants, to produce it in unlimited amounts and finally to insert it next to any other piece of DNA he chooses. For the first time man has developed the capacity for almost absolute control over the material of the genes.

Spooner's Other Isms

The Rev. William Archibald Spooner was an Oxford don for almost 60 years, and his name has a permanent place in the English language; his fame, however, rests not on pedagogy or scholarship but on his legendary verbal eccentricities. Spooner is known for his propensity to slips of the tongue in which words or parts of words are transposed: spoonerisms.

As warden of New College, Oxford, Spooner is supposed to have dismissed a student with the harsh words: "You have deliberately tasted two worms and you can leave Oxford by the town drain." During World War I he is reported to have said: "When the boys come back from France we'll have the hags flung out." In another anecdote he ventures the opinion that "the Lord is a shoving leopard."

The authenticity of these and many other errors attributed to Spooner is questionable, and some have doubted that he ever uttered a spoonerism. Arnold Toynbee concluded that most of the known examples are spurious, having been invented by "ingenious Oxford minds" whose wit was deliberate rather than accidental. Toynbee even offered a criterion for evaluating spoonerisms: "The wittier or more elegant the specimen, the less likely it is to be authentic." Nevertheless, the legend may not have been without basis in fact. For example, there is apparently good reason to believe the story that in chapel at New College, Spooner once announced a hymn as "Kinquering Congs Their Titles Take."

One of the difficulties in studying Spooner's speech is that an oral lapse is by its nature evanescent. John M. Potter of Oxford has recently collected what may amount to corroborative evidence of Spooner's eccentricity, uncontaminated by anecdotal embellishment. He has examined all known manuscripts in Spooner's hand, searching for telltale errors in his use of written language. Potter did not find an extraordinary number of errors, and there were no dazzling examples of unconscious wit, but the kinds of mistakes made do suggest an unusual pattern. Potter's findings are reported in *Proceedings of the Royal Society of Medicine.*

After the funeral of a bishop's widow, Spooner wrote in his diary: "We buried her in the grace close by her husband under the tall cross." The erroneous "grace," Potter suggests, could be regarded as "an interesting blend between 'grave' and 'place,' with perhaps a bonus for Freud of religious association." Another diary entry, on Spooner's 50th birthday, reads: "My wife has been in many ways a singularly peaceful and happy but more limited than I had once hoped." The word "one" is missing from the sentence (after "happy") but there is a more spectacular error: Spooner almost certainly meant to write "life" rather than "wife."

In 250,000 words of text Potter discovered 45 mistakes. Nine of these are straightforward errors in spelling. The remaining 36 have a more complicated nature.

In general, studies of large numbers of linguistic errors give a statistical distribution for various categories of verbal slips. By far the commonest variety are anticipations, in which a word appears redundantly before its normal place in a sentence, and the next commonest are perseverations, the opposite of anticipations. In one investigation these two categories accounted for 93 percent of all the errors.

Spooner's slips of the pen are distributed very differently. Of the 36 nonspelling errors found by Potter only eight were anticipations and only five were perseverations. "The remaining 23 errors were a hotch-potch of transpositions, contaminations and substitutions," all errors that are ordinarily uncommon.

Serious speech defects are a not uncommon result of neurological damage or disease, and it is possible, Potter speculates, that Spooner's more subtle linguistic abnormalities may have reflected an organic disorder. The evidence is too slight, however, even to prove there was an abnormality, much less to diagnose it. An additional point that is intriguing even though its pertinence cannot be ascertained is that Spooner was an albino, although his pigmentation was only reduced rather than being completely absent. In several mammalian species albinism is known to be associated with particular disruptions of certain nerve pathways.

Finally, there is a possibility that Spooner committed spoonerisms not only by word but also by deed. In one story Spooner suggested turning on a light before escorting a guest down a dark and slippery stairway. He then turned off the light and proceeded in complete darkness. On another occasion he is said to have acted out a reversal of a traditional method of spot removal: rubbing salt into a wine stain. At a dinner party Spooner had upset a saltcellar. He thereupon proceeded to pour some claret onto the spilled salt, drop by drop.

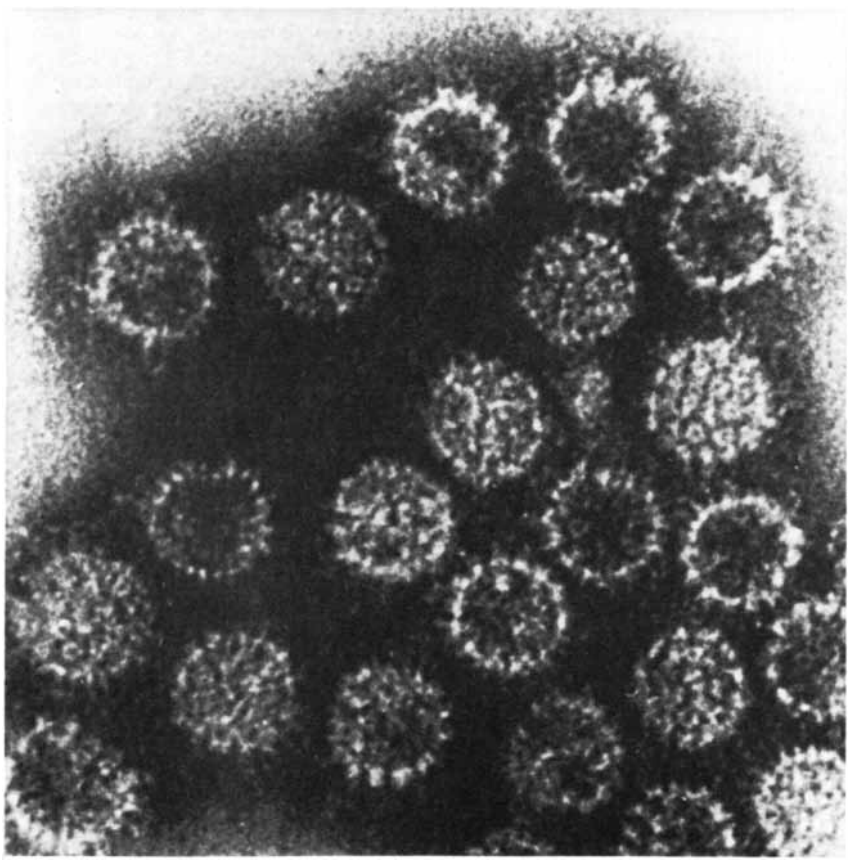
© 1976 SCIENTIFIC AMERICAN, INC

The Antibody Combining Site

When an antigen, or foreign substance, combines with an antibody molecule, it does so at a site that fits it precisely as a lock fits a key. The nature of this site has been worked out in considerable detail

by J. Donald Capra and Allen B. Edmundson

When the body is attacked by bacteria or viruses, it has two principal means of fighting back. First, white blood cells can move to the point of attack to engulf and destroy the foreign agents. Second, cells known as plasma cells can manufacture antibodies, highly specific proteins that are transported in the bloodstream to disable the invaders. The part of the attacking organism that is responsible for stimulating the production of antibody is called an antigen, and it is usually a small patch on the surface of a large protein or carbohydrate molecule associated with the outer structure of the microorganism. A single bacterium, and even a large protein molecule such as a bacterial toxin, can have several such antigenic sites on its surface, and hence it can induce the manufacture of several different specific antibodies. Once they are released into the bloodstream, these antibodies will combine with their corresponding antigens like the complementary pieces of a jigsaw puzzle, thereby modifying the outer structure of the invading microorganism so that a white blood cell can ingest it more readily. In addition antibodies alone can neutralize microorganisms or toxins by linking them together into large inactive complexes.



ANTIBODIES NEUTRALIZE INVADING VIRUSES by binding specifically to small antigenic sites on the outer surface of the virus. Each antibody molecule has two identical combining sites, so that it can bind to adjacent regions of the same virus particle or cross-link two virus particles in an inactive complex. In this electron micrograph antibody molecules form faint halos around spherical polyoma virus particles 400 angstroms in diameter. By binding to repeating antigens on a single virus, antibodies alter the virus's structure so that it can be more readily engulfed by white blood cells. Micrograph was made by June Almeida of the Ontario Cancer Institute and Bernhard Cinader and Allan F. Howatson of the University of Toronto.

Although antibodies are clearly essential for the survival of individuals and species, they are also potentially hazardous. If they are not highly specific for foreign antigens, they may attack the body's own molecules, many of which closely resemble those of invading microorganisms. Something like this may actually happen in the disorders known as autoimmune diseases, where the body apparently loses the ability to distinguish between "self" and "nonself" and makes antibodies that act against its own proteins or nucleic acids, often causing severe and sometimes fatal tissue damage.

The evolutionary success of antibodies has therefore required that they be highly sensitive to minor structural changes in the antigens they recognize. That this is the case was shown some 50 years ago, when Karl Landsteiner of the Rockefeller Institute for Medical Research found that antibodies could distinguish between optical isomers: molecules with identical chemical composition that differ only in the location of a few of their atoms in three-dimensional space. In recent years the application of the concepts and methods of molecular biology to the three-dimensional structure of the antibody combining site has yielded important new perceptions of how antibodies can act with such impressive specificity.

Structure of the Antibody Molecule

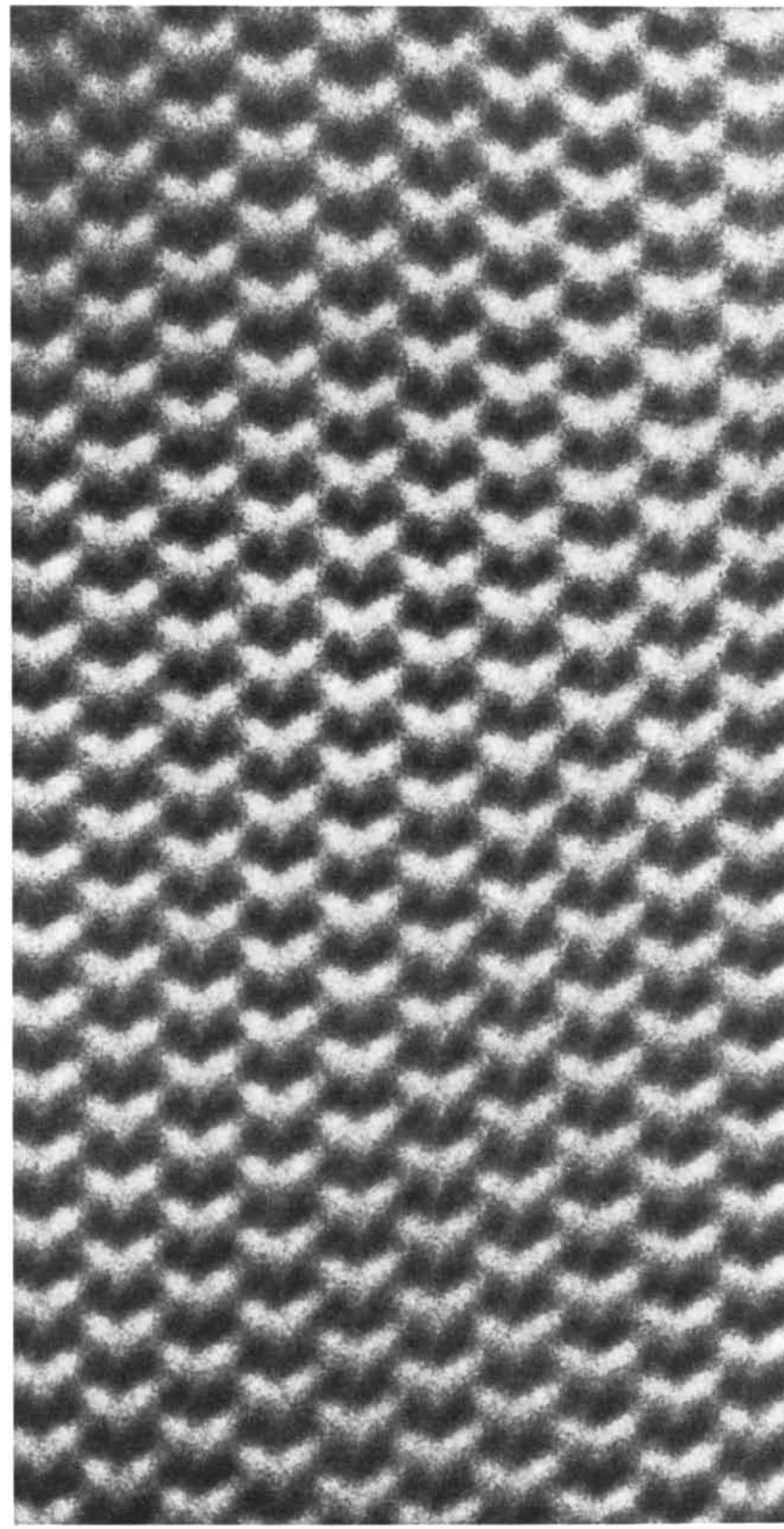
Antibodies belong to the class of proteins called immunoglobulins. Immunoglobulin gamma (designated IgG) is the

© 1976 SCIENTIFIC AMERICAN, INC

predominant antibody in the blood of most higher vertebrates and has been the most intensively studied of the five types of immunoglobulins. It has a molecular weight of about 160,000, which indicates that its molecule consists of some 23,000 atoms. In the early 1960's two groups, one under Rodney R. Porter in England and the other under Gerald M. Edelman at Rockefeller University showed that IgG is made up of four polypeptide chains, each consisting of amino acid units joined by peptide bonds. The four chains are paired so that the molecule consists of two identical halves, each with one "heavy," or long, chain and one "light," or short, chain. The two chains of each pair are cross-linked by covalent bonds between the sulfur atoms of the amino acid cystine. If these disulfide bonds are split, the heavy and light chains will remain bound to one another by noncovalent interactions. When they are put into a solution of acid or urea, however, they will dissociate into heavy and light chains, which can be separated by their difference in size.

Experiments conducted by Porter's group provided a view of the antibody molecule from the functional aspect. When IgG molecules are treated with the protein-cleaving enzyme papain, the molecule divides into three parts of about equal size. Two of them are identical and are designated Fab, for "fragment, antigen-binding." The Fab fragments of an IgG molecule each have a combining site. It is because the intact antibody molecule has two such sites that it is able to cross-link antigenic materials into inactive complexes. The third fragment produced by papain digestion is designated Fc because it crystallizes readily. It does not bind antigen, but it has other important biological activities.

The results obtained by many investigators in the 1960's were combined in the currently accepted model of IgG, in which the heavy and light chains are arranged in the shape of a *Y*. According to this model, the Fc fragment is the stem of the *Y* and consists of the lower half of the two heavy polypeptide chains, which are joined together by one interchain disulfide bridge or more. The two Fab fragments are the prongs of the *Y*, and each consists of one entire light chain and the rest of the heavy chain, with an antigen-combining site of identical specificity at the far end. Evidence from electron micrographs supported this scheme and also suggested that the Fc unit and the two Fab units of the intact antibody molecule are joined by a hinge that allows the angle between the Fab units to vary. This flexible geometry is surely more efficient for cross-linking foreign particles than a rigid structure would be, since the distance between the combining sites on the antibody can be adjusted to fit a variety of spacings be-



Y-SHAPED ANTIBODY MOLECULES, enlarged 1,450,000 diameters, are visible in this electron micrograph of a crystal of a human immunoglobulin obtained from a patient with multiple myeloma. Interference resulting from the diffraction of the electron beam by the crystal produced a photographic negative with low resolution, but in the making of a print the resolution was enhanced by an averaging technique. Louis W. Labaw and David R. Davies of the National Institutes of Health made a brief printing exposure, shifted the paper one unit in the pattern and then reexposed it. They repeated process several times to get image shown here.

© 1976 SCIENTIFIC AMERICAN, INC

tween antigenic sites (such as the repeating protein subunits of a virus).

Amino Acid Sequences

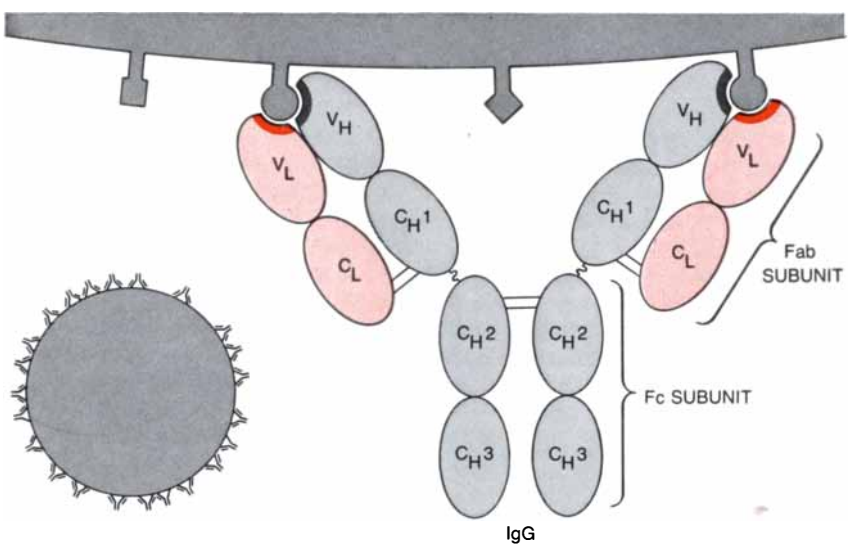
Studies of many proteins have led to the generalization that it is the sequence of amino acid units in the polypeptide chain that dictates how it folds into a three-dimensional structure and hence, in the case of an antibody, dictates the specificity of the combining site. Much about the structure of the antibody molecule can therefore be inferred from its amino acid sequence. In the case of IgG, however, there is a major technical difficulty to be overcome. Antibody molecules, even those of the same specificity, tend to be slightly different because of variations in their amino acid sequence. These differences are not large enough, however, for closely similar antibody molecules to be separated by the available methods. When that is the case, one cannot get a pure sample for determining the amino acid sequence or a crystal for analyzing the three-dimensional structure by X-ray diffraction.

The cancerous condition multiple myeloma provided the solution to the problem of adequate supplies of pure antibody. In this disease, which afflicts men and can be induced in mice, one antibody-producing cell propagates wildly and gives rise to a mass of daughter cells. A given antibody-producing cell makes antibodies of only one type, so that the myeloma cells, which are a clone of the single original cell, will all secrete an identical immunoglobulin molecule. In some cases whole IgG molecules accumulate in the serum; in others more light chains than heavy chains are produced, and the excess light chains are eliminated in the urine, either singly or in a dimer, or pair of chains. These urinary proteins were first observed by Henry Bence-Jones at Guy's Hospital in London in 1847 and have been known ever since as Bence-Jones proteins. Since Bence-Jones proteins can be obtained easily from the urine specimens of myeloma patients, without the need to draw blood, they were the first immunoglobulin components analyzed for their amino acid sequence.

The first complete amino acid sequences of Bence-Jones light chains, which consist of 214 amino acids, were determined in 1965 by Norbert Hilschmann and Lyman C. Craig of Rockefeller University and by Frank W. Putnam's group at the University of Florida. They revealed that Bence-Jones proteins from different myeloma patients had different amino acid sequences, and that the differences were mainly confined to the first half of the polypeptide chain. The light chain could therefore be divided into a variable region (amino acids 1 through 108) and a constant region (amino acids 109 through 214). Soon these observations were extended to the heavy chains, which consist of 446 amino acid units. Comparison of the sequences of heavy chains from different myeloma immunoglobulins showed that all the differences in sequence were located in 108 to 125 amino acids at one end. Thus the variable end of the heavy chain was about the same length as that of the light chain, whereas the constant end of the heavy chain was about three times longer.

The finding that the variation in amino acid sequences among IgG molecules is limited to approximately the same number of amino acid units at the same end of the heavy and light chains strongly suggested that it is in these regions that antigens are bound. The reasoning was that since this was the region in which individual IgG molecules differed from one another, it should be the region in which the specificity of the antibody is determined.

By 1968 Edelman's group had succeeded in working out the complete amino acid sequence of a myeloma IgG molecule. This sequence showed the striking similarities among segments of polypeptide chain in different parts of the molecule. The variable region of the light chain was similar in sequence to the variable region of the heavy chain. Moreover, the constant region of the heavy chain consisted of equal thirds that were similar in sequence and also resembled the constant region of the light chain. These homologies in amino acid sequence suggested that the IgG molecule is made up of compact domains, or subunits, that closely resemble one another in three-dimensional structure, a prediction that later proved to be correct.



STRUCTURE OF IMMUNOGLOBULIN G (IgG), the commonest of the five types of antibody found in the blood of higher vertebrates, is outlined in this schematic diagram. A large protein molecule, IgG is composed of two pairs of light and heavy polypeptide chains, linked by disulfide bonds into a *Y*-shaped structure. Each chain is folded into a series of globular domains; the heavy chain (*gray*) has four such domains and the light chain (*color*) two. Digestion of the IgG molecule with the protein-cleaving enzyme papain gave rise to three fragments: an Fc fragment (the stem of the *Y*-shaped molecule) and two identical Fab fragments (the prongs of the *Y*), each of which contains a combining site of identical specificity. The two domains at the tip of each Fab fragment (V_H *and* V_L) vary in amino acid sequence according to antibody specificity and are called variable domains; most of remaining domains on each chain do not vary and are called constant domains. Amino acid side chains from the two variable domains contribute to the formation of a binding cavity that is geometrically and chemically complementary to a single type of antigen, bringing about an intimate fit of the cavity with the antigen.

Hypervariable Sequences

In the early 1970's another important advance was made. When T. T. Wu and Elvin A. Kabat of the Columbia University College of Physicians and Surgeons analyzed the amino acid sequences of the variable regions in both human and animal light chains, they saw that three particular regions of the amino acid sequences varied more than the rest of the light-chain variable region. One of us (Capra, working in the laboratory of J. Michael Kehoe at the Mount Sinai School of Medicine in New York) later discovered four such "hypervariable" regions in the heavy chain. The hypervariable sequences make up about 25 of the 110 variable-region amino acid units in the light chain and about 30 of the 120 variable-region amino acid units in the heavy one. Because of the limited variability in amino acid sequence outside the hypervariable regions, many investigators predicted that the hypervariable regions alone would turn out to form the highly specific surface of the combining site. The major function of the rest of the variable region was believed to be the presentation of the hypervariable re-

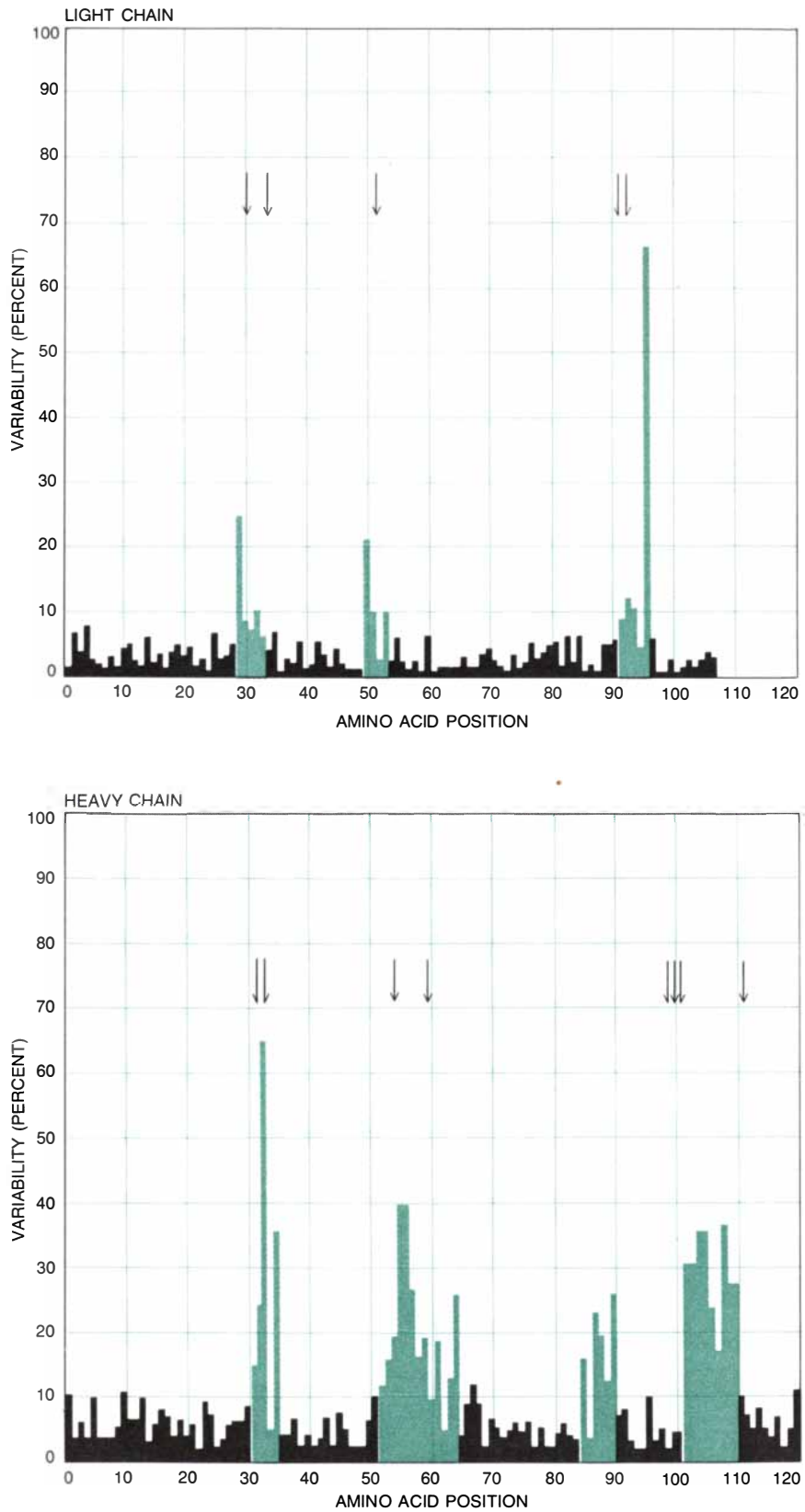
© 1976 SCIENTIFIC AMERICAN, INC

gions in three-dimensional space in such a way as to form a cavity or groove for the antigen to fit into.

This concept was supported by the affinity-labeling investigations pioneered by S. J. Singer and his collaborators at the University of California at San Diego. A common enzymological technique, affinity labeling involves attaching chemical labels to the polypeptide chain of the antibody molecule in the combining site or very close to it. The first step in the procedure is to produce antibodies against the affinity label. A small synthetic antigen, or hapten, such as a dinitrophenol group is joined to a carrier protein molecule and injected into an experimental animal. Antibody molecules are then elicited that have a special structural complementarity to the hapten, so that there is an intimate fit between one and the other. The usual hapten-antibody combination is noncovalent and reversible. To attach the affinity label the hapten is provided with a reactive chemical group, which will form a covalent bond with an amino acid unit in the combining site and so become irreversibly bound. The antibody molecule can then be separated into its constituent chains, and the position of the affinity label (which is usually made radioactive) on the chains can be determined. When such experiments were performed, the affinity labels were generally found in the hypervariable regions of both heavy and light chains, supporting the concept that the binding site is composed of hypervariable sequences contributed by both polypeptide chains.

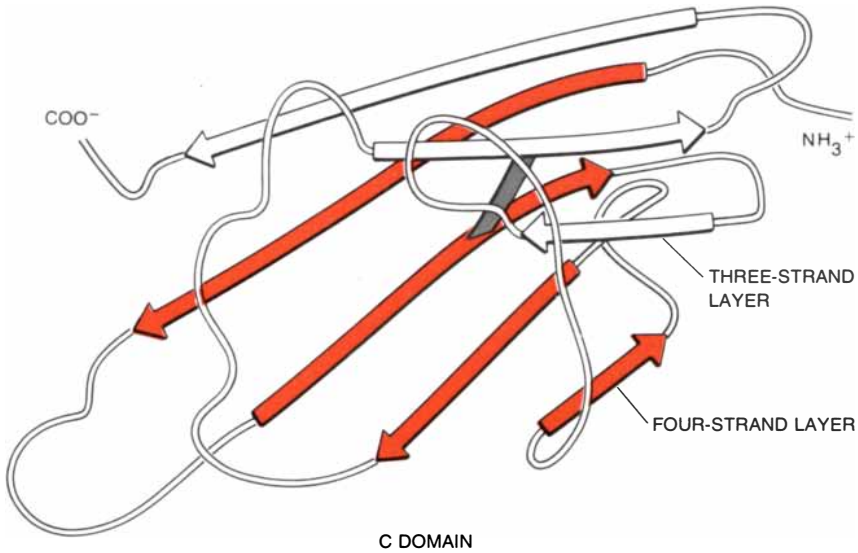
Structures of IgG Fragments

Over the past five years some components of immunoglobulin molecules have been crystallized so that their three-dimensional structure could be examined by X-ray diffraction. This work has opened the way to an understanding in three-dimensional terms of the significance of the amino acid sequences in the combining site. Roberto J. Poljak, Leon M. Amzel and their collaborators at the Johns Hopkins School of Medicine worked out the three-dimensional structure of a Fab fragment from a human myeloma protein (patient designated "New"). David R. Davies, David M. Segal, Eduardo A. Padlan and their co-workers at the National Institutes of Health determined the structure of a Fab fragment from a mouse-myeloma antibody designated McPC 603. One of us (Edmundson), in collaboration with E. E. Abola, K. R. Ely, R. L. Girling and Marianne T. Schiffer, completed the structural analysis of a Bence-Jones light-chain dimer (from a myeloma patient designated "Mcg") at the Argonne National Laboratory. All three structures were reported at the International Congress of Biochemistry in

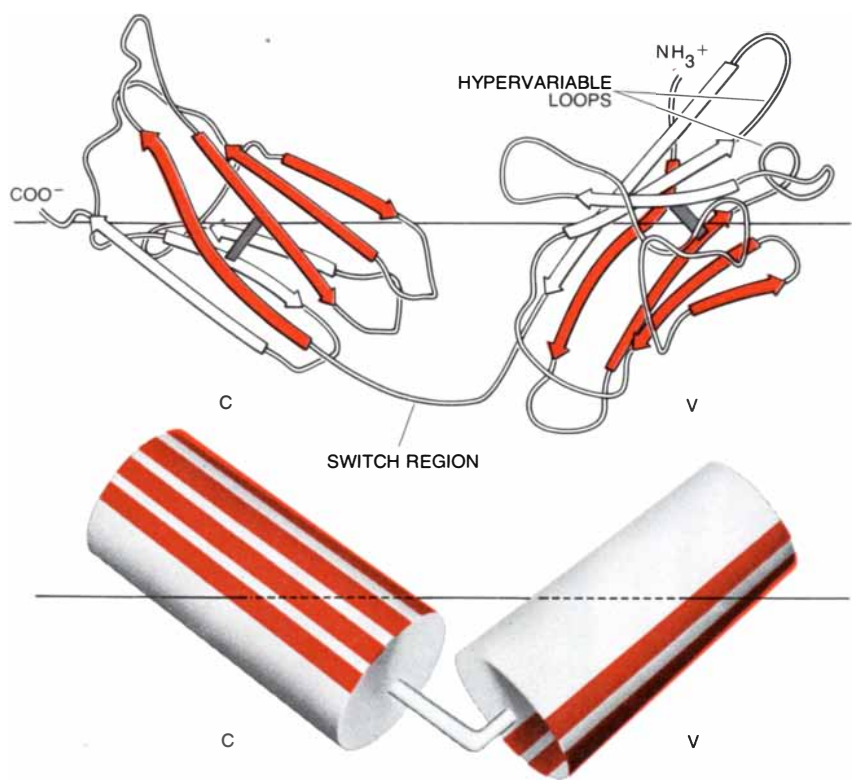


HYPERVARIABLE AMINO ACIDS (*colored bars*) in the variable domain of both the light and the heavy polypeptide chain vary considerably more from one antibody to the next than other sequences in the variable regions do, suggesting that they are responsible for the high specificity of the combining site. This hypothesis is supported by affinity-labeling experiments, in which synthetic antigens known as haptens are provided with reactive chemical groups that enable them to bind irreversibly to the combining site of an intact antibody molecule. When molecule is degraded into its component polypeptide chains, the affinity labels (*vertical arrows*) are always found to be bound to amino acid units in or near the hypervariable sequences.

© 1976 SCIENTIFIC AMERICAN, INC



BASIC FOLDING PATTERN of the polypeptide chain in the constant (*C*) domains of antibody molecule is cylindrical in shape and has a sandwichlike structure: the top layer is composed of three adjacent strands of polypeptide chain and the bottom layer of four (*colored arrows*), with the two layers held together by a disulfide (sulfur-sulfur) bridge. Although the variable domain of both the light and the heavy chain has an additional loop, the overall folding pattern of the variable and constant domains is highly similar, which was unexpected in view of the great differences in their amino acid sequences. The authors propose that this common domain structure, which has persisted in spite of the evolutionary divergence and functional specialization of constant and variable domains, may be due to presence in interior of both domain types of hydrophobic ("water-hating") amino acids that produce similar folding patterns.



LIGHT CHAIN of an IgG antibody from a myeloma patient folds into a constant (*C*) and a variable (*V*) domain. In the course of evolution the two domains have rotated with respect to each other so that their four-strand layers (*colored arrows*) face in different directions. This rotation has been accompanied by changes in the amino acid composition of the two domains that enable them to perform very different functions when they interact in pairs. For example, the association of two identical light chains in Bence-Jones dimers forms a cavity in which hapten molecules can bind. (Bence-Jones proteins are found in urine of myeloma patients and often exist as pairs, or dimers.) The dimers can thus be considered models for a primitive antibody.

Stockholm in 1973. These three-dimensional studies are continuing and have already made it possible to build precise models of the Fab fragments and the Bence-Jones dimer. They have confirmed the concept that antigens and antibodies can chemically and spatially combine in tight complexes.

The Fab fragments and the Bence-Jones dimer are similar in size, shape and general structural features. As was predicted from the homologies found in the amino acid sequences, the proteins consist of four subunits, corresponding to the variable and constant domains contributed by the light and heavy chains in the Fab fragments and the two light chains in the Bence-Jones dimer. The two domains in each chain are covalently linked by extended stretches of polypeptide chain called "switch" regions. Homologous domains in the two chains are paired, and each molecule as a whole has the shape of an elongated tetrahedron.

Domain Structure and Evolution

All four domains are cylindrical in shape and share a basic pattern of polypeptide-chain folding termed the immunoglobulin fold. In both variable and constant regions the domain structure consists of straight segments of polypeptide chain, mostly parallel to the longer axis of each subunit and arranged in two layers. The first layer is composed of three chain segments and the second layer of four chain segments. Adjacent segments within each layer run in opposite directions and are stabilized by hydrogen bonds between chains so that they take the three-dimensional form called a pleated sheet. The two layers surround a tightly packed interior of hydrophobic ("water-hating") amino acid side chains, and they are held together by a disulfide bond between polypeptide chains. Although the constant and variable regions differ substantially in amino acid sequence, the presence of key amino acids in the interior of the structure has caused the different domains to retain a common shape and folding pattern. These observations suggest that the variable and constant domains evolved as a result of the duplication of an ancestral gene, after which the duplicate genes diverged to some extent. The amino acid sequences specified by the genes thus became slightly different.

The evolutionary changes in amino acid sequence have markedly changed the way the variable and constant domains interact in pairs. In the constant domains the three-strand layers are on the outside of the molecule and the four-strand layers interact across a water-free zone to maintain the compact arrangement of the constant region. Conversely, in the variable domains the four-

© 1976 SCIENTIFIC AMERICAN, INC

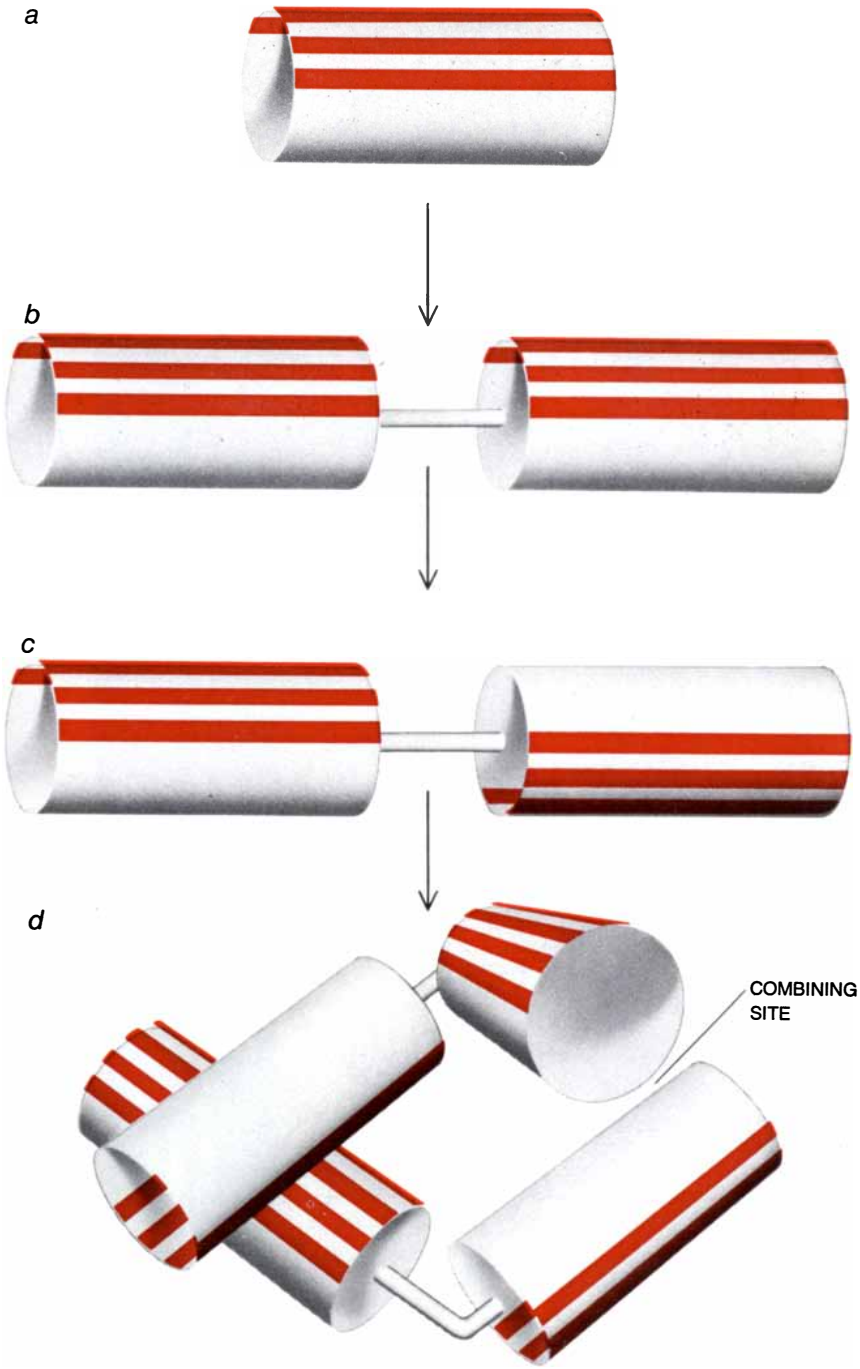
strand layers are on the outside, and the three-strand layers face each other across a water-filled channel to form the concave binding site. Thus it appears that in the course of evolution the two cylindrical domains in each chain have rotated 160 to 170 degrees with respect to each other, so that the surfaces that face each other in the variable region are different from those that face each other in the constant region. The two pairs of similarly shaped domains that make up the Fab fragments and the Bence-Jones dimer are therefore able to perform entirely different functions.

The differences in amino acid sequence between the constant and the variable domains can be correlated with their different binding interactions. For example, in the variable domains an alternating pattern of hydrophilic ("water-loving") amino acid units and hydrophobic ones in the three-chain layers is broken by the substitution of hydrophobic amino acids for the hydrophilic ones in positions important for the maintenance of the concave structure of the combining site. Similarly, the alternating hydrophilic-hydrophobic amino acid sequence in the four-chain layers of the constant domains is interrupted by the substitution of hydrophobic amino acid units for hydrophilic ones at sites of close interaction between the two constant domains. The modified sequences in each domain are in positions that are appropriate for their function but do not detract from the general structural integrity of the common immunoglobulin fold.

The fact that localized modifications in amino acid sequence have given rise to the functional differentiation of the constant and variable domains without changing their overall three-dimensional structure is a striking instance of how divergent evolution can operate at the molecular level. Indeed, the specialized interaction of paired domains appears to be a fundamental feature of antibody structure, and it allows antigen-binding regions to form even in molecules consisting of two identical light chains, as they do in the Bence-Jones dimers. X-ray crystallographic studies of the Mcg Bence-Jones dimer have shown that the spatial relations between the variable and constant domains in the two light chains are markedly different; the angle between the long axes of the two domains is about 70 degrees in one chain and about 110 degrees in the other. Comparisons of the structure of the paired light chains of the Bence-Jones dimer with that of Fab fragments indicate that the conformation of one chain resembles that of the heavy chain in Fab fragments, whereas the conformation of the other chain resembles that of the Fab light chain. Since the Bence-Jones dimer both looks and acts like a Fab fragment, we can regard it as a model for a primitive antibody.

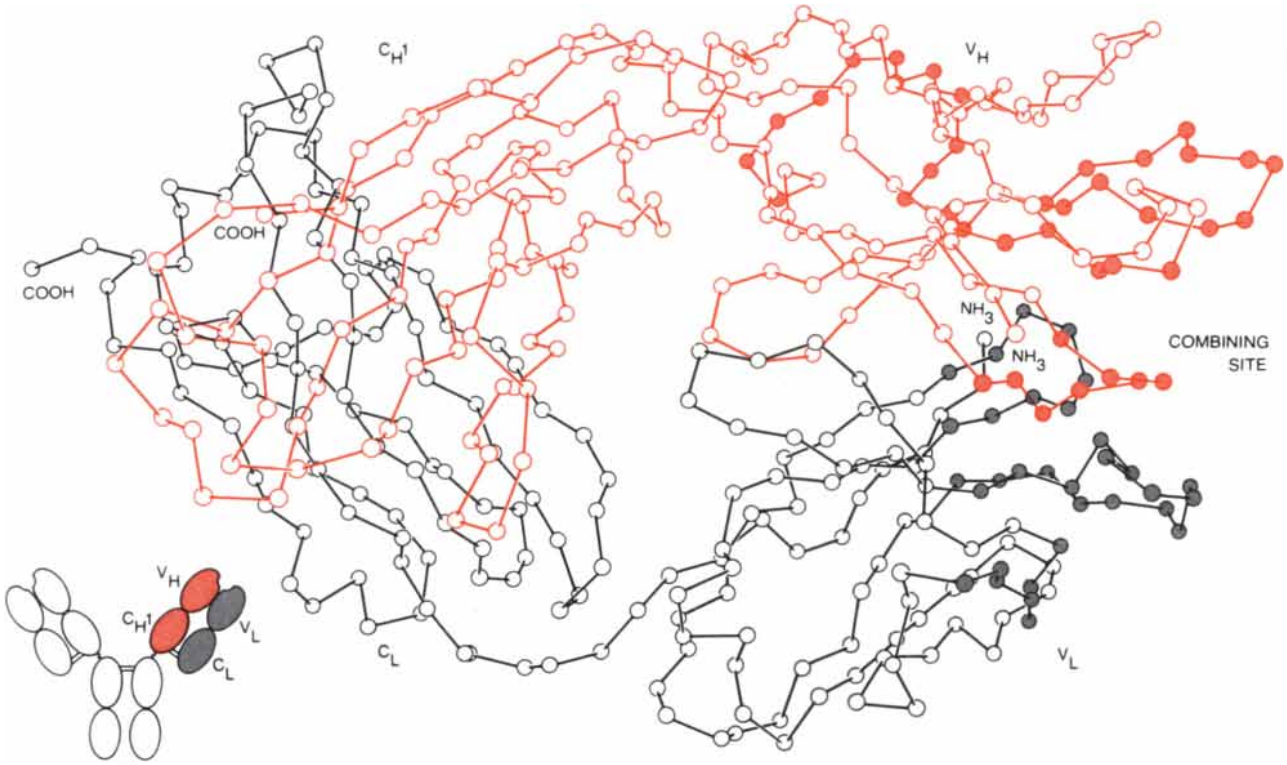
The Nature of Combining Sites

The combining sites of the different myeloma-antibody fragments studied so far differ markedly in size, shape and hapten-binding properties. For example, the human Fab fragment from patient New has a shallow groove that binds a molecule of hydroxylated vitamin K_1, whereas the mouse McPC 603 Fab fragment has a deeper cleft that binds the hapten phosphorylcholine. The Mcg Bence-Jones dimer has a still deeper and wider conical cavity connected to a pocket 17 angstroms from

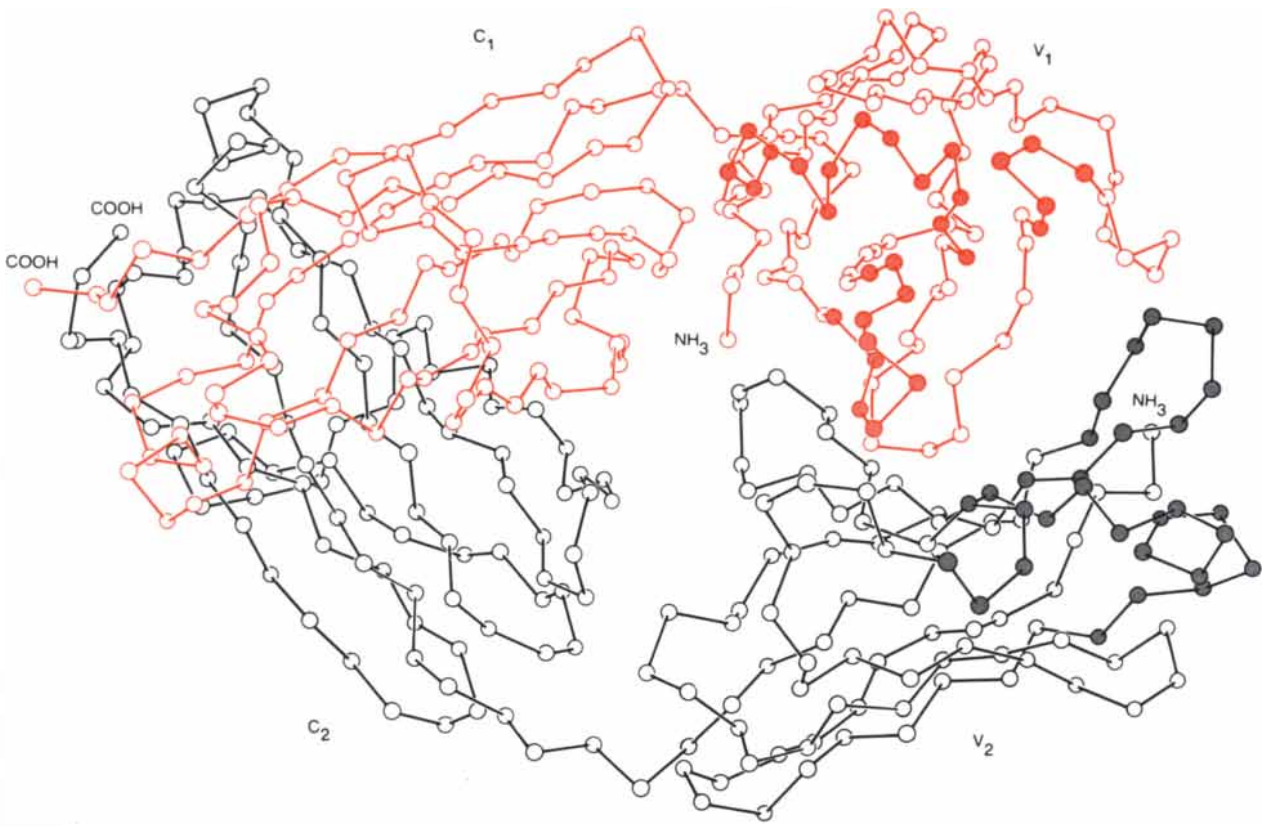


EVOLUTION OF PRIMITIVE ANTIBODY probably began with a single polypeptide chain some 110 amino acids long that folded into a cylindrical domain structure (*a*). The gene coding for this chain then duplicated, giving rise to two identical polypeptide domains joined by a short linear chain segment (*b*). Divergent evolution of the duplicate genes gradually changed the amino acid composition of the two domains, differentiating them into constant and variable regions and causing them to rotate with respect to each other when two entire chains associated (*c, d*). Resulting molecule had site between variable domains able to bind haptenlike antigens.

© 1976 SCIENTIFIC AMERICAN, INC



CARBON SKELETON OF Fab FRAGMENT (one prong of the antibody molecule) from a mouse myeloma protein designated McPC 603 shows how the entire fragment is divided into closely packed constant and variable regions composed of paired domains from the heavy chain (*color*) and the light chain (*black*) and joined by extended segments called "switch" regions. The combining site at the tip of the fragment is made up entirely of adjacent loops of hypervariable amino acid units (*filled-in balls*) contributed by both polypeptide chains.



STRUCTURE OF BENCE-JONES DIMER is strikingly similar to that of the Fab fragment (*top of page*). One of the two light-chain monomers (*color*) in dimer adopts same conformation as that of Fab heavy chain, whereas second monomer resembles the Fab light chain.

© 1976 SCIENTIFIC AMERICAN, INC

the entrance to the main cavity. When crystalline samples of this protein are compared by means of X-ray diffraction with similar samples in which dinitrophenyl or other haptenlike compounds are bound in place, it is apparent that there are three distinct binding sites, two in the main cavity and one in the pocket. Depending on the size and chemical structure of the haptenlike molecules, the sites can be occupied singly or in tandem. For example, compounds with two dinitrophenyl rings spaced an appropriate distance apart can bridge two of the binding regions with one ring in each site.

In every case the combining site is entirely made up of loops of hypervariable amino acid sequences fixed to the rigidly constructed, constant framework of the variable domains. The hypervariable loops that form the surface of the combining site are not large in the Fab fragments; they consist of at most 17 amino acid units. In the Bence-Jones dimer the binding regions continue from the hypervariable loops into more constant parts of the three-strand layers. All six hypervariable regions contribute to the formation of the main binding cavity of the Bence-Jones dimer, but only five of these regions are involved in the Fab fragments: the first, second and fourth hypervariable regions of the heavy chain and the first and third hypervariable regions of the light chain. Unique patterns of amino acid side chains on these loops come together to form a continuous surface capable of providing complementarity with a given antigen. The combining site of an immunoglobulin molecule is thus analogous to the die in a metal-stamping press in that it can differ from one antibody to the next without any change in the overall structure of the molecule.

Specificity-determining Factors

What are the factors determining the high specificity of a combining site for a hapten or an antigen? First, the shape of the site is complementary to that of the antigen, and it is concave in order to expose a large area of surface to the antigen. Second, the amino acid side chains within the cavity are precisely positioned to take full advantage of electrostatic, hydrogen-bonding and van der Waals interactions, all of which are weaker than covalent interactions. (In this regard the binding forces between an antibody molecule and an antigen resemble those between an enzyme molecule and the substance on which it acts.)

For example, the binding site for the hapten phosphorylcholine in the mouse McPC 603 Fab fragment is a wedge-shaped cavity, the walls of which are lined with hypervariable amino acid units. The phosphorylcholine occupies only a small part of the cavity and is

The Affordable Programmable

Never Before A Fully Programmable Calculator Available To The Scientific Community At Such A Low Cost!

by Mail Order Only $59.95

SUMMARY OF SPECIFICATIONS:

- **12 DIGIT RED L.E.D. DISPLAY – Display entries or results in 3 modes, scientific, fixed point or engineering**
- **PROGRAMMABLE – 72 Keystroke program storage**
- **TRIG FUNCTIONS – Calculates in radians or degrees or grads**
- **HYPERBOLIC FUNCTIONS**
- **Ln, log, e^x, 10^x**
- **POWERS Y^X: Roots $\sqrt[X]{Y}$**
- **FUNCTIONS OF X – 1/x, $\sqrt{\ }$ $\overline{X}$**
- **LINEAR REGRESSION: TREND ANALYSIS: SLOPE AND INTERCEPT**
- **STATISTICAL: $\Sigma+$, $\Sigma-$, averages, standard deviation, variance, permutations, combinations**
- **PERCENTAGE CALCULATION – addon/discount/yield△percentage**
- **π AND CHANGE SIGN KEY**
- **10 MEMORIES – 7 functions – M+, M–, MR, M×, M÷, STORE, 9M**
- **DEGREES/DEGREE, MINUTE, SECONDS CONVERSION**
- **RADIAN/DEGREE CONVERSION**
- **POLAR/RECTANGULAR CONVERSION**
- **8 METRIC CONVERSIONS**
- **SPHERICAL/CARTESIAN CONVERSIONS**

The Commodore PR (Personal Resource) 100 – Allows You To Personally Program According To Your Own Problem-Solving Needs!

WHY A PROGRAMMABLE?
The Commodore PR-100 Programmable Calculator introduces a new, innovative dimension in highly sophisticated mathematical and scientific problem-solving. It decentralizes and personalizes the decision-making power of the computer, bringing to you what was never before available in a programmable calculator with such features, yet at an amazingly affordable price.

Method of Entry: The Commodore PR-100 is designed for "Think and Touch" operation utilizing an algebraic mode of entry. You can enter equations which have up to **4 levels of parenthesis.** Add to these features the **10 memory registers,** and you can work the most complicated mathematical, scientific, engineering, business, statistical and combinational functions. Furthermore, you can also perform unit conversions: metric to English and English to metric.

Keyboard: The keyboard consists of 40 keys and 2 switches. 35 keys are for number and function entry. 10 of these keys are dual function (shifted keys).

SPECIALIZED FUNCTIONS:
Most of the important functions found on large scale computer systems are finally now available to you on the ultimate hand-held programmable calculator: iterative and recursive problem solving techniques; looping; conditional and unconditional branching.

ADVANTAGES OF THIS PROGRAMMABLE:
Consider for a moment the advantages of the Commodore PR-100. In terms of increased productivity you can now achieve the capability of: optimizing mathematical and scientific models; making trend and risk analyses; projecting and forecasting more accurately; performing statistical reductions; automating time-consuming "number-crunching."

The PR-100 is also a pre-programmed calculator that can achieve a multitude of functions: Basic + – × ÷ and = functions; clear entry or the entire calculation; Hyperbolic functions include: sinh, cosh, arc cosh and arc tan; Trigonometric functions include: sin x, cos x, tan x, arc sin x, arc cos x, and arc tan x. Calculates in radians, degrees or gradians. Enter x 0° to 360° (0 to 2π) or multiples of 360° (2π). Natural Log (Ln x), Common Log (log x), Exponential (e^x), Antilog (10^x); Algebraic Functions include: x, x^2, 1/x. Powers: (Y^x); π (pi). Change Sign (+/–); x↔y exchange function; Parenthesis ( ) enter equations that contain up to 4 levels of parenthesis without using temporary or intermediate storage.

Memory Functions: The PR-100 has 10 separate memory registers. For each memory there are 7 operations that can be performed: memory storage and clear, recall, addition, subtraction, multiplication, division and performs the desired functions on each memory.

NO LANGUAGE TO LEARN
Truly this programmable calculator is a powerful personal scientific and mathematical computing resource. And you don't need to know programming to make it work for you. There is no special language to learn. The entry system is so easy and flexible to use that you can apply it quickly to your own personal problem-solving techniques and style.

Special Functions of the PR-100: To allow you greater flexibility in programming, there are 3 special functions: 1. HALT – The program will halt and allow the user to read or enter data. 2. GOTO – Allows branching from one section of a program to another. 3. SKIP – Allows conditional flow of a program on an intermediate result; that is – the program goes one way if an answer is negative and another way is positive.

Loading, Editing and Modifying Programs: You can single step through a program to check it out, backstep through a program, easily read entered steps and modify a single step of a program.

Now personal programming is here. A step-function increase in capability over sophisticated slide-rule calculators. Capability you won't fully discover until you own a PR-100 yourself. Fill out the coupon below or call our toll free number for your two week trial.
The Commodore PR-100 is powered by rechargeable NiCad batteries and is furnished with an AC adapter/charger, leatherette carrying case and full instruction booklet. One year manufacturer's warranty.
Dimensions. 3¼" x 6½" x 1⅛" Weight: 4 oz.

Compute With It For Two Weeks – No Obligation
Please send _____ Commodore PR-100 Calculator(s) at the introductory price of only $59.95 (plus $3.95 shipping and insurance) each. If not completely satisfied, I can return it within two weeks for a prompt refund

☐ Instead, I/we would like to take advantage of your quantity price offering of $54.95 (plus $3.95 shipping and insurance) each. OFFER GOOD ONLY ON QUANTITY ORDERS OF 6 UNITS OR MORE! You can save $30.00!

☐ Check or M.O. Enclosed (Ill. residents add 5% sales tax)
☐ Please charge my credit card checked below:
☐ American Express ☐ BankAmericard ☐ Carte Blanche
☐ Diners Club ☐ Master Charge

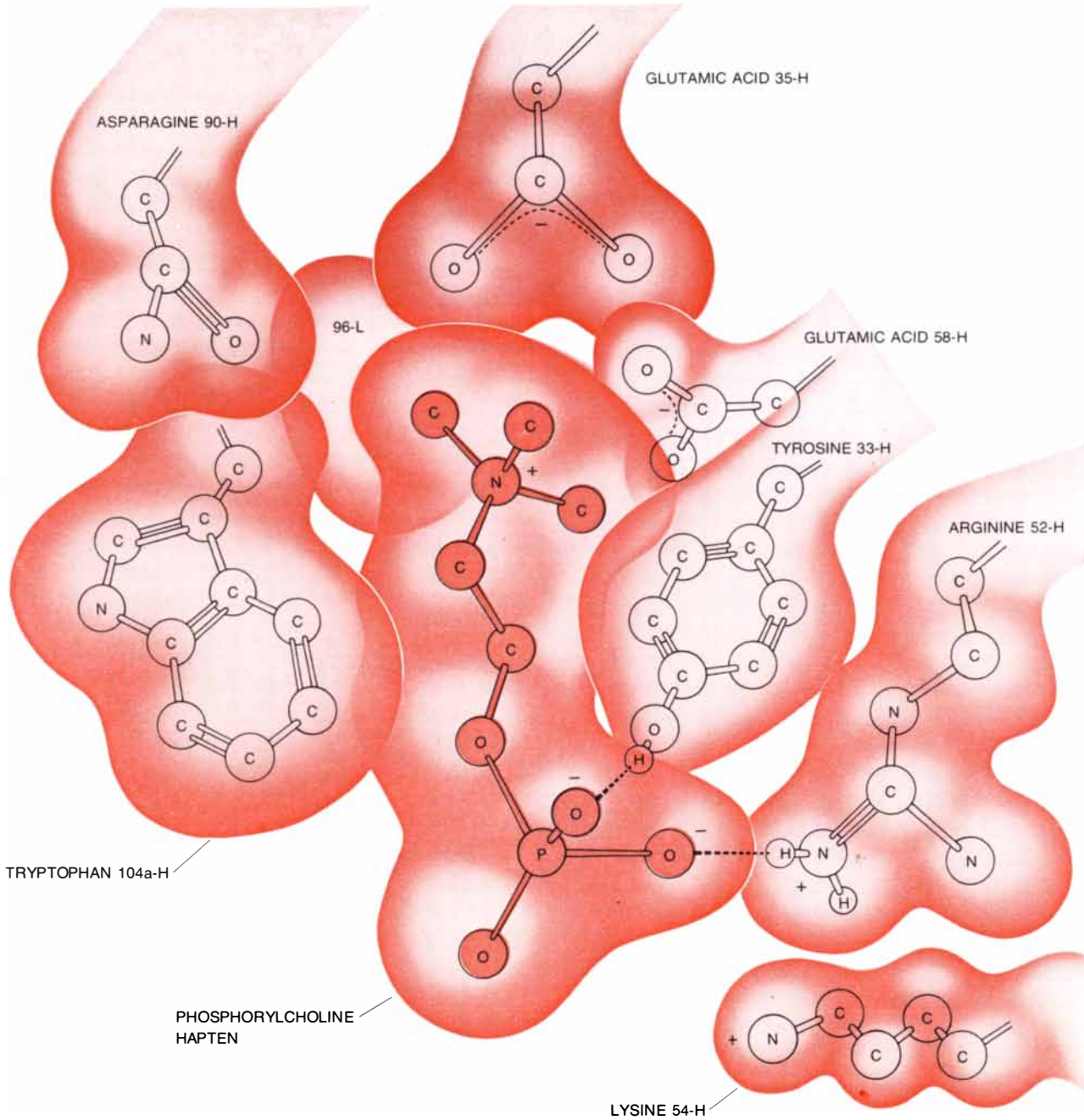
Credit Card # _____
Master Charge Bank # _____ Exp. Date _____
Name _____
Address _____
City _____ State _____ Zip _____
Signature _____
SCA-1/77

Send Coupon To
Contemporary Marketing, Inc.
790 Maple Lane, Bensenville, Ill. 60106
Call Toll Free: 800-621-8318 Ill. call: 312-595-0461

© 1976 SCIENTIFIC AMERICAN, INC

bound asymmetrically, being closer to the heavy chain than it is to the light one. The negatively charged phosphate group (PO_4) of the hapten is within hydrogen-bonding distance of the hydroxyl (OH) group in the combining site. It is also in close proximity to the positively charged guanidinium group of an arginine unit, and this proximity leads to a strong electrostatic interaction. The choline group of the hapten also appears to interact with hypervariable loops from both the heavy and the light chains. The positively charged nitrogen of the choline group is close to the negatively charged side chain of a glutamic acid unit, and there are van der Waals interactions between the choline group and the carbon atoms in both the heavy and the light chains.

These relations have been worked out in X-ray-diffraction analyses of crystalline proteins. One must bear in mind that the binding of a small hapten molecule to part of the combining site in a crystalline antibody is a highly artificial situation; in real life the antigen probably fills the binding site completely. Moreover, the hypervariable loops, particularly those that protrude into the water surrounding the antibody molecule, appear to be flexible, and the binding site may expand to enfold a larger area of antigen surface. The crystal studies of the Bence-Jones dimer show that the surfaces immediately around the binding cavity but not inside it can be approached sufficiently closely by large molecules to allow even more interactions. In the crystals these interactions hold a second Bence-Jones protein molecule in such a specific orientation that one of its amino acid side chains enters the cavity of the dimer and lodges in a site appropriate for antigen binding.



MODEL OF THE BINDING of the hapten phosphorylcholine (*colored balls*) to combining site of McPC 603 antibody illustrates how the shape of the site is precisely complementary to that of the hapten. In addition amino acid side chains lining cavity bind to hapten through weak, noncovalent interactions of following types: electrostatic (attraction of positively and negatively charged atoms), hydrogen-bonding (attraction of oppositely polarized hydrogen and oxygen atoms) and van der Waals (attraction between electrons and atomic nuclei).

© 1976 SCIENTIFIC AMERICAN, INC

As we have mentioned, the variable and constant domains in each polypeptide chain of the Fab fragments are covalently linked by stretches of chain called switch regions. In the Bence-Jones dimer this region is also bent to a larger extent in the heavy-chain analogue than in the light-chain one. R. Huber and his colleagues in Germany, however, have recently described an IgG protein in which the switch regions in the Fab parts of the molecule do not show this large difference between heavy and light chains.

These observations suggest that both the heavy and the light chains of some antibodies are capable of large hinge-like movements at the switch regions, but that these movements probably do not occur in all IgG's. Nevertheless, the capacity for conformational changes in such connecting segments may be very important in the mechanism by which the binding of antigen to certain antibodies initiates a variety of immunological responses, including the dissolution of blood cells by the blood factor known as complement, the release of histamine from mast cells and the activation of the synthesis of the antibodies themselves.

The Modification of Specificity

The study of precise structural models derived from X-ray-diffraction analyses has done much to suggest how the geometry of the antibody combining site, and hence its specificity for antigen, might be modified by making changes in the amino acid sequence of the hypervariable regions. For example, the simple substitution of a negatively charged glutamic acid unit for a positively charged arginine unit in a polypeptide chain of the combining site would probably not affect the overall topology of the hypervariable surface, but it would dramatically change its binding properties. In addition insertions and deletions within the hypervariable loops would alter their folding, and they could change the shape and size of the hypervariable cavity. Finally, there is the possibility that substitutions, deletions or insertions of amino acid units in the deeper parts of the interface between the variable domains might alter the interactions between the two domains, resulting in a shift in the relative positions of the hypervariable loops on the heavy and light chains and giving rise to significant changes in specificity.

In sum, the dramatic advances of immunochemistry in recent years have drawn a much clearer picture of how an antibody molecule fits itself to an antigen, and have provided much information on the evolution of antibody structure. A rational molecular explanation for the capacity of the body's immune system to respond to an apparently infinite variety of antigenic stimuli now seems to be at hand.

Garden Camera CALCULATORS

Hewlett Packard

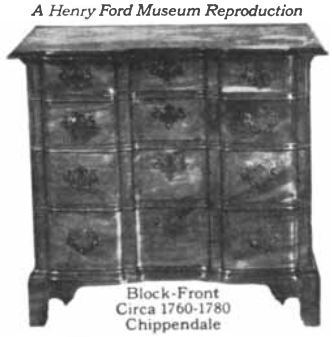
HP 25C..... 174.50
HP 25..... 124.50
HP 21..... 69.50
HP 22..... 114.50
HP 27..... 154.50
HP 55..... 249.50
HP 65..... 650.00
HP 91..... 384.50

Texas Instruments

SR 52.....217.50
SR 56..... 83.50
SR 50A..... 52.50
SR 51 II..... 59.95
T.I. 30..... 22.95
T.I. 40..... 32.95
PC 100..... 219.50
Business Analyst..... 39.50
The Little Professor.... 17.50

Speed your order **TOLL FREE!**
Call 1 (800) 223-5830
Or send postage and handling to
GARDEN CAMERA
345 Seventh Ave., N.Y., N.Y. 10001
NY. AK. HAW call (212) 868-1420
Master Charge Accepted.

A Henry Ford Museum Reproduction



Block-Front
Circa 1760-1780
Chippendale

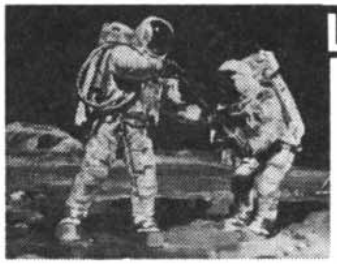
Ready to Assemble and Finish 18th Century Furniture

One of 24 Bartley classics, in hand crafted solid mahogany, oak or cherry. Totally authentic in design and beautifully constructed. Each kit is easily assembled and finished in your own home without tools. All pieces also offered completely assembled and hand finished. A $5.00 coupon included with catalogue.

Please send me your illustrated catalogue of 18th century antique furniture reproductions available in kit form or hand finished. I am enclosing $1.00 to cover postage and handling.

Name _____
Address _____
City/State/Zip _____

The Bartley Collection, Ltd.
747 Oakwood Ave., Dept. SA2 Lake Forest, Il 60045



LAST CHANCE TO BUY

APOLLO, SKYLAB
Slides and Movies

35mm Slides & 8mm Movies taken by astronauts on Apollo moon flights will only be available for a LIMITED TIME.

MOVIES & SLIDES MADE FROM NASA NEGATIVES.

A. Apollo 17, set of 24 Color Slides $10
B. Apollo 16, set of 24 Color Slides $10
C. Apollo 15, set of 48 Color Slides $18
D. Apollo 14, set of 48 Color Slides $18
E. Apollo 11, set of 24 Color Slides $10
F. Apollo 17, 400 Ft. Color Movie $40
G. Apollo 16, 200 Ft. Color Movie $20
H. Apollo 15, 400 Ft. Color Movie $40
I. Apollo 14, 400 Ft. Color Movie $40
J. Apollo 11, 450 Ft. Color Movie $50
K. Skylab 4, set of 60 Color Slides $25
L. Skylab 4, 400 Ft. Color Movie $40

APOLLO-SOYUZ FLIGHT
400-foot COLOR Movie **$40**

48 COLOR Slides, best of thousands by Apollo-Soyuz astronauts **$20**

"FOOTPRINTS ON THE MOON"
Spectacular 215 page HARD COVER Book by AP with 142 Full Color Photographs; many taken in space and on the moon.
ONLY **$5.95** POSTPAID

MOVIE NEWSREELS, 6269 Selma Avenue
Box 2589, Hollywood, Calif. 90028

SA

When Ordering 8mm Movies Please Specify ☐ Reg. 8 or ☐ Super 8.
Foreign Orders add 10% For AIRMAIL Postage.

☐ Enclosed find $_____ send following _____

Name _____ Address _____

City, State, Zip _____

OR-ORDER BY 24 HR TEL: (213) 467-2448 and charge to YOUR DINERS OR MASTER CHARGE CARD.

© 1976 SCIENTIFIC AMERICAN, INC

The Perception of Moving Targets

Highly specialized nerve circuits enable the human visual system to separate information about the direction in which an object is moving from the details of the object's pattern

by Robert Sekuler and Eugene Levinson

The ability to respond promptly and appropriately to a moving object is often a matter of life or death. Hence it is not surprising that the mammalian visual system has evolved highly specialized neural mechanisms to handle the job. Physiological experiments in which microelectrodes are implanted in the visual cortex of an animal's brain have provided investigators with a wealth of information about how this complex nerve circuitry works in various species. In order to study the corresponding function in man, however, less invasive techniques are obviously required. Although the procedures developed for the latter purpose are necessarily indirect, they have yielded some striking information about the mechanisms employed by the human visual system to perceive moving objects.

The first requirement of any such psychophysical procedure is the provision of special stimuli. The moving things one sees in everyday life are not very helpful here, because their impact on the visual system is both too complicated and too weak to be measured reliably. Therefore even though the ultimate objective is of course to understand why moving objects appear the way they do in everyday life, one must start by working with moving targets that are both less complicated and more powerful than ordinary objects.

As far back as early in the 19th century students of human visual perception realized that prolonged exposure to a simple, strong stimulus could reveal properties of the motion-perception system that could not be observed with more variable stimuli. The British investigator R. Addams, for example, noted that prolonged fixation on motion in one direction often resulted in a peculiar aftereffect. Addams stared at a waterfall near Loch Ness for several seconds. When he then shifted his gaze to the rocks next to the waterfall, the rocks seemed to move upward. Earlier the Czech physiologist Jan Purkinje had experienced a similar illusion after watching a parade of horseback riders for about an hour.

In our laboratory at Northwestern University we obtain comparable effects by means of electronically generated stimuli. One kind of stimulus we find useful consists of a moving pattern of small, bright dots "painted" on the screen of a cathode-ray tube with the aid of a computer. The dots are usually distributed across the screen in a random fashion. Opposite sides of the screen are functionally connected: when a dot disappears off one side of the screen, it immediately reappears at the other. From the point of view of an observer it is as though he were looking at an infinitely large sheet of dots moving behind a stationary aperture. Because the dots are distributed randomly on the sheet, they enable us to study the visual impact of their motion without the complicating effects of perceived contours.

Another suitable kind of target, also created on a cathode-ray tube, is a pattern of alternating dark and light parallel bars. Patterns of this type, called gratings, can be made to drift at any speed in a direction perpendicular to the bars. With either gratings or dot patterns we are able to generate movement that is continuous, repetitive and of any desired duration. The individual, momentary effect of the motion can be allowed to accumulate, producing a total impact large enough to be measured reliably.

For any target, including a moving one, the response of the visual system depends on the contrast between the dark and light parts of the image. The fact that a television set has separate controls for contrast and brightness serves as a reminder that the two variables are not the same. At any level of average brightness one can independently adjust the contrast, making the picture more or less "washed out." In order to measure how effectively a certain television image stimulates the visual system, one possible approach is to adjust the contrast until an observer can barely see that the television screen is not perfectly uniform in its luminance. If the particular image on the screen happens to be one to which the observer is very sensitive, it will remain visible even at a very low contrast.

It follows that the contrast threshold at which a target image disappears can be utilized as an index of an observer's sensitivity to that particular target. A low contrast threshold would mean a high sensitivity and vice versa. The same approach can be taken with an electronically generated grating or dot pattern: by reducing the contrast until the observer can barely see the pattern one can obtain a measure of the observer's contrast threshold for that pattern.

Ulker T. Keesey of the University of Wisconsin and Janos Kulikowski of the University of Manchester have shown that for many moving targets observers have not one but two distinct contrast thresholds. Suppose, for example, one is working with a low-frequency grating (that is, one with wide bars) moving across the screen of a cathode-ray tube. One starts with the contrast at zero and then gradually increases it until the observer can just detect that something is present. The observer does not actually see the grating; what he perceives is a procession of formless ripples. Although he can determine the direction in which the ripples are moving, the bars of the grating are themselves indistinct. If one now increases the contrast further, the bars too become visible. Evidently at a very low contrast a target's motion may be visible even though the characteristics of its pattern are not, whereas at a somewhat higher contrast both the pattern and its motion are discernible. The existence of separate thresholds for the perception of motion and the perception of pattern leads quite naturally to the suspicion that within the human nervous system there are two separate visual channels for analyzing these two aspects of a moving target.

Obviously there are limits to how much one can learn about the nervous system's response to motion from just watching a moving target. More analytic tools are needed. One such tool, frequently employed to study smell, hearing and taste as well as vision, is the psychophysical procedure known as

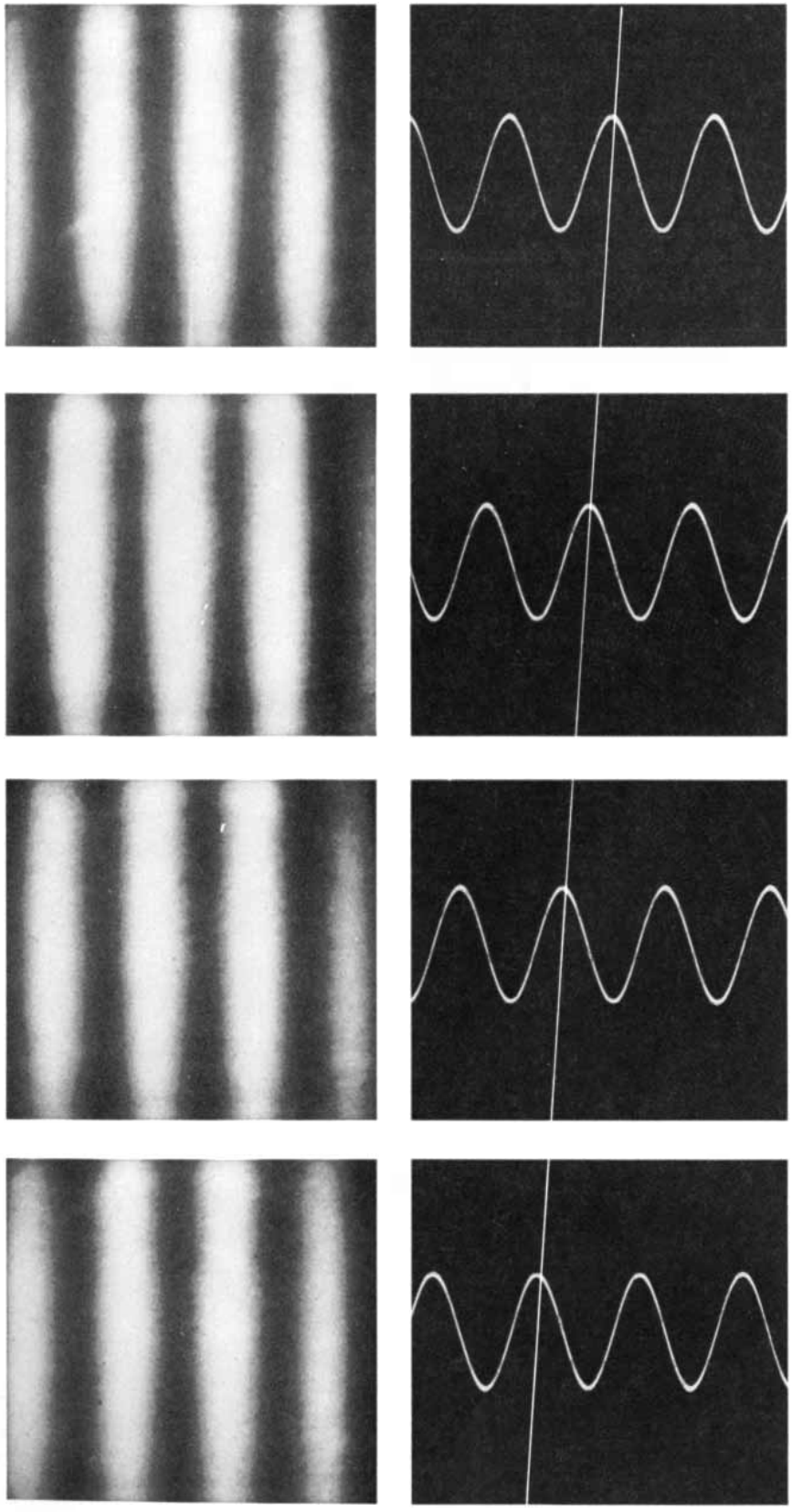
© 1976 SCIENTIFIC AMERICAN, INC

selective adaptation. This technique makes it possible to "dissect" any sensory system painlessly. In selective adaptation one compares the detection of various stimuli before and after exposing the sensory system to a strong level of a particular stimulus, termed the adapting stimulus. Imagine, say, that exposure to some adapting stimulus makes it hard to detect one test stimulus but has little or no effect on the ability to detect another test stimulus. It would be reasonable to conclude that the mechanisms needed to detect the first test stimulus overlap the mechanisms affected by the adapting stimulus but that the mechanisms needed to detect the second test stimulus do not. The adapting stimulus is usually thought of as fatiguing, or reducing the sensitivity of, selected mechanisms in the nervous system; reduced sensitivity makes it more difficult to detect the stimuli that are processed by the affected mechanism.

About a dozen years ago selective-adaptation experiments established that at some stage of neural processing the human visual system makes use of mechanisms that are directionally selective. Leo Ganz, who is now at Stanford University, and one of us (Sekuler) had observers view vertical gratings that moved either to the left or to the right. The adapting stimuli were of high contrast; they were designed to selectively desensitize mechanisms that responded to only one direction of motion or the other, if indeed such mechanisms existed. After a short adapting period con-

"WHITE SPIRAL," one of a set of 12 rotating optical illusions called Rotoreliefs designed by the artist Marcel Duchamp in 1935, demonstrates an important property of the human visual system: the independence of the neural channels responsible for perceiving opposite directions of motion. When a copy of the design is rotated clockwise at a speed of about 33 revolutions per minute (on a phonograph turntable, say), a viewer should be able to see two superposed spirals: a dark one with many narrow turns, which appears to drift toward the center of rotation, and a white one with fewer, broader turns, which seems to move outward from the center. The two interwoven spiral patterns are actually fixed with respect to each other; nevertheless, when they are viewed in rotation (preferably with one eye), the two directions of motion separate clearly. This reproduction is from a set of Duchamp Rotoreliefs in the collection of Anne d'Harnoncourt.

© 1976 SCIENTIFIC AMERICAN, INC



MOVING PATTERN of alternating dark and light parallel bars can be created on the screen of a cathode-ray tube in order to provide a suitable target for psychophysical experiments with human observers. In the sequence of photographs at the left, made in the authors' laboratory, a grating is shown drifting to the left across the screen at a constant rate. Sinusoidal curve associated with each photograph is a plot of light intensity across the screen at the moment of each exposure. Long slanted line connects a given peak on each curve at successive intervals.

trast thresholds were measured for a grating that moved either in the same direction as the adapting stimulus or in the opposite direction. There was a significant difference between the visibility of the two opposite directions of test movement: after adaptation it was much harder to see the grating that moved in the same direction as the adapting grating. This effect—called direction-specific adaptation—has been re-created many times since.

Naturally we wondered whether the mechanisms we were dealing with were related to single neural cells that show a strong preference for direction of motion. A directionally selective neuron discharges strongly when a properly oriented stimulus drifts in one direction through its visual field; it fires at a much reduced rate if the same stimulus moves through the field in the opposite direction. Direction-specific cells are abundant in the visual cortex of both the monkey and the cat. Because this kind of neuron is able to discriminate between directions of motion, it could provide the basis for the human ability to see moving targets. Although it is not certain that the human brain has exactly the same kinds of cells as a monkey's brain, it is known that in areas of vision not directly concerned with motion man's visual abilities are nearly identical with those of certain monkeys. Moreover, monkeys appear to experience motion aftereffects in much the same way that human beings do.

Several investigators have tried to determine whether or not the electrical activity of the human brain also shows directional selectivity. Electrodes can be fastened to an observer's scalp, and after proper amplification and signal processing the massed response from the brain tissue lying under the electrodes can be measured. The direction-specific adaptation experiment can then be done with the massed brain response. For example, one investigator measured the brain's response to a pattern moving first in one direction and then in the opposite direction. Adaptation produced a change in the brain's response to motion in the same direction as the adapting motion; the response to motion in the opposite direction was not affected.

Selective-adaptation experiments, in short, provide evidence that the human visual system contains some direction-selective mechanisms. The directional selectivity shown with adaptation techniques, however, does not tell one everything one would like to know about how the system responds to a moving stimulus. For example, the kind of selectivity observed in these experiments does not necessarily mean that there are separate detectors for opposite directions of motion. Direction-specific adaptation might reflect activity in directionally selective neurons whose outputs are subsequently pooled at a detection stage

© 1976 SCIENTIFIC AMERICAN, INC

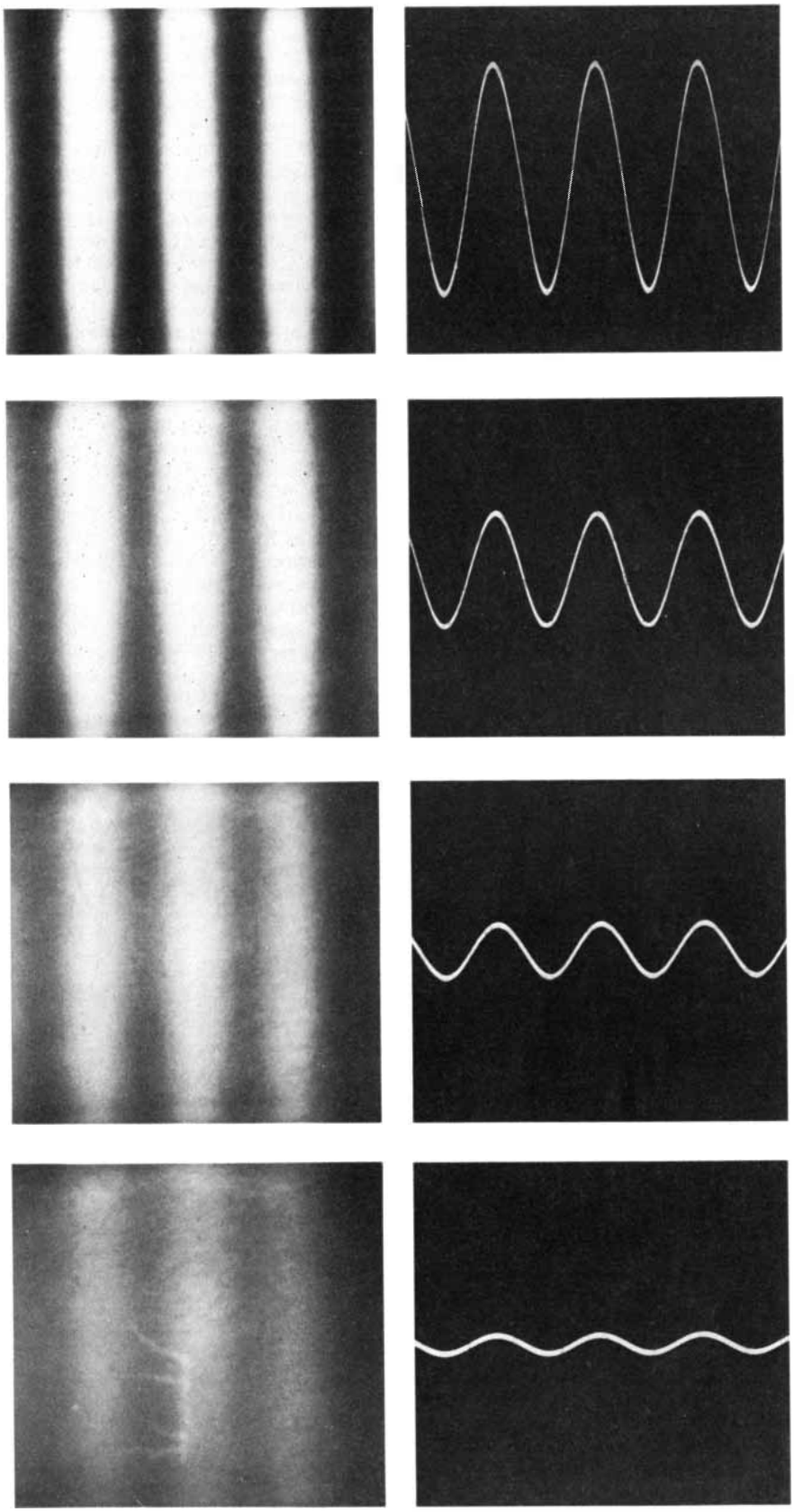
that is not direction-specific. Adaptation experiments confirm that the visual system analyzes direction of motion, but one would like to know how the information is treated after this preliminary analysis.

To test any system's ability to respond independently to either of two inputs, an engineer might examine how that system responds to the sum of the inputs. The basic idea is to compare the system's response to some stimulus alone with the system's response to the same stimulus when it is accompanied by an additional stimulus. If the presence of the additional stimulus neither helps nor hinders one's ability to detect the original target, the two stimuli would have to be detected by independent mechanisms.

We have exploited this approach in the following way. Starting with two separate cathode-ray tubes, on one we generate a grating that drifts to the right while on the other we produce a similar grating that drifts to the left. The images of the two tubes are combined optically (with a half-silvered mirror) and viewed by an observer, who adjusts the contrast of the combination grating to the threshold level.

Although we have made many measurements with the aid of this basic procedure, there is one special condition that is of particular interest: the case in which both gratings have the same bar width and contrast, drift at the same speed but move in opposite directions. In this case the combination stimulus shows neither direction of drift; it simply reverses repetitively in phase! At any one point on the observer's retina the combination grating changes from light to dark and back again, without any obvious drift in either direction. In fact, when an observer takes a quick look at such a high-contrast combination grating, he sees no movement; the grating simply appears to pulsate, with the dark bars getting light and vice versa.

Suppose now that an observer views one of the moving component gratings, adjusting its contrast until something is barely visible. He then repeats this procedure with the other moving component grating, also presented alone. The two thresholds will appear at about the same contrast value. Next the observer views the combination grating, always keeping the contrast of each moving component at its own threshold value. If gratings moving in opposite directions are detected independently, the combination grating would also be exactly at threshold. In order for the combination grating to be visible at all each of its oppositely moving components would require enough contrast to be visible in its own right. The actual appearance of the combination grating at the threshold level suggests that this line of reasoning is correct: when the combination is just



CONTRAST THRESHOLD, the level of contrast between the dark and light parts of a target image at which the image can barely be seen, serves as an index of an observer's sensitivity to that target. In the sequence of photographs at the left, for example, the contrast of four otherwise identical stationary gratings is progressively reduced from top to bottom. For moving targets human observers have been found to have two distinct contrast thresholds: one for perception of motion and the other for perception of pattern. Curves at right indicate light intensity.

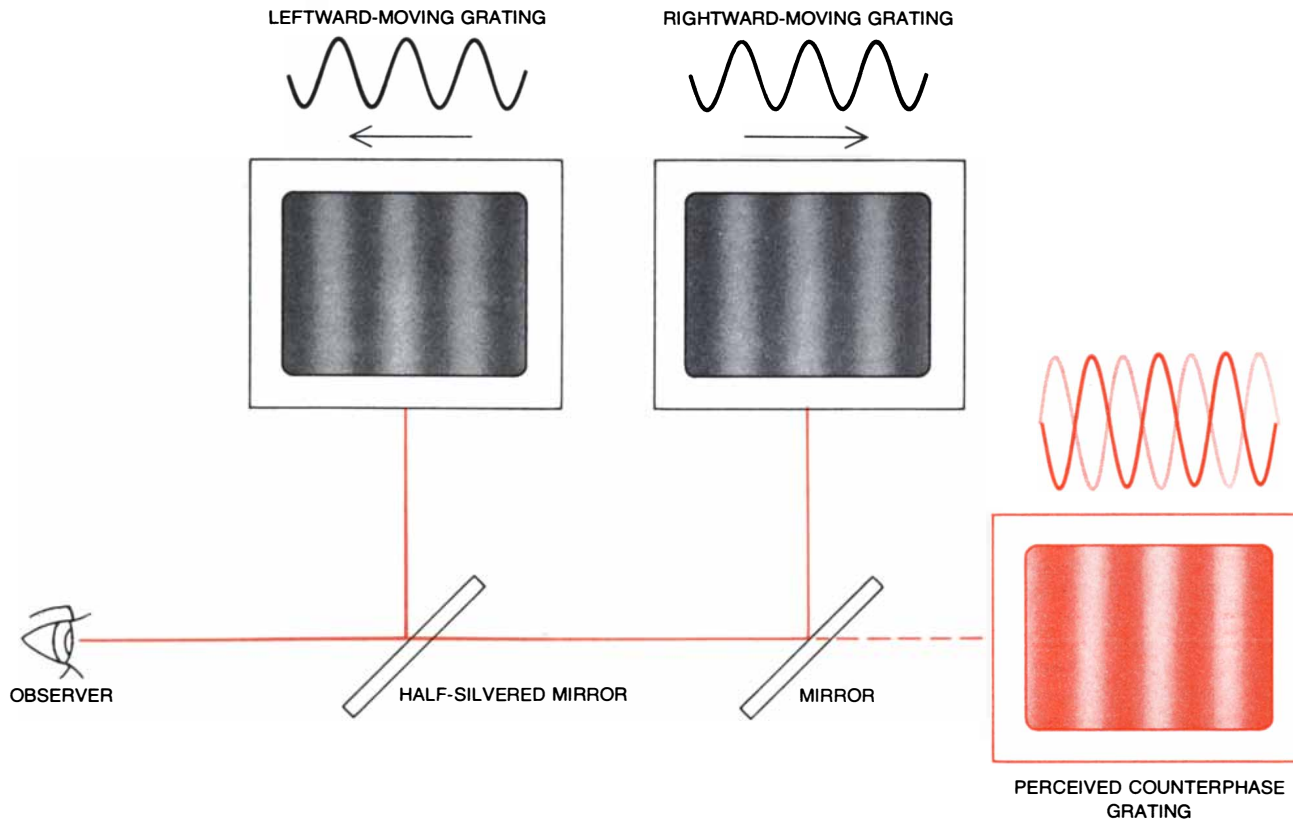
© 1976 SCIENTIFIC AMERICAN, INC

at threshold, the observer sees first one and then the other component. The detection of the combination stimulus seems to depend on two independent visual mechanisms, each one responsible for detecting just one directional component, and these mechanisms fluctuate slightly in sensitivity over time.

The possibility that oppositely moving components in a combination grating are detected by independent direction-sensitive mechanisms leads to an interesting prediction: The contrast threshold for detecting a combination grating should be exactly twice the contrast threshold for detecting one of the moving components presented by itself. One expects this twofold difference because, as we have already explained, if either component of the combination grating is at the contrast threshold, then both components are at the contrast threshold, giving the combination grating as a whole a total contrast of two threshold units. Accordingly in our experiments we have compared the contrast threshold for a combination grating with the contrast threshold for either of its moving grating components presented alone, repeating the measurements at different bar widths and drift rates. The results are unequivocal: twice as much contrast is needed to detect a combination grating as is required to detect one of the moving grating components of the combination stimulus.

Since the combination grating made of oppositely moving components of equal contrast is so important in the study of vision, it has been given a special name: counterphase, or phase-alternating, grating. To examine other uses of the counterphase grating, we can start by relating our psychophysical results to some recent studies of single neurons in the cat's visual cortex. John G. Robson and Graham Cooper of the University of Cambridge have recorded the responses of single cells to unidirectional moving gratings and to counterphase gratings. A cell that responds equally well to either direction of grating drift discharges as strongly to a counterphase grating as it does to a moving grating of the same contrast. A cell that is directionally selective, however, responds about half as well to a counterphase grating as it does to an equal-contrast moving grating. The direction-selective cell, then, seems to respond only to the half-contrast component of the counterphase grating that drifts in the cell's preferred direction. Parallel psychophysical measurements have been made with cats: relying on a behavioral technique, John M. Camisa, Randolph Blake and one of us (Levinson) have found that the cat's contrast threshold for a counterphase grating is twice as high as it is for a unidirectional moving grating. These results are similar to our psychophysical observations with human subjects.

The twofold difference between the contrast thresholds for counterphase gratings and unidirectional moving gratings is not the only reason for believing the human visual system decomposes the counterphase grating into its separate moving components. Prior adaptation to a grating moving in one direction can radically alter the appearance of a counterphase grating. If, for example, one stares at a grating moving to the left—thereby desensitizing mechanisms sensitive to leftward motion—a subsequently seen counterphase grating initially appears to drift to the right. After several seconds the grating again assumes its normal pulsating appearance. Presumably the time during which the counterphase grating appears to drift to the right is the time during which the leftward-selective channel is less sensitive than the rightward-selective channel. As the effect of adaptation decays, the sensitivities of the channels regain parity and the pulsating appearance returns. It does seem somewhat peculiar that the visual system can extract the two moving components from a stimulus that hides movement as well as a counterphase grating does. Similar ex-



COUNTERPHASE GRATING, also known as a phase-alternating grating, is made by combining two oppositely moving gratings that have the same bar width and contrast. In the experimental setup shown here two separate cathode-ray tubes are used; on one the grating moves to the right while on the other the grating moves to the left. The images on the two tubes are combined optically with a half-silvered mirror and are viewed by an observer. The resulting visual stimulus (*color*) shows neither direction of drift; it simply reverses repetitively in phase. Nevertheless, the eye can readily decompose flickering image into its two oppositely moving component images.

© 1976 SCIENTIFIC AMERICAN, INC

The exhibit you've been waiting for since 1325 B.C.

TREASURES OF TUTANKHAMUN

The elegant solid gold funerary mask of the boy-king Tutankhamun. One of fifty-five treasures on display from King Tut's fabled hoard.

National Gallery of Art
Washington, D.C.
November 17, 1976—March 15, 1977

New Orleans Museum of Art
New Orleans, Louisiana
September 15, 1977–January 15, 1978

Seattle Art Museum
Seattle, Washington
July 15, 1978—November 15, 1978

Field Museum
Chicago, Illinois
April 15, 1977—August 15, 1977

Los Angeles County Museum of Art
Los Angeles, California
February 15, 1978—June 15, 1978

Metropolitan Museum of Art
New York, New York
December 15, 1978—April 15, 1979

This exhibition, organized by the Metropolitan Museum of Art, is made possible through the assistance of the National Endowment for the Humanities and Exxon Corporation.

EXXON

© 1976 SCIENTIFIC AMERICAN, INC

Announcing the 136th use for polarizing

By our personal count, the 136th use for Polaroid polarizing filters is in Variable Light Transmission windows. VLT for short.

Polaroid VLT windows let in as much or as little light as you select. So you can easily eliminate the glare from your view. You can even eliminate the *view* from your view (by shutting out 99.95% of the light).

Two Polaroid light-polarizing filters are used, one stationary, the other free to rotate. Turning the inner filter moves the axes of polarization increasingly out-of-parallel which decreases light transmission.

VLT polarizers also protect against infra-

© 1976 SCIENTIFIC AMERICAN, INC

lters: Adjustable sunglasses for airplanes.

red and ultraviolet radiation. And they're free of optical distortions.

Now, VLT windows are being used in private planes like the Beech King Air. But we think you'll find them in commercial planes some day. (The reasons are easy to see.)

Meanwhile, if you have a light-filtration problem, and you'd like to consult with the people who wrote the book, write Dept. A186, Cambridge, Mass. 02139. Or call, toll-free, 800-225-1618 (in Massachusetts, call collect, 617-547-5177).

Polaroid Corporation

Polaroid® © 1977 Polaroid Corporation

© 1976 SCIENTIFIC AMERICAN, INC

© 1976 SCIENTIFIC AMERICAN, INC

How your car can lose 2000 pounds on the Thornel diet.

Your present car weighs a lot because there's a lot of metal in it.

Now, imagine replacing most of that metal with a material that can lead to the reduction of the weight of a car, by half.

That's the Thornel® diet.

Thornel is a remarkable graphite fiber, produced by Union Carbide, which when used to reinforce high performance plastics, creates a material that can have five times the strength and stiffness, for equal weight, of metals.

In addition, Thornel absorbs shock, dampens vibration and will not rust or corrode.

Right now, there is a study showing the feasibility of taking a 4000-pound car and reducing it to 2000 pounds by using graphite materials for such parts as hood and trunk lids, doors, body panels, chassis and suspension members, wheels and bumpers.

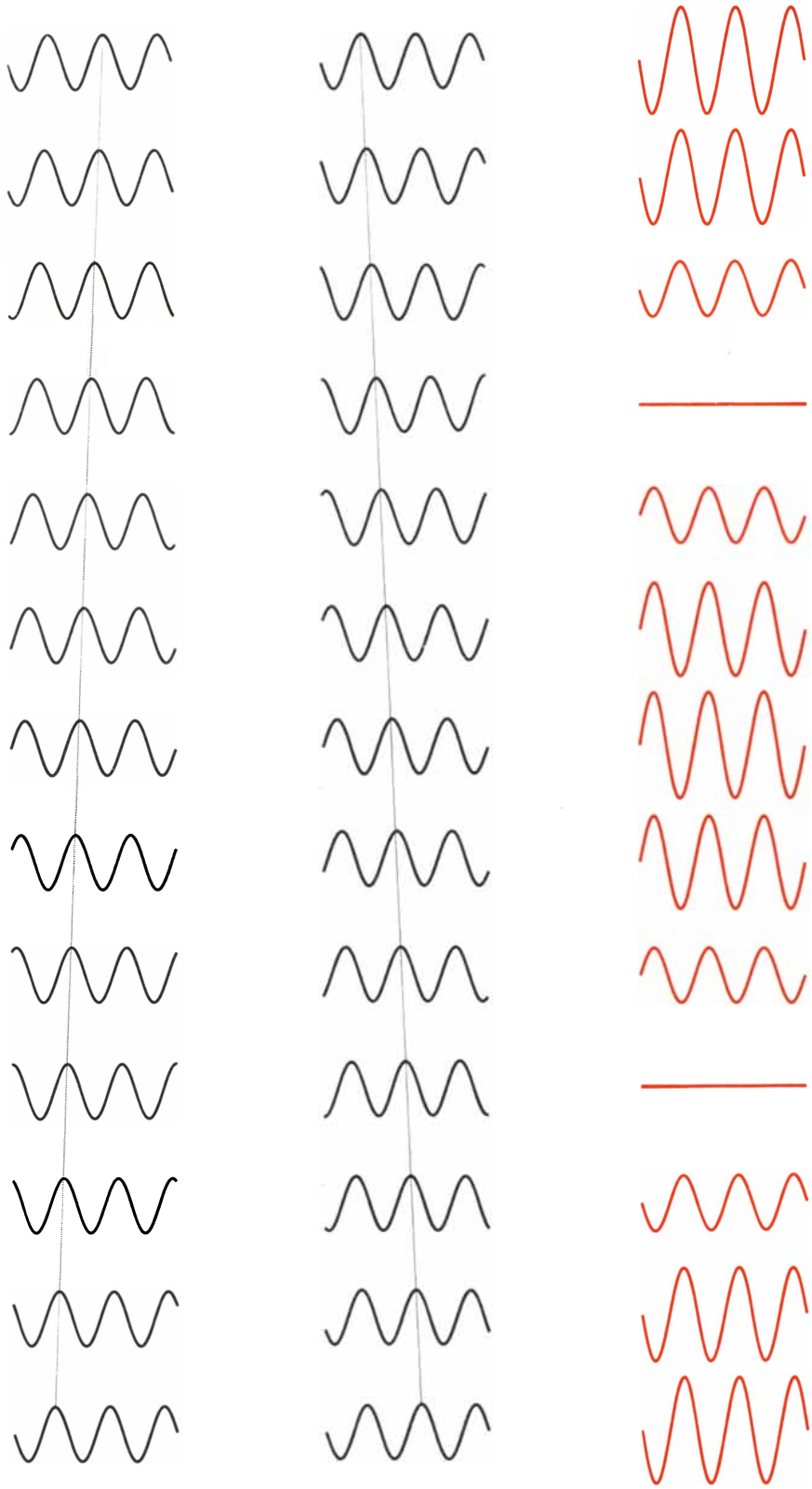
While this is being developed for the future, there are things happening now. Thornel fiber is being used in space vehicles, airplane components, bicycle frames, golf clubs, tennis rackets and fishing rods.

Thornel. One day it will make a big car with the weight of a small car.

And that's no small achievement.

Today, something we do will touch your life.

© 1976 SCIENTIFIC AMERICAN, INC



DECOMPOSITION OF A COUNTERPHASE GRATING into two oppositely moving gratings is represented graphically. The first column shows successive light-intensity curves for a sinusoidal grating drifting to the left at a constant rate. The second column shows the corresponding curves for a similar grating drifting to the right at the same rate. Each of the colored curves in the third column gives the algebraic sum of the other two curves in the same row; the curves in this column, representing the counterphase grating, go through one complete cycle of phase reversal in the 12 sampled time intervals. The peak-to-peak amplitude of the top curve in the third column is twice the amplitude of either of the other two curves in that row. This ratio is symbolic of the fact that the contrast threshold for detecting a counterphase grating is exactly twice contrast threshold for detecting either of two components presented by itself.

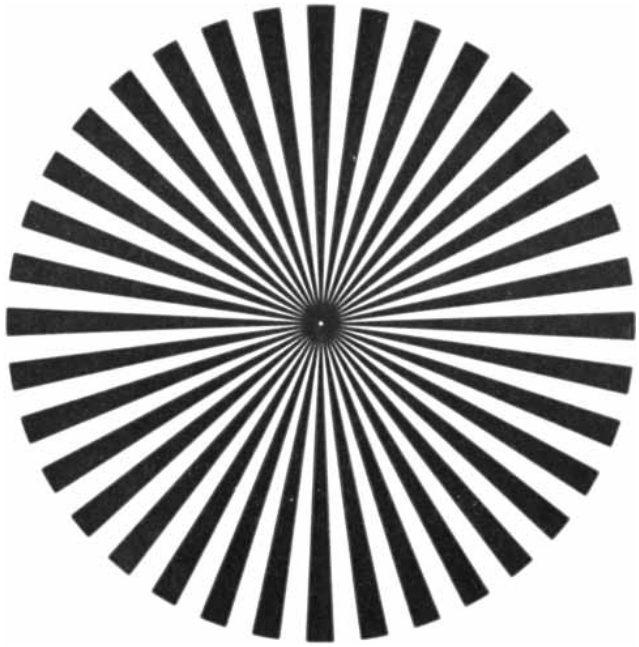
tractions are known, however, in other sensory domains. The decomposition into moving components is reminiscent of the ear's ability to separate a chord into its individual harmonic tones.

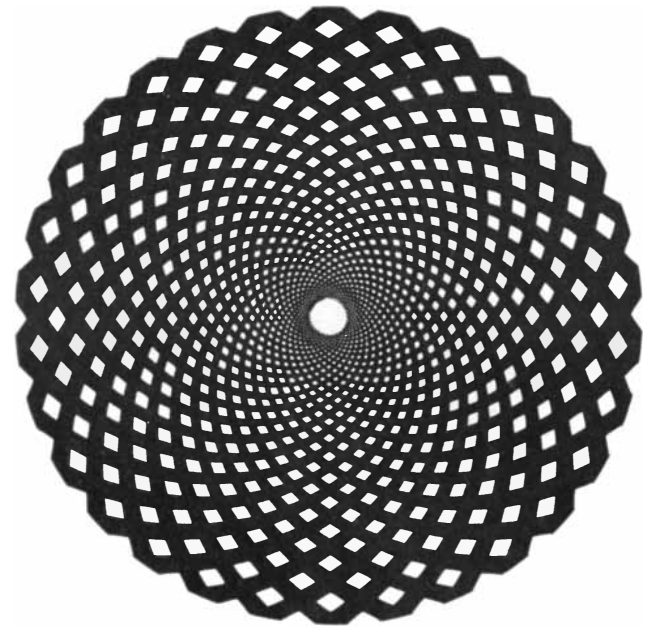
The ability of the visual system to use independent detectors to analyze a complex moving pattern into its components is no more surprising than the ability of the system to perform the reverse operation: the synthesis of such a pattern from its components. To demonstrate this synthesizing ability we presented a leftward-moving grating to one eye and a rightward-moving grating of equal contrast to the other eye, allowing the observer's brain to combine the two. The resulting synthetic percept was identical with a counterphase grating that could have been created by combining the two moving gratings with a half-silvered mirror. This synthesis reinforces the idea that the visual system can put complex patterns together as well as take them apart.

Most of the work we have discussed so far has focused on the role of human neural mechanisms in detecting motion, that is, in determining whether or not a moving target is visible. A more difficult and intriguing problem remains: the contribution of such mechanisms to the appearance of moving targets that are clearly visible. Our recent work at Northwestern shows that an observer's immediate perceptual history determines not only whether a moving target will be visible but also how the target will look if it does become visible. For example, we have found that adaptation to a sheet of random dots drifting uniformly in one direction can drastically change the apparent direction of subsequently viewed moving test dots. The shift in the perceived direction, which can be as much as 10 degrees of arc, is maximal when the adapting direction and the test direction are similar; the resulting shift is always away from the adapting direction. This direction shift is easily understood if one assumes that the neural code for perceived direction depends on a set of direction-specific mechanisms, each tuned to a somewhat different direction. The dots in the adapting pattern change the pattern of responsiveness among the mechanisms, thereby altering the apparent direction of the test stimulus.

It is now clear that with low-level stimulation at the contrast threshold the visual system decomposes complex stimuli and responds to each of the components without interference from the others. These separation processes do not tell the whole story, however. In everyday situations most stimuli are far above the contrast threshold and as a result the visual system's operation changes: interactions rather than independent responses are the rule, and the interactions tend to be inhibitory.

© 1976 SCIENTIFIC AMERICAN, INC





NEGATIVE AFTEREFFECT in the perception of motion can be demonstrated with these two designs. Place a copy of the radial "adapting" pattern at the left on a phonograph turntable. Stare at the center of the pattern as it spins at a speed of $33\frac{1}{3}$ r.p.m. Then look at the center of the stationary pattern at the right. The second pattern will briefly appear to be rotating in opposite direction. Experiments conducted with electronically generated moving stimuli reveal similar adaptation effects (*see bottom illustration on next page*).

To see how such inhibition works we turn to the horseshoe crab *Limulus polyphemus*. A great deal is known about the inhibitory activity in the faceted eye of *Limulus*. If one photoreceptor in that eye is strongly illuminated and a neighboring receptor is weakly illuminated, for example, the more strongly illuminated receptor reduces the neighbor's response. The inhibition is reciprocal: each photoreceptor can inhibit its neighbors and be inhibited by them. The net mutual inhibition depends on the relative strength of the illumination and the proximity of the photoreceptors to one another.

One important feature of such inhibition is that it does not operate all the time. For example, in the *Limulus* eye there are levels of illumination that are adequate to stimulate a photoreceptor but are inadequate to make that photoreceptor inhibit any of its neighbors. In order for inhibition to become apparent, stimulation must be increased. Thus at low levels of stimulation each receptor acts as an independent entity, whereas at higher levels of stimulation inhibitory interactions show up.

What role might inhibition play in human motion perception? Fred Attneave of the University of Oregon has pointed to the possible role of inhibition in various reversible or multistable figures. He cites the Necker cube, a visual illusion that seems to flip back and forth between orientations, as a good example of this phenomenon. There are reversible figures in motion perception too, and the counterphase grating is one of them. We mentioned above that a counterphase grating appears to pulsate or flicker when it is briefly viewed. When a low-frequency counterphase grating is observed for a longer time, however, its appearance changes in a way that is reminiscent of the alterations of more common reversible figures. The counterphase grating sometimes seems to drift in one direction and then in the other. The grating produces a multistable percept, and multistability, as Attneave has shown, is a symptom of inhibition [see "Multistability in Perception," by Fred Attneave; SCIENTIFIC AMERICAN, December, 1971].

We have also made more objective measurements of reciprocal inhibition between channels tuned to opposite directions of motion. In one experiment we measured the contrast threshold for a rightward-moving grating after adaptation to two different kinds of stimulus. One was simply a rightward-moving grating; the other was a combination of that same rightward-moving grating with a leftward-moving one. We knew that the rightward-adapting grating by itself would decrease the visibility of a test grating moving in the same direction. We wanted to determine whether an added leftward component could reduce the effectiveness of the rightward-adapting stimulus. If adding increasing amounts of the leftward component produces more inhibition, the adapting power of the rightward component should be reduced. As it turns out, that is exactly what happens.

Work on the physiology of motion sensitivity in lower animals complements our observations on inhibition in human motion-perception mechanisms. In general inhibition serves the function of "neural sharpening"; in other words, it accentuates the differences between a mechanism's responses to different stimuli. The responses of directionally selective neurons in the visual cortex of the cat to their nonpreferred directions of movement are often lower than the maintained, unstimulated activity of those cells. A number of investigators have recently shown that this reduction in response can sometimes be eliminated after administration of the drug bicuculline, an antagonist for the putative inhibitory neurotransmitter gamma-aminobutyric acid. In this case inhibition increases the difference between a neuron's responses to motion in its preferred and its nonpreferred directions and thereby improves the directional selectivity.

Inhibition between direction-specific units is apparently weak at birth and its development requires some visual experience. John D. Pettigrew, now at the California Institute of Technology, has found a lack of inhibition in the visual response to movement in the nonpreferred direction in the cortex of a newborn cat or a visually deprived adult cat. Consistent with this observation are experiments reported by Max Cynader at Dalhousie University. Cynader and his colleagues raised cats in a visual environment in which the cats were exposed to only one direction of motion. The vast majority of direction-selective cortical neurons of these "unidirectional" cats preferred movement in the direction to which the cats had been exposed during rearing. This bias in direction preference contrasts sharply with the

© 1976 SCIENTIFIC AMERICAN, INC

SPECIAL STIMULUS employed by the authors in their laboratory to study the perception of motion under controlled conditions consists of a moving pattern of small, bright dots electronically generated on the screen of a cathode-ray tube with a computer. To an observer the stimulus looks like an infinitely large sheet of random dots moving behind a stationary aperture.

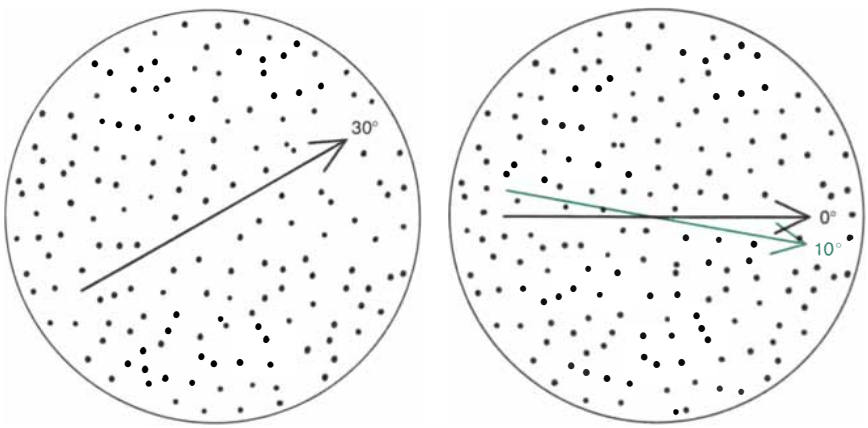
comparatively uniform distribution of direction preference in normally reared cats. Although it is not yet known what perceptual deficits might accompany this abnormal distribution of direction preference, it is clear that the physiological characteristics of motion-perception mechanisms depend critically on the kind of visual world the young cat experiences. One can speculate that man's comparatively uniform sensitivity to different directions of movement might result from frequent exposure to all directions during infancy.

We mentioned above that there might be separate visual channels for analyzing the movement characteristics and the pattern characteristics of a stimulus. We have recently established that of these separate channels only those concerned with motion show directional selectivity. When we measured the contrast thresholds for both motion perception and pattern perception after adaptation to movement in one direction, we found that the motion thresholds were raised more when the test stimulus and the adapting stimulus moved in the same direction than when they moved in opposite directions. Pattern thresholds, however, were equal for all test directions. In other words, although the adapting stimulus was exerting a strong directionally selective effect on the motion-perception system, it was having no directional effect on the pattern-perception system.

These observations suggest that the neural mechanism for detecting motion is independent of the mechanism for detecting pattern. Many other experiments support this view. We have already noted that the motion and pattern aspects of a moving target may have different contrast thresholds. Allan Pantle of the University of Miami has differentiated the two mechanisms in another way, by comparing their respective responses to variations in contrast. He relied on various procedures to determine how human direction-sensitive mechanisms respond to variation in the contrast of a moving target. All the procedures agreed: the two mechanisms have a very limited range of response to stimulus contrast. Their response increases linearly with the logarithm of stimulus contrast in the low-contrast region, but it is essentially independent of stimulus contrast once the contrast exceeds the threshold level by more than a factor of five or six. The response of the pattern-sensitive mechanisms does not appear to saturate at low contrast.

Additional evidence for the separation of the two systems comes from the examination of cases of brain pathology. Whitman A. Richards of the Massachusetts Institute of Technology has begun to accumulate evidence that some people who have normal pattern vision have difficulty seeing certain kinds of movement. Whether or not the difficulties are genetically determined remains to be seen, but the specificity of the loss is consistent with the idea of specialized machinery for the perception of motion. Moreover, certain kinds of neural damage (caused by gunshot wounds to the head, for instance) can result in a blind region in the visual field, called a scotoma. Observations of people with such neurological damage show that in areas of the visual field where stationary contours cannot be seen some residual sensitivity to moving stimulation is preserved.

A scotoma is often ignored by the person who has it; the visual system apparently extrapolates from the areas surrounding the functionally deficient region and fills it in with whatever the intact adjacent areas see. A dramatic new



ADAPTATION TO A SHEET OF RANDOM DOTS drifting uniformly in one direction (*pattern at left*) can drastically change the apparent direction of a subsequently viewed sheet of moving test dots (*pattern at right*). The shift in the perceived direction (*colored arrow*) can be as large as 10 degrees of arc. The resulting shift is invariably away from adapting direction.

© 1976 SCIENTIFIC AMERICAN, INC

form of extrapolation in normal observers that depends on a moving stimulus was recently discovered in our laboratory by Paul Tynan and one of us (Sekuler). To demonstrate this completion effect a grating with a low spatial frequency is made to drift slowly across a cathode-ray tube. A section of opaque black construction paper, roughly half the height of the screen, is taped across the middle of the screen from one side to the other. Since the paper is opaque, none of the grating can actually be seen in the region covered by the paper. Nevertheless, a normal observer who stares at the center of the black paper will not see grating sections at top and bottom with a middle part missing; instead he will see a grating covering the entire field! The middle, illusory part is a good deal dimmer in overall luminance than the top and bottom parts, but it will be seen quite vividly. These moving "phantoms" appear to travel at the same speed and in the same direction as the pattern that induces them. They are different in this respect from other well-known illusions of motion. Moreover, very small strips of moving pattern separated by large distances are able to produce vivid phantom gratings extending across the entire blank space between them. If the grating stops drifting, the phantom section disappears. The illusion, which can also be created with a rotating radial grating, requires movement of the real grating sections [*see illustration on this page*].

Floyd Ratliff of Rockefeller University has suggested that the visual system might economize on the amount of data it transmits by extrapolating information from certain select points in the visual field [see "Contour and Contrast," by Floyd Ratliff; SCIENTIFIC AMERICAN, June, 1972]. A highly repetitive pattern is therefore treated as though it covered the entire field, even when the pattern actually has gaps. If the information going to the brain were compressed and abbreviated for economy of transmission, the brain would have to fabricate much of our visual experience, working from scant clues. Redundancies in the visual world usually allow this fabrication to pass unnoticed, but illusions can result when certain unusual stimuli are compressed and then reconstructed. The moving phantoms may clarify the nature of the extrapolation process.

Although the visual system normally discards some of the information it receives, the lack of certain forms of information will seriously degrade its performance. The mathematical theory of the "ideal detector" shows that if an observer does not know what target to expect, and where and when to expect it, he cannot detect the target as easily as he could if he had that knowledge. The ideal detector represents the best detection performance any sensory system can achieve; failure to reach that level can provide clues about a sensory system's operation. With this possibility in mind Karlene Ball and one of us (Sekuler) have compared the detectability of dim, moving dots under two conditions: when the direction of movement was kept constant and when the direction of movement varied unpredictably from trial to trial. The observer's uncertainty about the direction of movement produced a dramatic loss in detectability. Analogous results have been obtained with uncertainty about the velocity of a target's motion. By examining the effect of various kinds and amounts of uncertainty we have improved our estimates of the direction-tuning and velocity-tuning functions of motion-sensitive mechanisms. Moreover, work with stimulus uncertainty provides a bridge between laboratory data and the motion perception of everyday life, in which stimulus uncertainty is pervasive and serious enough to play a major limiting role.

PHANTOM-CONTOUR EFFECT, a form of visual extrapolation in normal observers that depends on a moving stimulus, was discovered recently at Northwestern by one of the authors (Sekuler) and his colleague Paul Tynan. When one stares at the center of this radial pattern as it rotates on a phonograph turntable at 33⅓ r.p.m., one will see dim but distinct contours filling in the dark, blank region that interrupts the pattern. The phantom contours, which move in phase with their real neighbors, are enhanced by viewing the spinning disk in dim light.

The effects we have observed with moving visual stimuli resemble analogous effects in other senses. Accordingly, although the perception of motion has interesting and unique properties of its own, the study of this particular ability is also valuable for what it reveals about vision and the other senses more generally. Other visual abilities, for example, seem to make use of coordinated mechanisms, each designed to do a particular job. In the human eye exquisite sensitivity to just a few photons of light coexists with excellent spatial resolving power. This coexistence is made possible in part by the operation of two classes of retinal photoreceptors: the rods and the cones. Similar divisions of labor are known in touch, hearing and taste.

It makes a good deal of sense for the human visual system to separate information about pattern and motion, and about different directions of motion as well. Often during the evolution of our species it must have been much more important to respond to the direction in which some poorly defined form moved than to appreciate the details of that form. Any pedestrian who has dodged traffic in a large city appreciates this fact and should be thankful that the mechanisms for motion perception and pattern perception can coexist in a way that enables each to do what it alone does best.

© 1976 SCIENTIFIC AMERICAN, INC

Exoelectrons

They are electrons emitted by a fresh metal surface. Such surfaces are created by wear or by the cracking associated with metal fatigue; thus exoelectrons have become useful in the study of those processes

by Ernest Rabinowicz

Investigators who want to study the surface of a metal for wear, cracks or other defects can call on a wide variety of techniques: optical microscopy, electron microscopy, scanning electron microscopy and autoradiography, among others. Thus a graduate student doing a doctoral dissertation on some aspect of surface science commonly ends up with a collection of lovely pictures and then finds it quite difficult to extract any numbers that can be plotted on curves, evaluated and tested critically. Furthermore, he cannot be sure that the regions visible in the pictures are truly representative of the surface as a whole. This unsatisfactory state of affairs is being remedied by a new technique that involves the detection of exoelectrons, electrons that are emitted from surface atoms under certain conditions that provide enough energy to provoke their release. It turns out that such conditions are closely correlated with the surface changes produced by wear, cracks and material fatigue in general. The exoelectrons can readily be counted, so that the technique provides what pictures normally lack: numbers that can be displayed, plotted and compared.

FATIGUE CRACK IN ALUMINUM looks like this in an optical microscope. Although optical and other standard imaging techniques are usually effective in revealing cracks and other surface flaws, they give no information about how recent they are or how they are progressing.

The discovery of exoelectrons is related to a malfunctioning of Geiger tubes, which are used to detect the radiation, particularly electrons, produced by nuclear reactions or the decay of radioactive nuclei. Geiger tubes are essentially metal cylinders, typically about the size and shape of a large cigar, with a fine insulated metal wire running down the middle and a thin foil window sealing off one end. A high-voltage difference, commonly 1,500 volts, is applied between the center wire and the grounded cylinder. If a high-energy electron enters the window, it triggers an electric discharge that can be recorded and counted. Soon after the introduction of the Geiger tube in the 1920's it was observed that a freshly made tube gave a high and irregular counting rate over the first few hours or days and after that performed normally. A little later it was found that when any freshly machined surface was introduced into a Geiger tube, that too increased the initial counting rate.

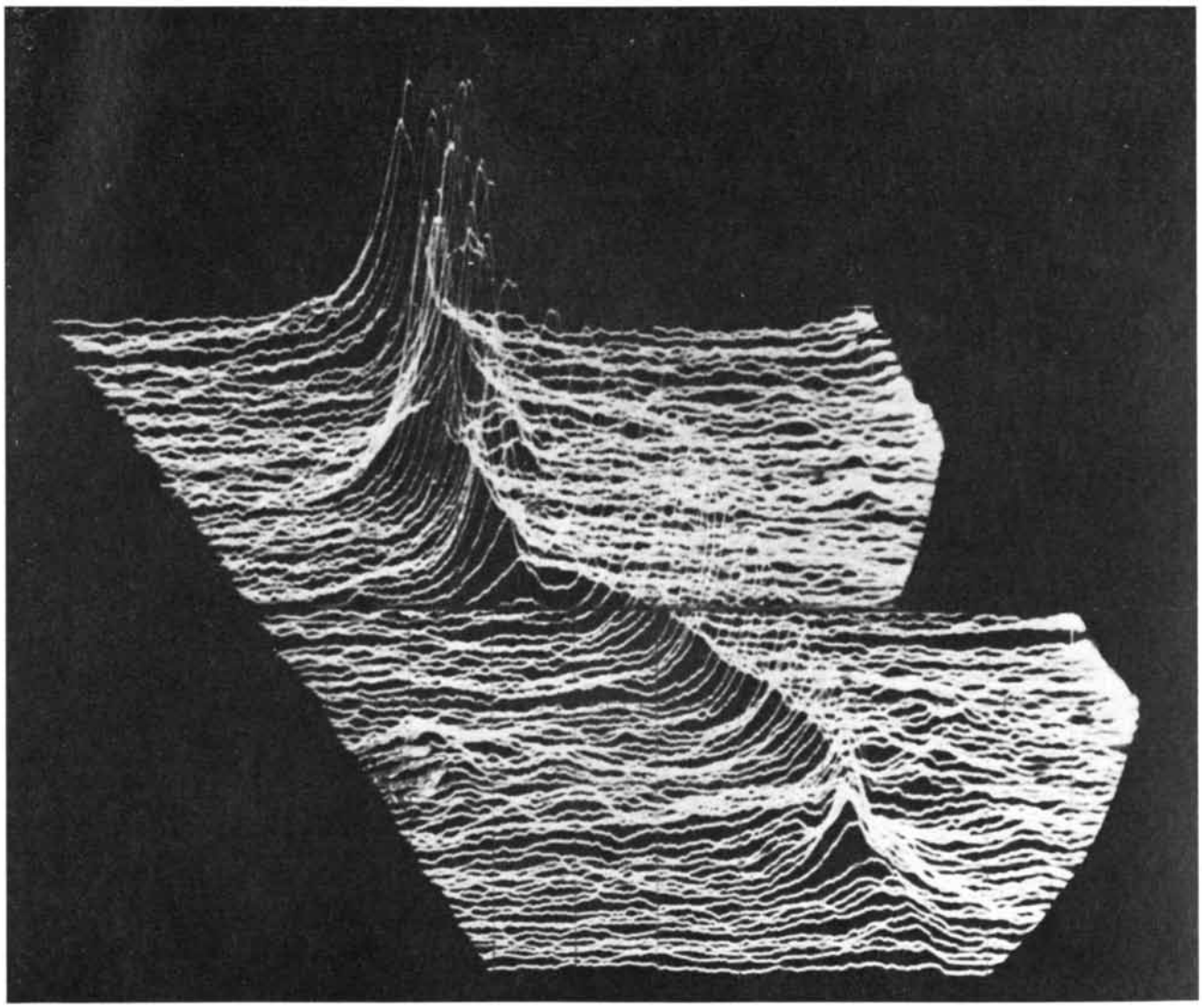
The phenomenon was first carefully investigated by a German physicist, Johannes Kramer, in the 1940's. He showed that all freshly prepared metal surfaces had the ability to give off electrons and that if the electrons were emitted inside a Geiger tube, they would of course trip the counting mechanism. Kramer found that the emitted electrons had an energy of only about one electron volt and that the emission lasted for only a few hours or days after the fresh surface was produced. Hence the increased initial counting rate of new Geiger tubes was finally explained.

© 1976 SCIENTIFIC AMERICAN, INC

The announcement of this discovery surprised the physics community. The conditions under which electrons could be emitted from a metal surface had been well established for many years, as the result of studies that had become classic, and the phenomenon discovered by Kramer simply did not fit. Briefly, it was well known that in order to remove electrons from any surface a specific amount of energy had to be supplied. For a metal such as copper the removal of one electron requires the expenditure of 4.3 electron volts; therefore one speaks of the work function of copper as being 4.3 electron volts. One well-known method of removing electrons from a surface is to heat the surface to a high temperature, thus enabling an electron to acquire enough energy by thermal excitation to exceed the work function and leave the surface. The process is known as thermionic emission. Another method consists in shining ultraviolet radiation on the surface. If the wavelength of the radiation is short enough for the energy in each photon to exceed the work function, a photon can transfer its energy by colliding with an electron, and the electron springs from the surface. This process is known as photoelectric emission. (When Albert Einstein received his Nobel prize in 1921, it was for his 1905 paper explaining the photoelectric effect.)

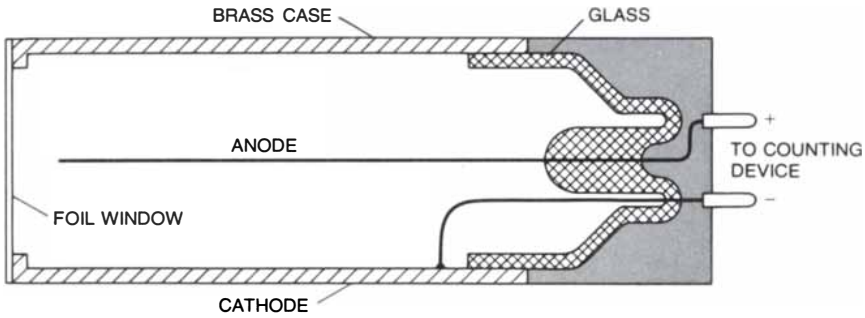
Kramer's discovery of the spontaneous emission of electrons was surprising precisely because it described the spontaneous occurrence of something that was known to require the expenditure of considerable energy. A good analogy might be a housing development in a part of the country where the water table is 4.3 meters below the surface. To get water to the surface every householder has to drill a well at least 4.3 meters deep and pump the water up from that depth. Then someone comes along who drills a new well and finds that water spurts out of the ground spontaneously to the height of a meter!

What mechanism supplies the energy that is required to overcome the work function and cause the emission of an electron? Kramer thought the energy was thermal in nature. For example, if an old metal surface is abraded, its oxide layer is removed and a naked metal surface is left. The fresh surface would immediately start to oxidize again, lib-



EXOELECTRON IMAGE of the fatigue crack shown optically on the opposite page provides information that is both more graphic and potentially more useful. In general the fresher the surface is, the more exoelectrons are emitted. The exoelectron image shown here was made by scanning the aluminum surface with a fine beam of ultraviolet radiation, which stimulates the emission of exoelectrons. The intensity of emission is simultaneously displayed on the face of a cathode-ray tube by an electron beam that follows the scanning pattern of the ultraviolet beam. The vertical deflection of the cathode-ray image is proportional to the intensity of exoelectron emission. The exoelectron image and the optical image on the opposite page were made by C. C. Veerman of the Delft University of Technology.

© 1976 SCIENTIFIC AMERICAN, INC



ANOMALOUS BEHAVIOR OF GEIGER TUBES led to the discovery of exoelectrons. Perhaps the best-known detector of charged subnuclear particles, a Geiger tube consists of a metal cylinder surrounding a fine metal wire, which are insulated from each other so that a difference of about 1,500 volts can be applied between the two. When an electron or another kind of charged particle enters the tube through a thin foil window, it triggers a brief electric discharge. Such discharges can be recorded and counted. It had been observed that newly made Geiger tubes often give spuriously high counts for hours or days. The cause was internal exoelectrons.

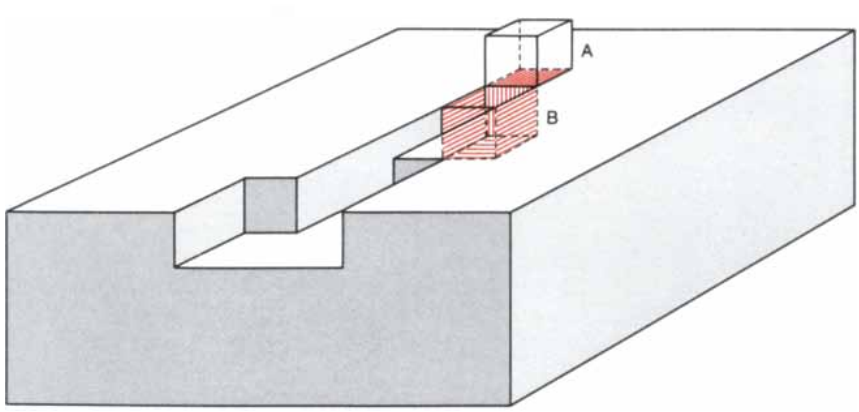
erating a considerable amount of chemical energy in the form of heat. Electron emission was also observed when a liquid metal cooled and solidified. There the latent energy of melting is given off in the form of heat, and conceivably some of the released thermal energy could give rise to the emission of electrons. Since such processes are exothermal, or heat-emitting, Kramer called the electrons exoelectrons. Although his explanation for the mechanism of electron emission is not accepted today (for example, the exoelectron emission in solidification is believed to be associated with changes in volume and the accompanying breakup of surface layers), the term exoelectron has survived. Exoelectrons are also, however, sometimes called Kramer electrons.

Kramer's book of 1950 describing his findings stirred up great excitement among investigators interested in surfaces, who perceived that exoelectrons might be useful in their work. Physicists also expended considerable effort trying to explain just how and why exoelectrons are produced. As a result exoelectrons were soon enlisted in the study of a wide variety of surface phenomena: friction, wear, metal-cutting, grinding, ball-milling, catalysis, corrosion, fracture, plastic deformation and many others. In addition there were numerous studies making use of physical and chemical techniques to elucidate the properties of exoelectrons and to discover the mechanism of their production. Some of this work was published; much was not.

I shall discuss the scientific and engineering aspects of exoelectrons in turn. Concerning the production of exoelectrons there can be no doubt that if a freshly machined metal surface is to give off an electron, there must be a source of energy. A freshly prepared surface has two obvious sources that an old surface does not have. One is the chemical-bond energy of the metal. Let us consider one greatly simplified way this energy could be made available. When a fresh surface is created by machining, an atom may be left in an exposed position, so that several of its chemical bonds are unsatisfied and, as it were, dangling. Later, when the atom forms new bonds with other atoms in the metal, the number of dangling bonds is reduced and the energy of the new bonds is made available, either in the form of heat or possibly in the form of exoelectrons. The other energy source is the one I have already described: the energy of oxidation or analogous processes.

One can estimate the amount of energy involved in each process. As an extreme example of the first case consider an atom of copper that is left dangling on a fresh surface so that it is held by a bond in only one direction, compared with the six directions in which it would be held if it were embedded in the crystal lattice of the metal. If the atom dropped into a nearby slot in the lattice, it might be held by bonds in four directions. In its initial state it was five-sixths of the way toward being evaporated (since an evaporated copper atom has no bonds to other copper atoms); after dropping into the slot it is only two-sixths evaporated (since we have assumed that bonds in two directions are still missing), so that the energy made available by the movement of the atom is the difference: three-sixths, or one-half, of the energy of evaporation of copper. The energy of evaporation of copper at room temperature is 2.9 electron volts per atom, and so we conclude that the energy resulting from the atomic rearrangement, being about half the energy of evaporation, might amount to 1.45 electron volts. Consider now the energy released by oxidation. If one atom of copper oxidizes to CuO (copper can also oxidize to Cu_2O with the liberation of a similar amount of energy), the energy liberated is .8 electron volt.

It is hard to see on the basis of these calculations how an exoelectron acquires the 4.3 electron volts it needs to overcome the work function and escape from a copper surface, to say nothing of the additional electron volt corresponding to the kinetic energy a typical exoelectron possesses. One possibility is that a copper atom might oxidize, adopt a more highly bonded lattice position and acquire an unusual amount of thermal energy all at the same instant. Another is that after the copper starts to oxidize, a very thin oxide layer with a reduced work function, say two electron volts, might be formed at the surface. This makes exoelectron emission easier because the energy requirement is re-



ONE THEORY OF EXOELECTRON EMISSION invokes the energy that is made available when an atom drops from a location in which it is loosely bound on a metal surface (*A*) to one where it is more tightly bound (*B*). In this schematic example the atom on the surface is bound in one direction only, to a neighboring atom below it, leaving it with free bonds in five directions. When the atom drops into a slot in the surface, it is held by bonds in four directions, leaving only two bonds free. Energy released in the change is roughly equivalent to about half the energy needed to evaporate the atom and could help to eject an electron from the surface.

© 1976 SCIENTIFIC AMERICAN, INC

duced, but at the same time the exoelectrons must now traverse the oxide layer in order to reach the surface. Another possible phenomenon to be considered is the adsorption onto the solid surface of gas molecules from the environment. Such adsorption could contribute energy, but the amount would be limited.

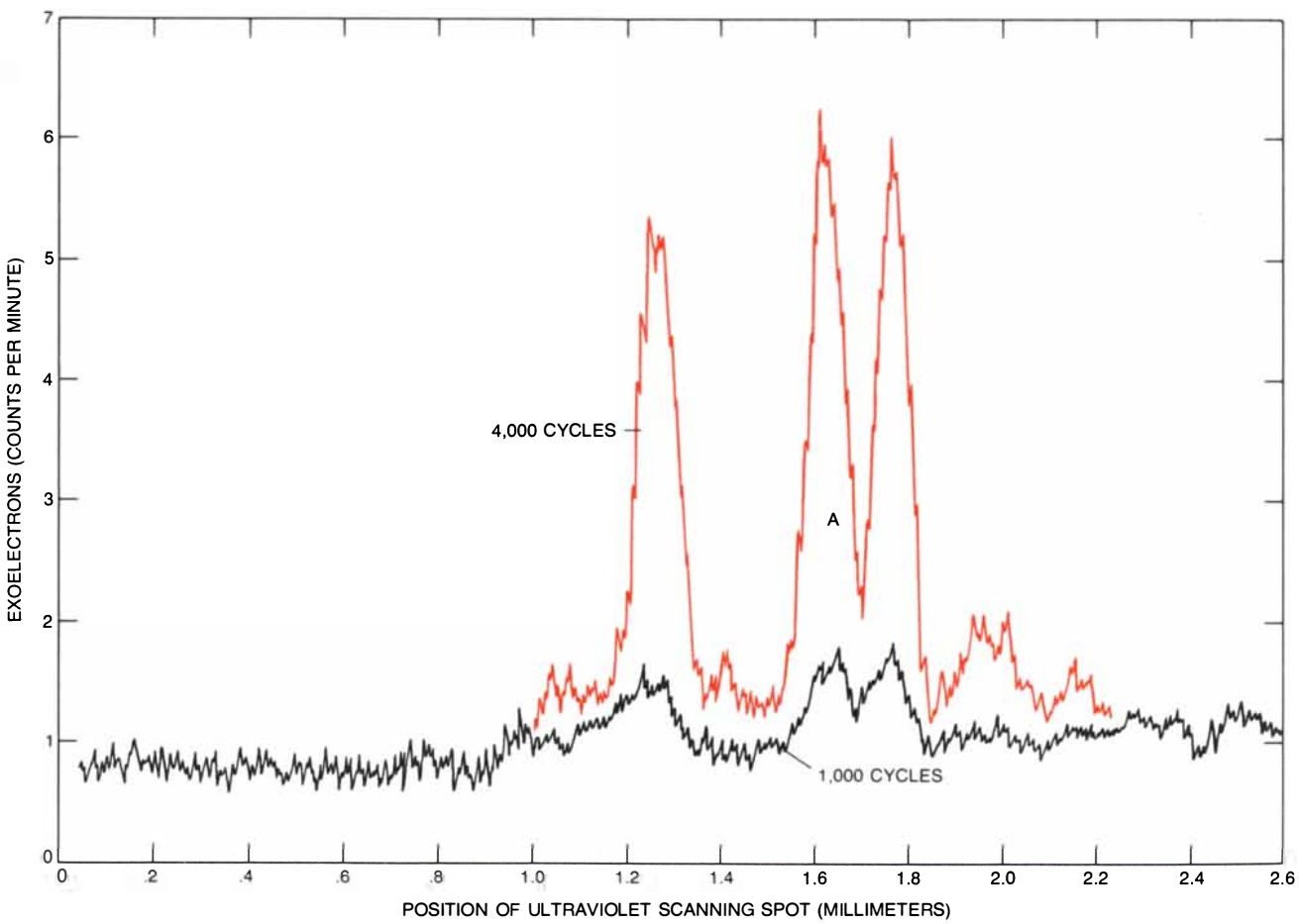
After considering such possibilities, one is left with the impression that exoelectron emission seems to depend on an unusual combination of favorable circumstances. In fact, calculations based on the rates of exoelectron emission confirm that the production of an exoelectron is a rare event. If one abrades a square centimeter of a metal surface, about 10^{15} metal atoms are exposed. If exoelectron emission were normal surface behavior, one might expect each surface atom to be the site for the emission of at least one electron, so that in the course of several days a total of 10^{15} exoelectrons might be emitted from each square centimeter of freshly exposed surface. The total number of exoelectrons is much smaller, amounting to no more than about 10^{8} per square centimeter even in favorable circumstances. Thus only about one surface site in 10 million sites is subjected to the special set of conditions required for the emission of an exoelectron.

So far every expert on exoelectron phenomena seems to have his own theory about the exact combination of conditions needed for exoelectron emission. One of the problems is that each worker runs his tests under somewhat different conditions, and he naturally finds that the results he gets do not quite match those of other workers. It seems likely that different combinations of the available mechanical energy, chemical energy, adsorption energy and thermal energy, balanced against a work function that changes as a result of chemical action and surface contamination, are operative in different circumstances.

The users of exoelectrons quickly found that exoelectron emission could provide information about many surface-related effects. For example, when a coarse material is put into a ball mill (essentially a rotating drum filled with steel balls) in order to reduce it to a powder, at first there is a steady reduction in the size of the particles but then the rate of reduction decreases and eventually levels off entirely. The phenomenon is faithfully mirrored by the rate of electron emission from the particles, which is fairly high in the early stages of the milling process but decreases steadily.

Another example is what happens when two metals are slid over each other under various loads; they emit more exoelectrons at high loads than at low loads. This was hardly surprising, since one would expect a greater surface area to be disturbed at high loads than at low ones. Although the correlation between exoelectron emission and load was interesting, no one could see any way to exploit the emission to gain a significant new understanding of the phenomenon of friction.

Basically the problem was that if exoelectrons were to be used to obtain knowledge about some other phenomenon, it was necessary to know under precisely what conditions exoelectrons are produced. As we have seen, how-

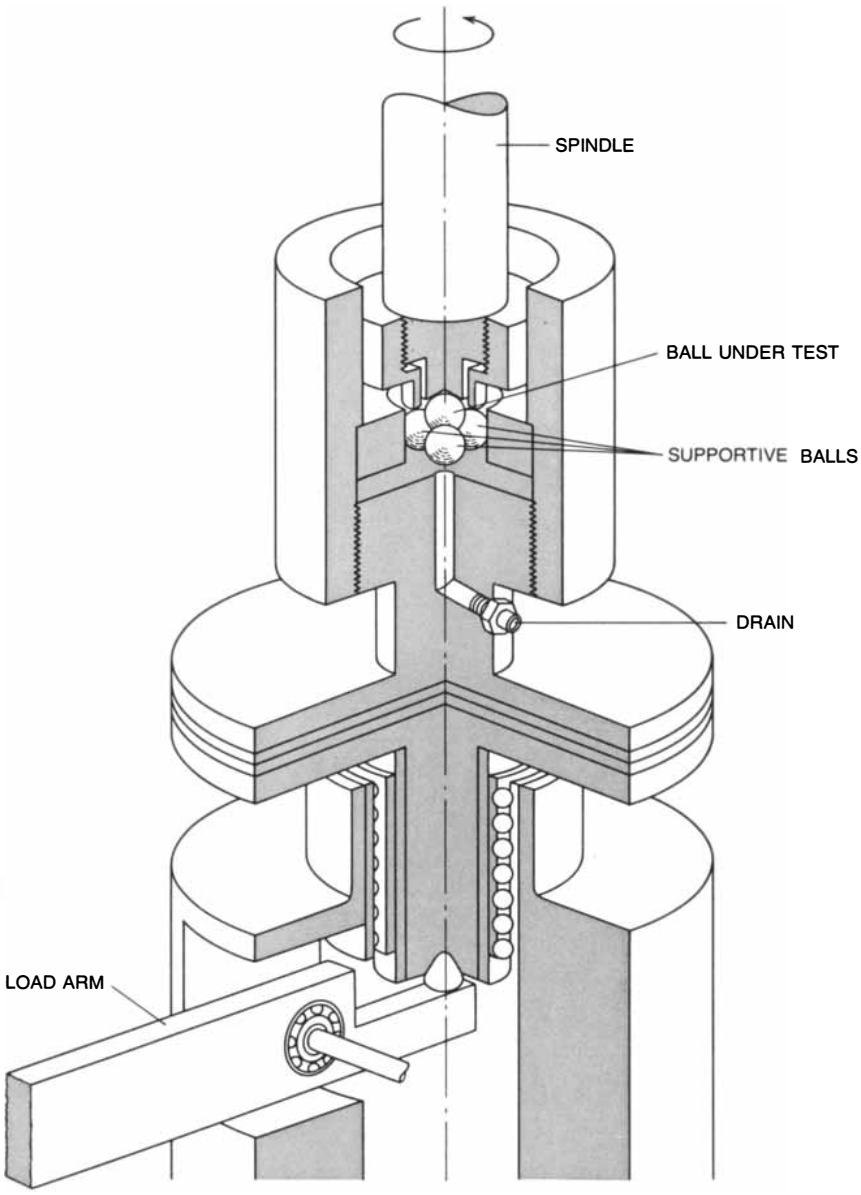


GROWTH OF CRACKS in an aluminum test strip is revealed in an exoelectron-emission study made by William J. Baxter of the Research Laboratories of the General Motors Corporation. The strip was stressed by repeated bending in a vacuum. At intervals exoelectron emission was stimulated by directing ultraviolet radiation at the surface. Traces show increase in emission after 1,000 and 4,000 stress cycles. The sample failed at location of peak *A* after 140,000 cycles. This showed that exoelectrons can reveal the site of growing cracks.

© 1976 SCIENTIFIC AMERICAN, INC

ever, this is exactly what has proved to be so difficult. Interest in exoelectrons reached a peak in 1957, the year of an international conference on the subject, and then, mainly for lack of progress in explaining the mechanism of their emission, declined to a low point in the mid-1960's. Starting in the late 1960's, however, interest again quickened with the discovery of potential new uses of exoelectrons, some of which I shall now describe.

One of the newer uses exploits the emission of exoelectrons as an index of radiation damage or radiation exposure. It has been found that if certain inorganic materials, such as calcium sulfate, lithium fluoride and beryllium oxide, are exposed to nuclear radiation or other types of high-energy radiation, their crystal structure is damaged in proportion to the length of the exposure. The damage can subsequently be reversed by annealing the material at a specific temperature, usually about 1,000 degrees Celsius. Simultaneously exoelectrons are emitted in rough proportion to the amount of radiation the specimen absorbed. The phenomenon is being considered for possible application in a reusable device for monitoring the radiation exposure of workers in nuclear power plants and similar places. The method promises certain advantages over other monitoring techniques, which measure the leakage of electrostatic charge or the exposure of photographic film.



BEARING-BALL FATIGUE is studied with a device in which a test ball, riding on three lower balls, is rotated at 3,560 revolutions per minute. The balls are continuously bathed with fresh lubricant. Contact is confined to a narrow track on the upper ball. The amount of stress on the test ball is governed by the weight on the load arm. The test is stopped when a microphone detects a level of vibration signifying that a ball has developed a serious surface defect.

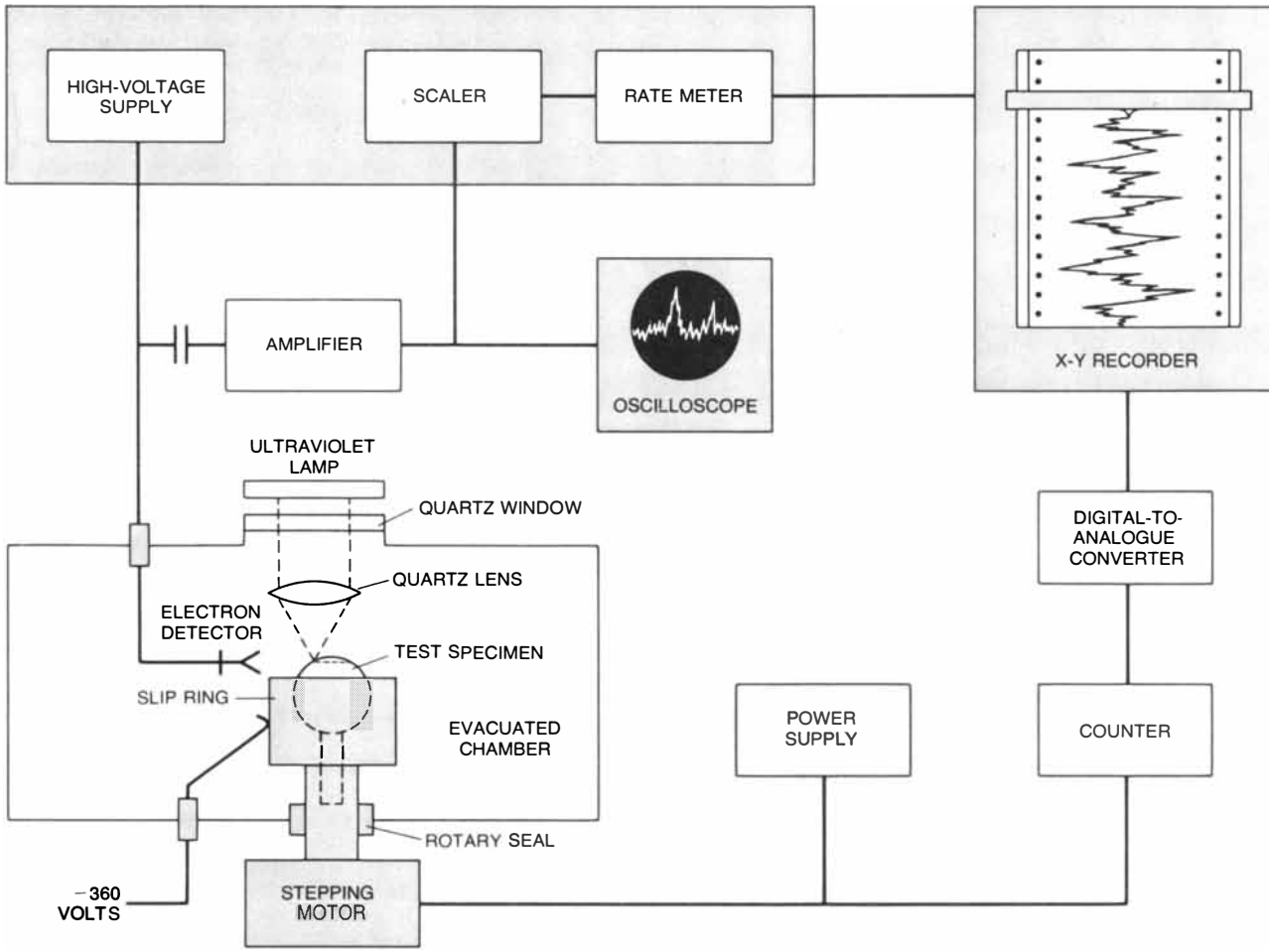
Another recent use of exoelectrons is for studying the process of crack growth in solids, particularly in the course of fatigue. Fatigue is a mode of failure in metallic solids that are placed under cyclically varying stress. Suppose a metal rod is repeatedly stretched and compressed. If the stresses are less than the yield strength of the metal, it will survive the first applications of stress without suffering damage. If the applications of the alternating stresses are continued, however, the rod will eventually break as the result of fatigue.

Earlier studies had shown that in such cases tiny cracks form in the metal, and that they gradually grow during the successive cycles of stress until the specimen breaks in two. In the early stages of fatigue not only are the cracks usually small and not easy to detect but also at that time there was no way to predict which ones would grow and which ones would remain as they were. Thus the study of crack growth had long presented a difficult problem. Since a growing crack constitutes a freshly exposed surface, however, one would expect it to emit exoelectrons, and it does. As a result exoelectrons are now being exploited to study crack formation and crack growth.

In such studies one should like to be able to locate precisely where the exoelectrons are being emitted. One method of location is purely mechanical. A mask with a small hole is moved slowly across the surface of the specimen in a scanning pattern so that one can tell at each instant the source of exoelectrons reaching the detector. As it happens, a much more elegant technique was available, as a result of earlier work with exoelectrons. One of the diagnostic techniques used by early workers to investigate the exoelectron process was to apply various types of energy (heat, light, mechanical vibration) to surfaces giving off exoelectrons and study the consequences. In general it was found that if external energy is applied, the number of electrons is increased. Electrons released in this way are called stimulated exoelectrons. The most effective stimulator is ultraviolet radiation. The frequency of the radiation must be kept low enough so that the energy in each photon is less than the work function, otherwise photoelectric emission will occur as well as exoelectron emission.

With ultraviolet radiation the number of exoelectrons emitted from a surface can be increased by a factor of more than 10,000 over the normal emission rate. This means that if one small region of a surface is illuminated while the

© 1976 SCIENTIFIC AMERICAN, INC



EXOELECTRON EMISSION from bearing balls was measured and recorded with this apparatus in the author's laboratory at the Massachusetts Institute of Technology. After being removed from the device shown on the opposite page balls are cleaned and mounted in a vacuum chamber so that a narrow beam of ultraviolet radiation will fall on the track of principal wear as the ball is slowly rotated. The beam stimulates the emission of exoelectrons. The exoelectron-emission rate is displayed on oscilloscope and plotted on an *x-y* recorder.

exoelectron emission from the entire surface is measured, nearly all the emission will be generated in the illuminated area. As a result one can sweep a beam of ultraviolet radiation over the surface and simply measure the electron emission as a function of time. The exoelectron count at different times indicates the emission from different areas of the solid surface.

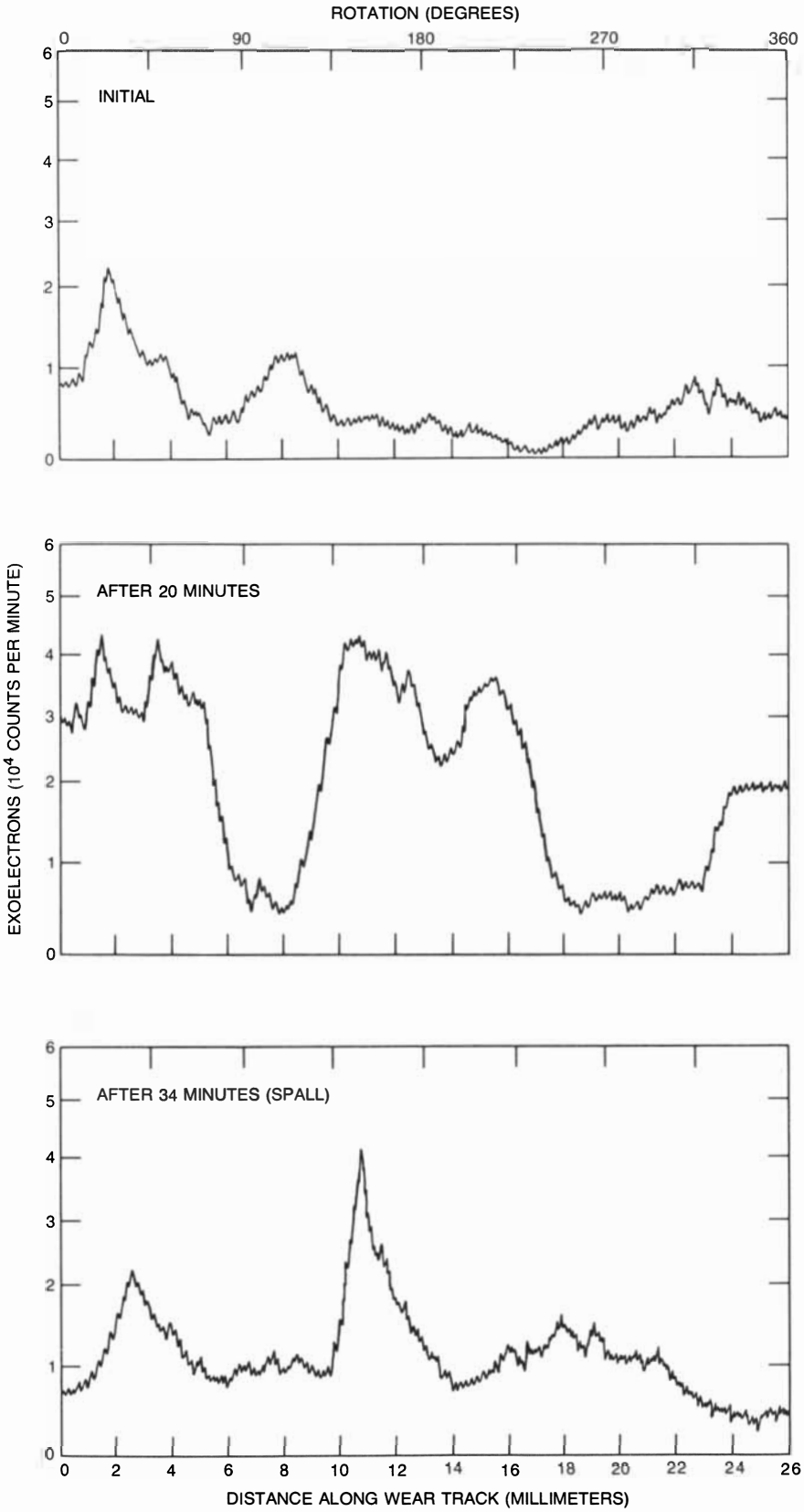
An early demonstration of the information that can be obtained with this technique was published by William J. Baxter of the Research Laboratories of the General Motors Corporation. He placed an aluminum strip in a vacuum chamber containing an electron detector and stressed the strip cyclically by bending it. Simultaneously he slowly scanned the surface with a sharply focused beam of ultraviolet radiation. Before the mechanical stresses were applied there was no exoelectron emission from the surface. After 1,000 cycles three sites gave above-average emission. After 4,000 cycles the active sites had become much more prominent. One of them eventually led to a failure after 140,000 stress cycles. It is clear that the active sites indicate locations where cracks form, grow and extend. Baxter's results were exciting because up to that time no methods were available for identifying fine cracks in metal surfaces and watching them grow to failure.

Baxter's work prompted us at the Massachusetts Institute of Technology to try to apply the same techniques to the problem of crack growth in rolling-contact devices, a process generally referred to as surface-fatigue wear. Surface-fatigue wear is the main cause of failure of ball and roller bearings and an important cause of wear in gears and in wheels rolling on rails. In ball bearings surface fatigue develops when a crack forms in one of the balls and slowly grows until eventually a sizable chip spalls off. Once this happens the bearing is quickly destroyed.

Surface-fatigue wear is an awkward phenomenon to analyze because there is a very large statistical scatter in the time required for failure to occur. Thus if 100 identical ball bearings are tested to failure, the times measured between the first and the last failure may differ by a factor of 100 or more. The problem is made all the more difficult because on a macroscopic scale there are no useful measurements one can make while waiting for a ball to fail. On a microscopic scale one can examine the surfaces for crack formation, but the growing cracks are tiny and hard to locate. At high magnification, to be sure, many fine cracks are visible, but there is no ready way to tell which one is growing rapidly and will eventually lead to a failure. It therefore seemed to us that exoelectrons might provide a way of identifying significant cracks and watching them grow.

The experimental problems seemed considerable. Since exoelectrons travel only very small distances in air before they are stopped, it is virtually essential to conduct experiments with them in a vacuum system. On the other hand, to be meaningful ball-bearing tests must be carried out under the conditions in

© 1976 SCIENTIFIC AMERICAN, INC



HISTORY OF BEARING-BALL FAILURE can be reconstructed by observing the rate of emission of exoelectrons and their location at intervals during a fatigue test. These records were made with the apparatus illustrated on the preceding page. The top trace shows the emission of exoelectrons before the test began. The middle trace shows the emission rate along the wear track after 20 minutes of testing. The ball finally failed by spalling, or losing a flake of material, after 34 minutes. The sharp peak in third trace coincides with location of the spall.

which the bearing usually works: in air with lubricated surfaces. Hence to measure bearing balls for exoelectron emission would call for running tests in a bearing tester in the presence of oil, cleaning off the oil with a solvent, evaporating off the solvent, putting the ball in a vacuum chamber, pumping the chamber down and hoping that all this would not interfere with exoelectron emission. When my colleagues were informed about the experiment, many of them thought it was highly unlikely that after all these manipulations any exoelectrons could be detected, but we decided to go ahead anyway. The U.S. Army Research Office offered to sponsor the work.

Two pieces of apparatus were used, a fatigue tester in which the bearing balls were stressed and a vacuum chamber in which the exoelectrons were to be detected. The fatigue tester was of the Barwell four-ball type, in which a test ball, held by a spindle and rotated at 3,560 revolutions per minute, rides on three lower balls that are constrained by a ring and disk to travel in a circular track. All the balls are half an inch in diameter. (It has been established that tests run in this apparatus, with four balls rolling together, give results that agree well with behavior in a conventional ball bearing, which has an inner race, an outer race and a large number of balls between them.) Lubricant is dripped continuously into the test section. A microphone attached to the four-ball tester measures vibration and automatically shuts off the power when any one of the balls begins to spall. The geometry is such that contact is confined to a relatively narrow track on the upper ball, so that damage and failure generally occur there. Thus only the upper ball need be examined for exoelectron emission.

The evacuated exoelectron-detection apparatus uses ultraviolet radiation for the stimulation of exoelectron emission. The ultraviolet rays pass through a quartz window and are focused on the test specimen with a quartz lens. For the results reported here the spot size was two millimeters by .27 millimeter. The emitted electrons are detected in an electron multiplier. The output pulses are summed and displayed on the y axis of an x-y recorder.

The initial measurements were made of the exoelectron emission from lubricated bearing balls that had been scratched with a diamond scriber and cleaned with a Freon degreaser. The exoelectron-detection apparatus worked as we had hoped, revealing sharp emission peaks when the ultraviolet beam illuminated the scratched region on the ball. We then carried out a number of actual wear tests. In one typical test steel balls were run in the Barwell tester at a

© 1976 SCIENTIFIC AMERICAN, INC

stress level where the anticipated life was about 40 minutes. Every 10 minutes or so the top ball was removed, tested for exoelectron emission and returned to the tester for further stressing. The test ball spalled after a total of 34 minutes of testing. When the wear track on the ball was tested at 20 minutes, it had shown three big peaks of exoelectron emission. The site of one of the peaks proved to be the eventual spall site.

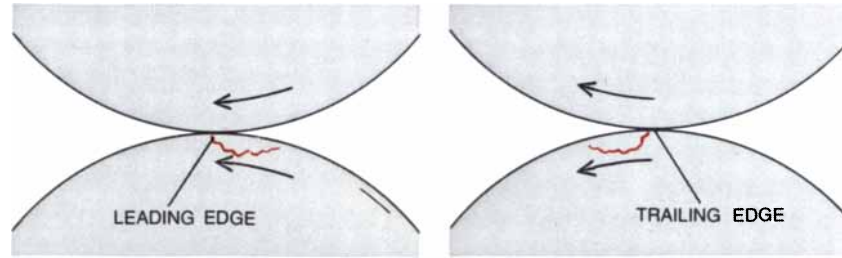
Besides enabling us to locate the position of an eventual surface-fatigue failure, exoelectrons throw light on the mechanism of the failure process. For example, if a bearing ball is stressed in a test apparatus (or in actual use) so that the ball always rolls against the adjacent surface in the same direction, a crack that develops at the surface typically grows inward at an angle and can eventually lead to a spall. The question arises: Does the crack grow in the direction of rolling or away from it? The question can be studied by observing how the exoelectron emission near a crack that grows into a spall varies with time.

Our studies have shown that the leading edge of the eventual spall shows a prominent peak early in the test but that the peak gradually decays, presumably as the crack extends under the surface and parallel to it. The trailing edge shows hardly any exoelectron emission above the background rate until spalling takes place. Hence the crack develops from the leading edge to the trailing edge, confirming earlier studies that had provided less direct evidence.

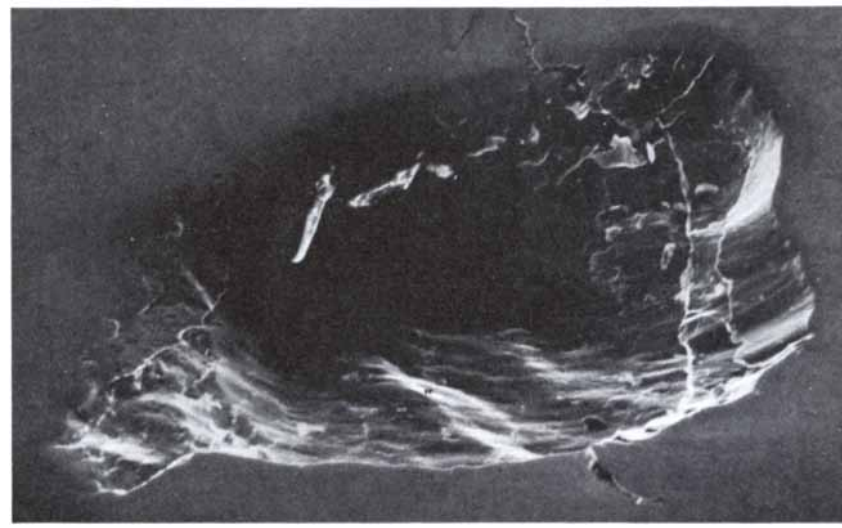
In our work with bearing balls the ultraviolet beam can be directed along a single track because of the way stress is localized in the Barwell tester. One can readily imagine, however, a technique in which the surface of a specimen is swept with a fine ultraviolet beam in a raster pattern while the exoelectron emission is displayed synchronously as the vertical deflection of an electron beam scanning the face of a cathode-ray tube. The first demonstration of such a technique was made in 1969 by C. C. Veerman of the Delft University of Technology in the Netherlands. In his exoelectron images a fatigue crack in an aluminum specimen resembles a mountain ridge [*see illustration on page 75*].

In a similar scanning method the electron-emission rate is made to modulate the intensity of the electron beam in the manner of a television image. In such a picture light areas correspond to regions emitting exoelectrons.

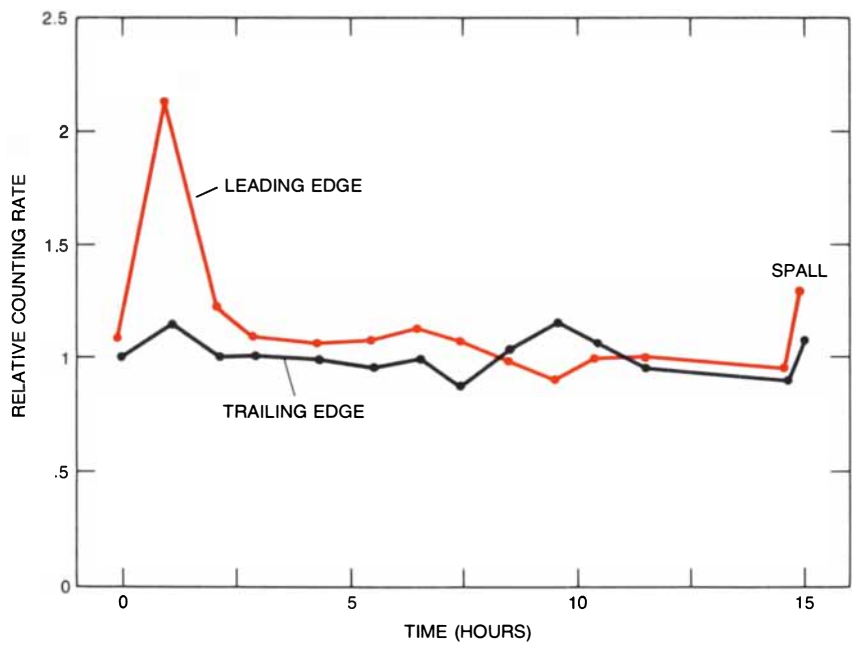
It is risky to predict where future developments of exoelectron-emission phenomena will lead. Within the next decade it should be possible, with special techniques for detecting exoelectrons in air, to make routine examinations of the entire wing structure of an airplane, searching for growing fatigue



CRACK LEADING TO SPALL in a bearing ball might grow from the leading edge to the trailing edge (*left*) or in the opposite direction (*right*). Before development of exoelectron-emission technique there were only indirect methods for establishing the direction of crack growth.



TYPICAL FATIGUE SPALL in a bearing ball is shown in a scanning electron micrograph that was made by the author. The spall is 1.8 millimeters long; the leading edge is at the left.



EXOELECTRON EMISSION from the ball that ultimately developed the spall shown in the micrograph above this illustration enabled the author to establish that the initial fatigue crack appeared at what was to become the leading edge. Exoelectron emission from the leading edge is in color; emission from the trailing edge is in black. Evidently crack starts early at leading edge, then runs under surface, where it cannot be detected, and finally emerges at trailing edge.

© 1976 SCIENTIFIC AMERICAN, INC

We'll let someone else tell you how good our belt-drive turntables really are.

"...in the same class with a number of more expensive products, including many of the direct-drive record players we have seen."

This quote, from the Hirsch-Houck Labs' report in *Stereo Review*, refers to the Dual 510, a semi-automatic belt-drive turntable. Since direct-drive models (especially our own) are accepted as *the* standard of performance, Hirsch-Houck's comparison is not to be taken lightly.

The 510 also benefits from comparison with other semi-automatic turntables. Dual's unique sensor locates the 12-inch and 7-inch lead-in grooves for you. You don't have to guess where they are. And there's no way to drop the tonearm accidentally; the cue-control lifts it automatically at the end of play.

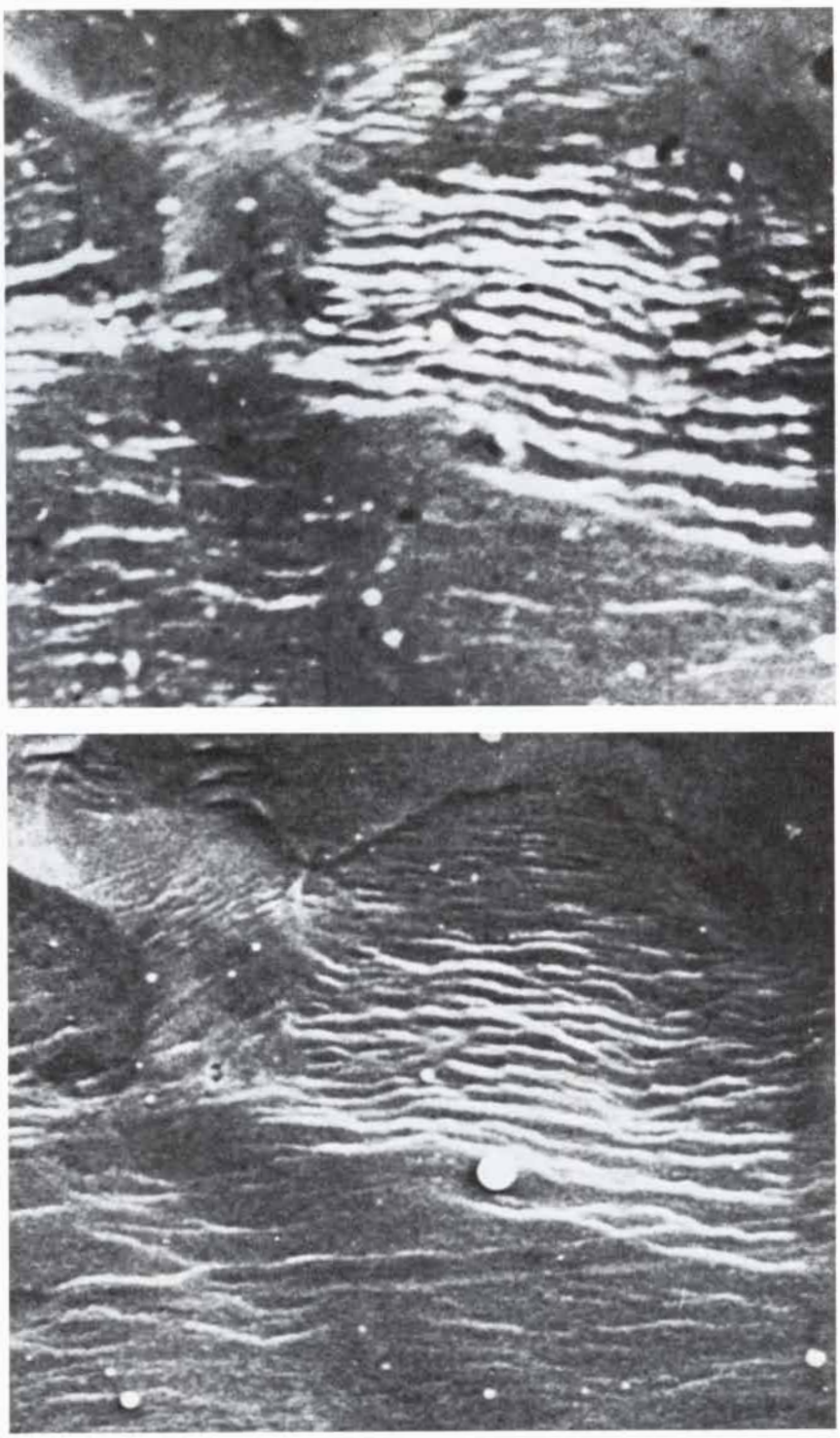
You might also compare the 510 with your present turntable, or any other you may be considering. Keep in mind the 510's many other features and refinements described below.

Your old records will sound better, your new ones last longer.

Dual 510.
True four-point gimbal tonearm suspension. Synchronous motor, precision-ground belt, unique Vario-pulley, dynamically-balanced platter. 6% pitch-control, illuminated strobe. Lead-in groove sensor. Cue-control viscous-damped in both directions. Price: less than $200.
Dual 502, similar except less sensor, pitch-control and strobe, less than $160.
Dual 1249, fully automatic single-play/multi-play, less than $280.

Dual® **United Audio Products**
120 So. Columbus Ave.
Mt. Vernon, N.Y. 10553
Exclusive U.S. Distribution Agency for Dual

cracks. As the speeds of passenger trains increase it will be important to develop fast and reliable methods for detecting incipient flaws in wheels and rails; perhaps exoelectron techniques will serve the purpose. It is conceivable that exoelectrons could be applied in earthquake prediction by indicating which cracks in the surface of the earth or in subsurface rock structures are actively growing. I suspect that as we learn more about the exact conditions underlying the emission of exoelectrons, uses will be found that we cannot imagine today.



TWO TECHNIQUES FOR VISUALIZING CRACKS in a steel surface that had been subjected to plastic strain are compared. The image at the top is an exoelectron micrograph in which brightness is proportional to the emission of exoelectrons. The image at the bottom depicts the same region in a conventional scanning electron micrograph; surface cracks show up as bright lines. The pattern of exoelectron emission closely follows the pattern of cracks. Micrographs were made by Baxter and Stanley R. Rouse of General Motors Research Laboratories.

© 1976 SCIENTIFIC AMERICAN, INC

Unique full-function 8-digit wrist calculator... available only as a kit.

A wrist calculator is the ultimate in common-sense portable calculating power. Even a pocket calculator goes where your pocket goes—take your jacket off, and you're lost!

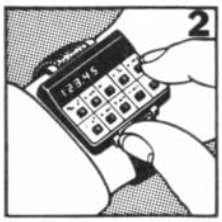
But a wrist-calculator is only worth having if it offers a genuinely comprehensive range of functions, with a full-size 8-digit display.

This one does. What's more, because it *is* a kit, supplied *direct* from the manufacturer, it costs only a very reasonable $19.95. And for that, you get not only a high-calibre calculator, but the fascination of building it yourself.

How to make 10 keys dothe work of 27

The Sinclair Instrument wrist calculator offers the full range of arithmetic functions. It uses normal algebraic logic ('enter it as you write it'). But in addition, it offers a % key; plus the convenience functions $\sqrt{x}$, $1/x$, x^2; plus a full 5-*function* memory.

All this, from just 10 keys! The secret? An ingenious, simple three-position switch. It works like this.



1. The switch in its normal, central position. With the switch centered, numbers—which make up the vast majority of key-strokes—are tapped in the normal way.

2. Hold the switch to the left to use the functions to the left above the keys...

3. and hold it to the right to use the functions to the right above the keys.

The display uses 8 full-size red LED digits, and the calculator runs on readily-available hearing-aid batteries to give weeks of normal use.

Assembling the Sinclair Instrument wrist calculator

The wrist calculator kit comes to you complete and ready for assembly. All you need is a reasonable degree of skill with a fine-point soldering iron. It takes about three hours to assemble. If anything goes wrong, Sinclair Instrument will replace any damaged components *free*: we want you to enjoy assembling the kit, and to end up with a valuable and useful calculator.

Contents
Case and display window.
Strap.
Printed circuit board.
Switches.
Special direct-drive chip (no interface chip needed).
Display.
Batteries.

Everything is packaged in a neat plastic box, and accompanied by full instructions. The only thing you need is a fine-point soldering iron.

All components are fully guaranteed, and any which are damaged during assembly will be replaced free.

The wrist-calculator kit is available only direct from Sinclair Instrument. Take advantage of this 10 day money-back offer.

Send the coupon today.

Sinclair Instruments Inc.
6 Commercial Street
Hicksville, N.Y. 11801

To: Sinclair Instruments Inc.
6 Commercial Street
Hicksville, NY 11801

Please send me

_____ Sinclair wrist calculator kit(s) at $19.95 $_____

Sales tax (NY residents) _____

Shipping & handling ($2.50 per unit) _____

Enclosed is check/MO (payable to Sinclair Instruments Inc.) TOTAL $_____

Name _____

Address _____

City _____

State _____ Zip _____

I understand that you will refund the purchase price of the kit in full if I return it in saleable conditiion within 10 days of receipt.

© 1976 SCIENTIFIC AMERICAN, INC

Cratering in the Solar System

The age of space exploration has shown that craters such as those on the moon are a feature of all the bodies in the inner solar system. Cratering is therefore one of the main keys to the system's history

by William K. Hartmann

Every planet and satellite in the inner solar system is pitted with craters. The craters on the moon were discovered by Galileo when he turned his telescope on the sky in 1609. The craters on the earth have mostly been identified by geologists in recent decades. The craters on Mars were discovered by the spacecraft *Mariner 4* in 1965. The craters on Mars's two moons, Phobos and Deimos, were photographed by *Mariner 9* in 1971. The craters on cloud-shrouded Venus were revealed when the surface of the planet was mapped by radar in 1972. And the craters on Mercury abound in the pictures sent back by *Mariner 10* in 1974. No planet or satellite in the outer solar system has yet been photographed with sufficient resolution to reveal whether or not it is cratered.

For students of the solar system the ubiquity of the craters on the surface of the inner planets and satellites is fortunate. The form and number of the craters provide a wealth of information about both the astronomical and the geological processes that shaped the planets. There is every reason to believe that the craters are the vestige of ancient impacts of massive meteorites on the surface of the planets; craters are still being made by such meteorites in the present geological period. As scars left by impacting interplanetary bodies the craters provide evidence about the population of those bodies in the interplanetary environment over hundreds of millions of years. As features modified by geological processes such as volcanism, tectonic activity and erosion they are clues to the surface environments of the planets. Moreover, if the rate at which the craters are made is known, then counting the number of craters that have accumulated on a surface reveals the age of the surface. The craters are thus a key to reconstructing the history of each of the inner planets.

The crater record is least altered on the surface of a body such as the moon, where erosive forces are the least complex. On the moon's atmosphereless and waterless surface no sediments have been blown into crater bottoms and no waves have gradually cut away crater walls. Once lunar craters are formed they are likely to last for a long time. There are three main ways, however, that they can be obliterated. First, an early crater may be destroyed by a later impact (or impacts) directly on top of it. Second, older craters may be covered by material ejected in the formation of younger craters. Third, craters may be buried by lava flows.

The most direct way to gain information about the population of interplanetary bodies that made the craters is to study the craters in young geological provinces where they are so widely scattered that the complexities of obliteration are negligible. At the same time the craters must not be so sparsely scattered that their numbers are too few for sound statistical analysis. On the moon the most nearly ideal geological provinces for such studies are the lava plains making up the maria, or "seas." Samples of lunar rocks indicate that the lava flowed some 3.5 billion years ago, wiping clean the slate of preexisting craters and creating a fresh surface to record new ones. No major internal activity has disturbed the maria since they were formed. In that time enough large craters have accumulated on the maria to provide a good sample without presenting problems of overlapping or obliteration.

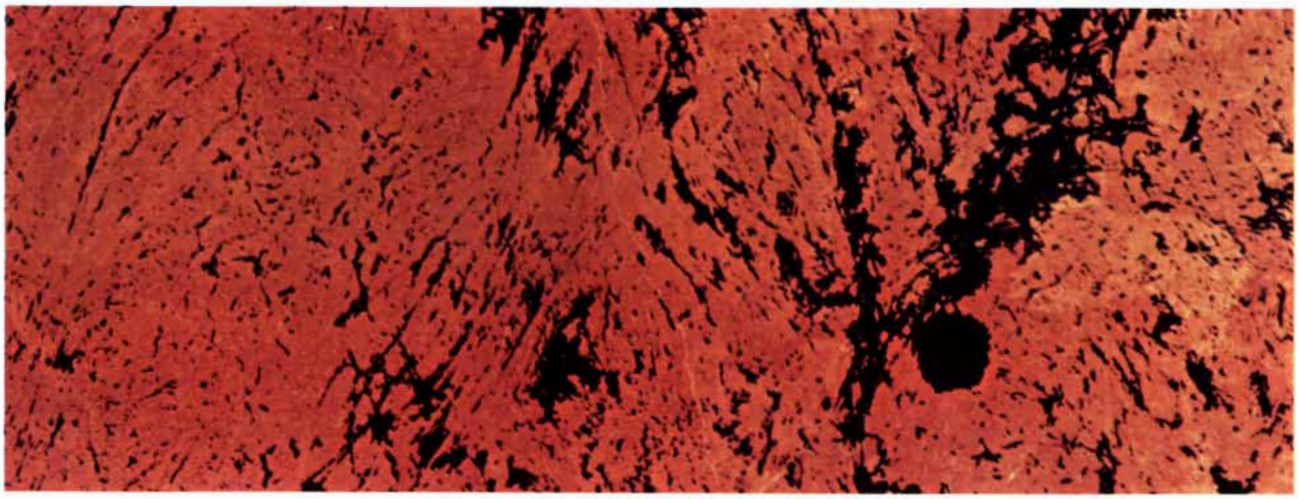
When a meteorite several kilometers in diameter struck the smooth lava surface of a lunar mare, the results were vividly recorded. A huge explosion lifted up layers of lunar rock to form the crater's rim. High-speed jets of material shot outward from the explosion and fell to the surface as long streamers of dust and glass beads: the bright "rays" visible through telescopes on the earth. Low-speed ejecta fell back as a thick blanket of hummocky terrain surrounding the crater for several times its diameter. The rain of material was heaviest on the crater's rim of uplifted rock, adding to the height of the rim.

Many isolated blocks of debris also fell back, simulating meteorites and themselves forming small craters. It is therefore necessary to distinguish between the primary cratering by meteorites and the secondary cratering by blocks of surface material. Although the secondary craters created by an impact are always smaller than the primary crater, some of them are as large as a few kilometers in diameter. Accordingly it is trickier to analyze the smaller craters than it is to analyze the larger ones. Furthermore, smaller craters can be confused with volcanic pits obscured by debris from adjacent impact craters. Most craters larger than four kilometers in diameter probably are directly related to the population of meteoritic bodies in interplanetary space, whereas most smaller craters probably are not. I shall therefore confine this discussion to what can be learned from the larger craters.

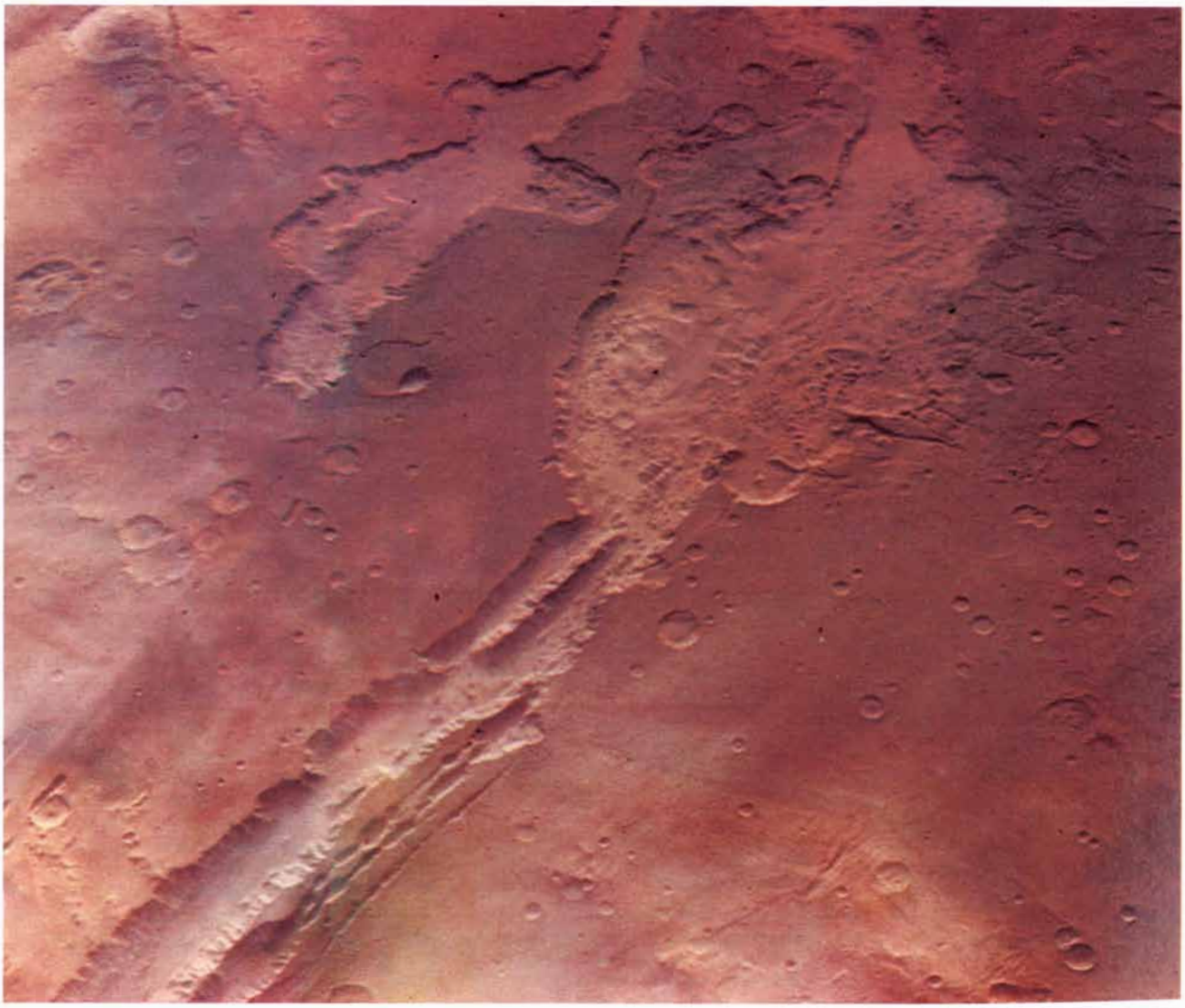
For example, from the larger craters something can be learned about the impacting meteorites themselves. If one plots the number of craters on a lunar mare with respect to their size, one finds that the number of craters is roughly proportional to the inverse square of their diameter: if there are a certain number of craters larger than 10 kilometers across, a fourth of them will be larger than 20 kilometers across, a ninth will be larger than 30 kilometers across and so on. Hence it can be said that the primary craters are distributed in a spectrum of sizes that follows a power law with an exponent of -2. If a graph of the number of craters with respect to their size is plotted on log-log paper, the graph is an almost straight line with a slope of -2 [*see illustration on page 93*].

Such a graph has several interesting consequences. First, the mass of the meteorite needed to form a crater of each size can be calculated approximately from test impacts on the earth. From those calculations the spectrum of crater sizes can be converted into a spectrum of meteorite diameters. One finds that the diameters of the meteorites that made the craters are also distributed in a

© 1976 SCIENTIFIC AMERICAN, INC



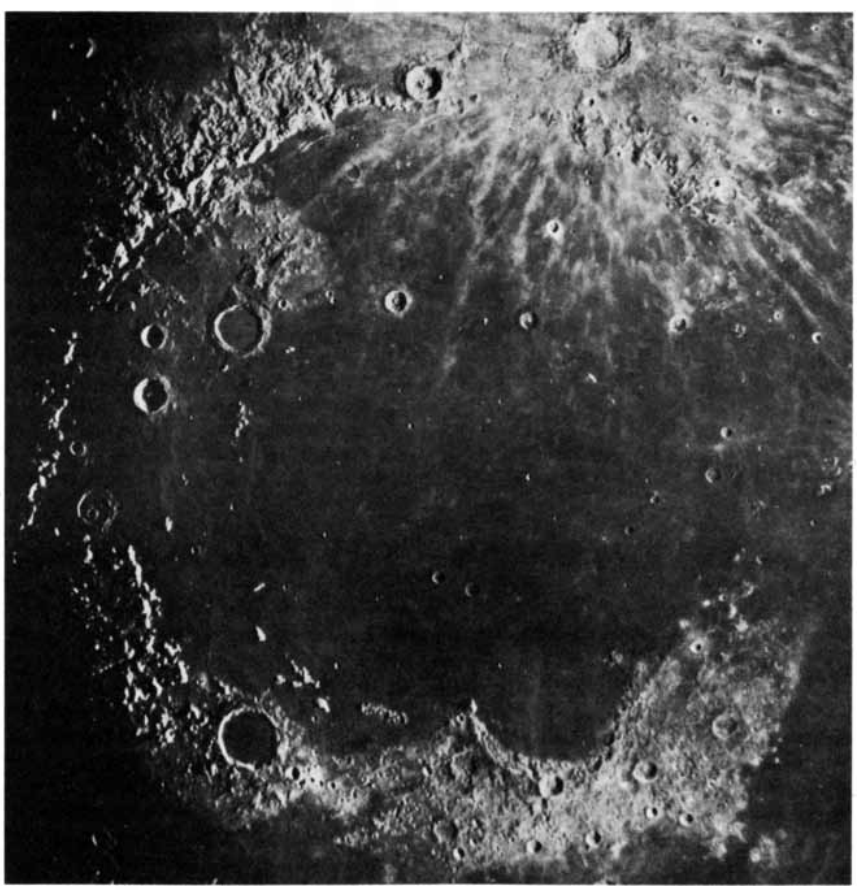
EARTH CRATER in Saskatchewan was photographed in false color by Skylab astronauts. It is Deep Bay Crater at south end of Reindeer Lake. It is nine kilometers across and 100 million years old. Most terrestrial craters are quickly worn away by geological activity, but many Canadian craters are in an area that may have been protected by sediments removed by glaciers. Round lake is inside crater rim.



MARS CRATERS appear in this color mosaic made by the U.S. Geological Survey in Flagstaff, Ariz., from 15 frames of imagery sent back by the *Viking 1* orbiter. The craters, their profiles degraded by erosion and the deposition of sediment, are the round objects. The canyons encroaching on them are part of the great Valles Marineris system near the Martian equator. Picture is 1,800 kilometers across.

© 1976 SCIENTIFIC AMERICAN, INC

CRATERS IN THE HIGHLANDS OF THE MOON were photographed by the *Apollo 10* astronauts. The largest crater, some 40 kilometers across, is relatively fresh, and its floor is covered by deposits that have apparently slumped from its inner walls. The surrounding terrain is made up of material ejected from craters and of overlapping ancient craters in various stages of degradation. Size of smaller craters in picture continues down to limit of visibility.



CRATERS IN THE LOWLANDS OF THE MOON were photographed with the 100-inch telescope on Mount Wilson. Because this region is near the north pole of the moon the original photograph showed the floor of the large circular basin occupying most of the picture substantially foreshortened. It is seen here from vertically overhead because the image of the basin was projected onto a globe and then rephotographed from directly above the basin floor. The basin is Mare Imbrium, the largest well-preserved impact structure on the moon. It is believed to have been formed by an asteroid that was between 100 and 150 kilometers in diameter. The basin was later filled by lava flows, and newer craters of various ages formed on the lava floor. At upper right is the large, prominent crater Copernicus. Long bright "rays" extending away from it are made up of material thrown out of the crater by the impact of the body that made it.

spectrum following a power law with an exponent of about -2.

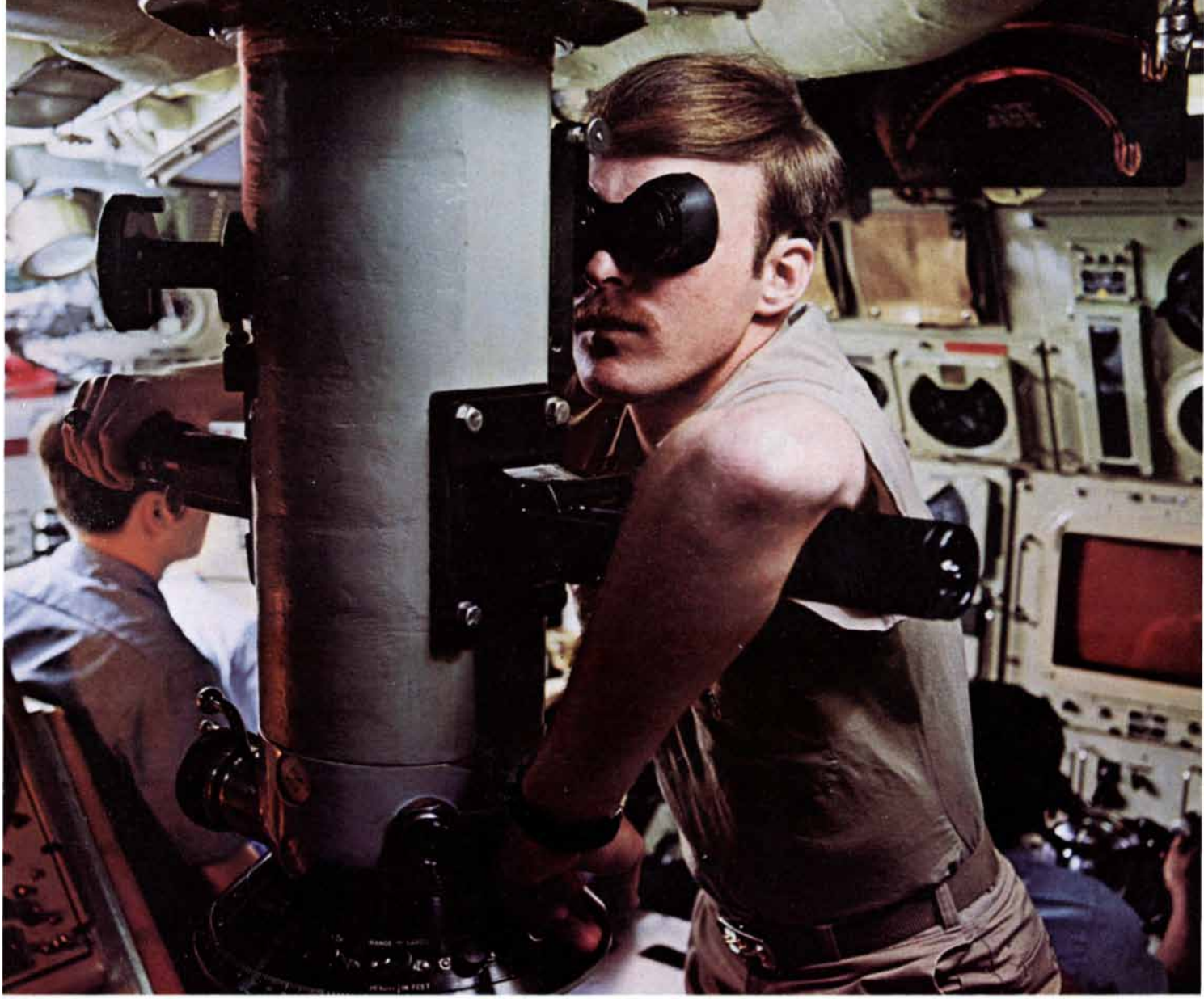
It turns out that the spectrum of sizes of the meteorites that formed the craters on the moon and the planets is essentially the same as the spectrum of sizes of both the asteroids in the asteroid belt and the meteorites that fall on the earth. This finding is strong confirmation of the hypothesis that the great majority of the craters on the planets and the satellites were made by asteroidlike and meteoritelike interplanetary bodies. Moreover, this type of mass spectrum is characteristic of the mass spectrum of fragments created when rocks collide at a high speed, say a few kilometers per second. Asteroids collide at such speeds today. Thus one can infer that the bodies that made the craters were the fragments of asteroidlike bodies that had earlier collided and broken up. Such an inference is consistent with studies of meteorites and asteroids, which indicate that there were many asteroidlike bodies in the ancient solar system. The study of craters on the moon is hence a direct link between astronomical studies of bodies in interplanetary space and geological studies of the evolution of the crust of the moon.

The spectrum of crater sizes further indicates that the impacting meteorites must have ranged in size up to bodies more than 100 kilometers across. These bodies must have looked like asteroids. If one adds up all the craters of all sizes surviving on the inner planets, one finds that the number of impacting meteorites must have come to at least a few percent of the number of bodies traveling in the asteroid belt today. Many additional objects, however, must have made craters that have since been obliterated by erosion, lava flows and so on. It is estimated that for all the inner planets to have been cratered to the same degree as the moon, about half the number of the bodies in the present-day asteroid belt would have been required. That number may have represented only the last fraction of the interplanetary material involved in the formation of the planets.

One can investigate older, more cratered regions of the moon's surface, regions outside the maria, by combining studies of lunar craters with analyses of the lunar rocks. The astronauts gathered samples from several geological provinces, some much older than the maria. By counting the number of craters in a province whose age is known from the rocks, it is possible to determine the rate at which the moon was cratered at different eras in the history of the solar system.

This approach reveals that earlier than four billion years ago the rate at which the moon was cratered was hundreds or even thousands of times greater than the rate at which it is being cratered today. Moreover, the rate declined rap-

© 1976 SCIENTIFIC AMERICAN, INC



NAVY NUCLEAR PROPULSION. THE FASTEST WAY UP IN NUCLEAR ENGINEERING.

If you want to get into nuclear engineering, start by getting into the Nuclear Navy.

The Navy operates more than half the reactors in America. So our nuclear training is the most comprehensive you can get. You start by earning your commission as a Navy Officer. Then we give you a year of advanced nuclear technology, training that would cost you thousands if you could get it in graduate school. During your career, you'll get practical, hands-on experience with our nuclear powered fleet. Maybe you'll work on a nuclear submarine, maybe a nuclear cruiser. But wherever you work, you'll really get to prove your worth—as a young Nuclear Propulsion Officer entrusted with the most advanced technical equipment known to man.

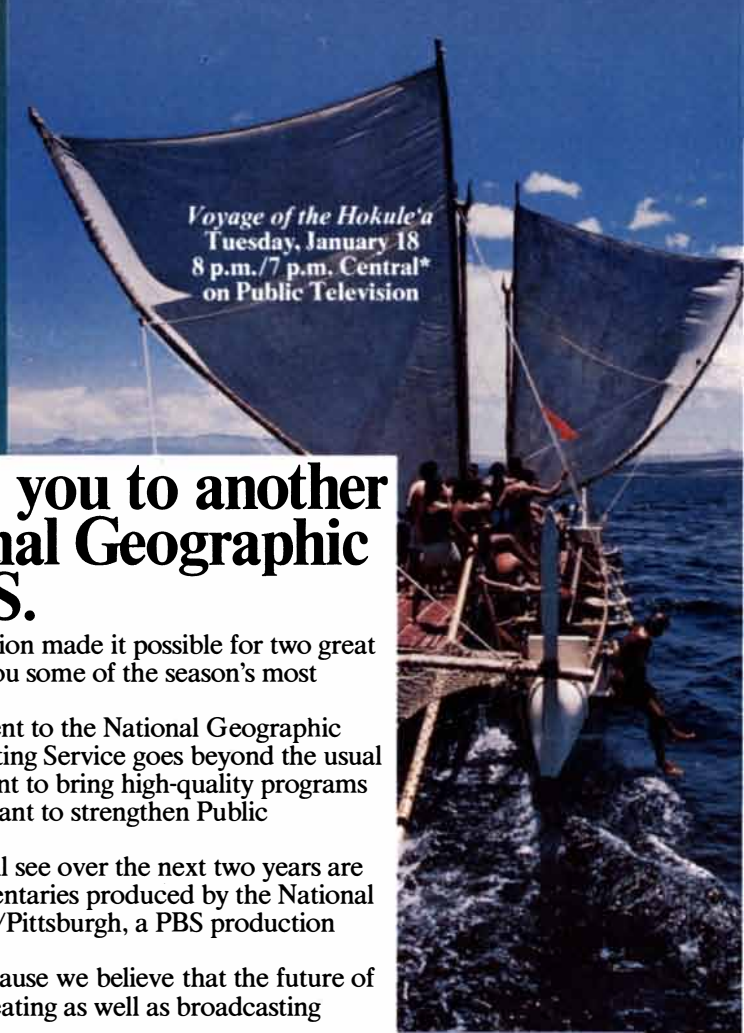
If that sounds like the kind of responsibility you're looking for, speak to your Navy recruiter. He can tell you if you qualify as a Nuclear Propulsion Officer Candidate. Or call toll free **800-841-8000.** (In Georgia, 800-342-5855.)

Navy Nuclear Propulsion Officer. Some men wait for the future. He lives it now.

NAVY OFFICER. IT'S NOT JUST A JOB, IT'S AN ADVENTURE.

© 1976 SCIENTIFIC AMERICAN, INC

Treasure!
Tuesday, December 7
8 p.m./7 p.m. Central*
on Public Television



Voyage of the Hokule'a
Tuesday, January 18
8 p.m./7 p.m. Central*
on Public Television

Gulf welcomes you to another exciting National Geographic Season on PBS.

Last year Gulf Oil Corporation made it possible for two great American institutions to bring you some of the season's most exciting television.

Gulf's three-year commitment to the National Geographic Society and the Public Broadcasting Service goes beyond the usual corporate grant. We not only want to bring high-quality programs to the American audience; we want to strengthen Public Television, too.

The eight new specials you'll see over the next two years are original American-made documentaries produced by the National Geographic Society and WQED/Pittsburgh, a PBS production center.

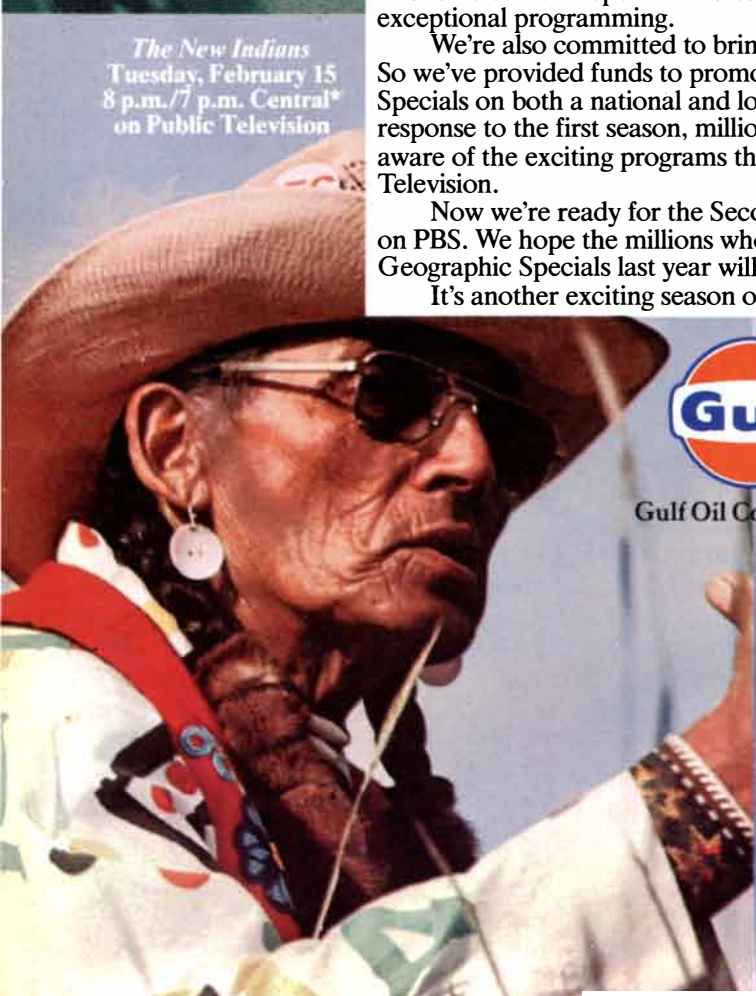
Gulf funded the project because we believe that the future of Public Television depends on creating as well as broadcasting exceptional programming.

We're also committed to bringing a wider audience to PBS. So we've provided funds to promote the National Geographic Specials on both a national and local level. Judging from the response to the first season, millions of Americans are now more aware of the exciting programs they can see only on Public Television.

Now we're ready for the Second National Geographic Season on PBS. We hope the millions who enjoyed the National Geographic Specials last year will be back this year.

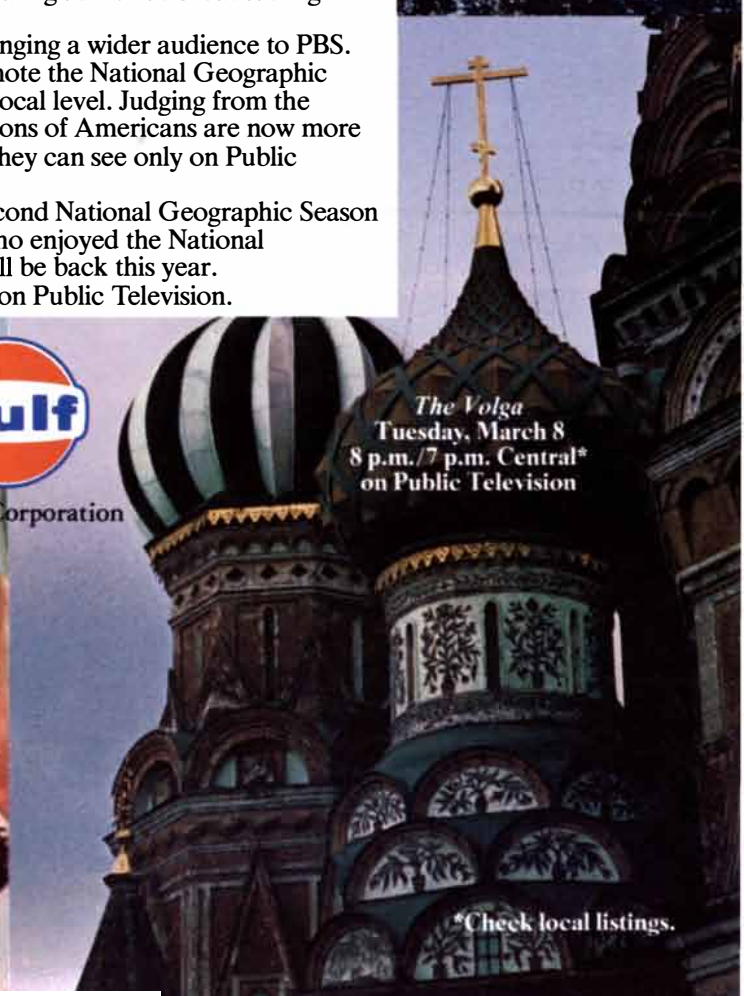
It's another exciting season on Public Television.

The New Indians
Tuesday, February 15
8 p.m./7 p.m. Central*
on Public Television



Gulf

Gulf Oil Corporation



The Volga
Tuesday, March 8
8 p.m./7 p.m. Central*
on Public Television

*Check local listings.

© 1976 SCIENTIFIC AMERICAN, INC

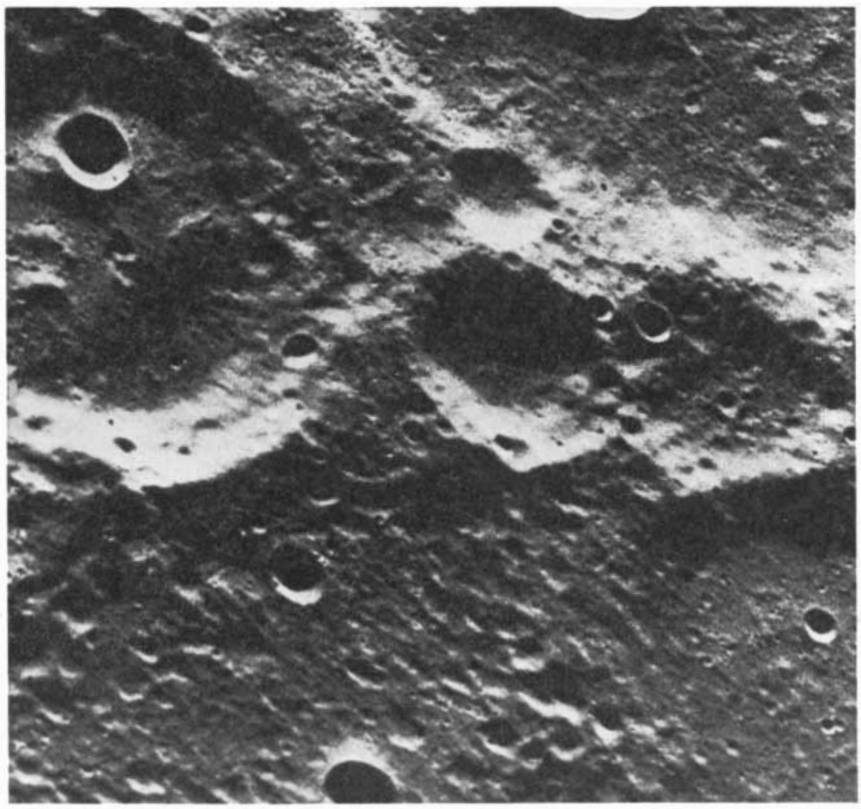
idly from that early time until about three billion years ago, when it reached the current level. The rate has been relatively constant ever since. It is believed that the other planets and satellites experienced a similar sequence.

Why were the planets so heavily bombarded early in the history of the solar system? According to one hypothesis the answer is directly related to how the planets and their satellites came into existence. The planets formed when matter in the cloud of gas and dust surrounding the primordial sun gradually collected into large solid objects. As the planets and satellites accumulated they must have been intensely bombarded by remaining objects both large and small. Even when 99 percent of the primordial material was concentrated into large bodies, the remaining 1 percent would have been frequently deflected by the gravitational field of those bodies and would have been perturbed into orbits that could intersect the orbit of a planet and collide with it.

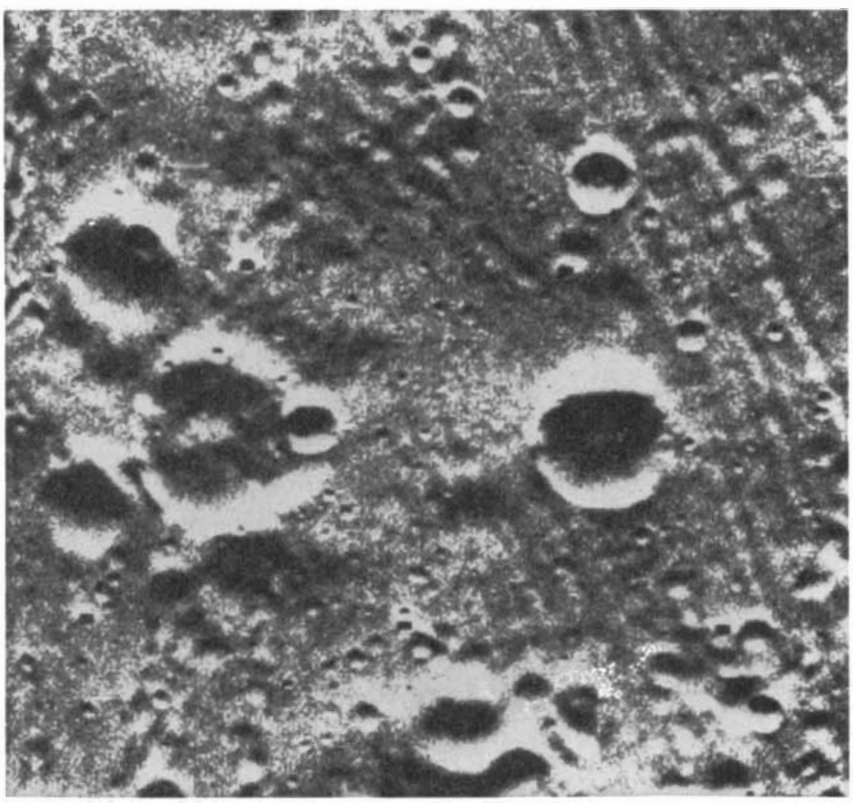
Apparently the period during which the planets were heavily bombarded was the final stage in the sweeping up of the debris left over from the primordial nebula. Dynamical calculations show that more than four billion years ago the half-life of the debris, that is, the time required for the planets to sweep up half of it, might have been as short as a few million years, as is indicated by the rapid decline in the early cratering of the moon. The material would therefore have been thinned out in a relatively short time. Today the typical half-life of bodies in trajectories that intersect the orbits of planets is closer to 100 million years or even longer. As those bodies are swept up new material is continually supplied by perturbations and collisions of bodies in the solar-system "reservoir," the asteroid belt.

The dating of lunar rocks has led to an alternative hypothesis of what caused the intense early cratering. Various Apollo analysts have found that the age of many samples of lunar rocks cuts off rather sharply at four billion years. Few older rocks have survived. Gerald J. Wasserburg and his colleagues at the California Institute of Technology have suggested that a short cataclysmic episode of intense bombardment four billion years ago destroyed older rocks and structures on the surface of the planets. This episode may have been the result of the breaking up of a large interplanetary body that passed close to the earth and was torn apart by tidal forces. G. W. Wetherill of the Carnegie Institution of Washington's Department of Terrestrial Magnetism has calculated that such bodies might have been deflected from the region of the giant planets in the outer solar system into the inner solar system about four billion years ago.

The truth may lie between the two hy-



CRATERS ON MERCURY, photographed by the spacecraft *Mariner 10*, strongly resemble those in the highlands of the moon in being in various stages of degradation and overlap. Shapes of the craters are distorted out of round because they were photographed at an oblique angle.



CRATERS ON PHOBOS, the larger of the two moons of Mars, also resemble those in the lunar highlands. This picture was made on September 18 from a distance of 800 kilometers by the *Viking 2* orbiter. Grooves were an unexpected discovery. They may be fractures created by an impact on Phobos or by an impact on a hypothetical parent body of which Phobos was a part.

© 1976 SCIENTIFIC AMERICAN, INC

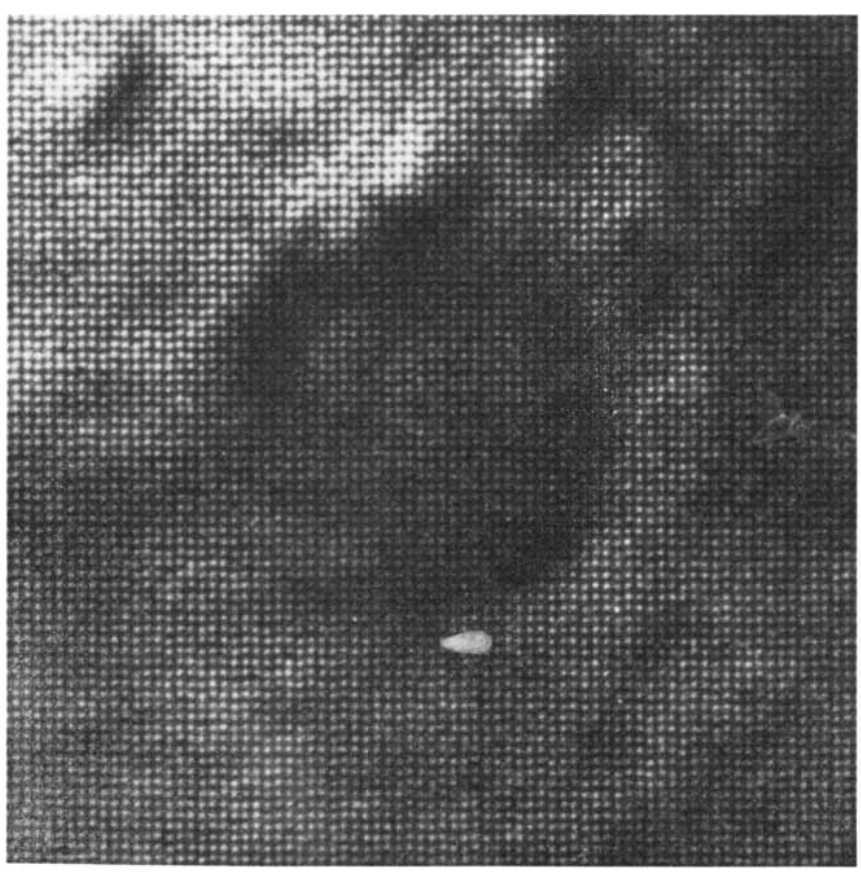
potheses. The declining curve plotting the rate of cratering almost certainly has "sawteeth" in it, representing bursts of groups of fragments from disrupted asteroids or comets. Whether or not such sawteeth were of cataclysmic proportions, dramatically affecting the evolution of a planet's surface, is uncertain.

Any model of the primitive geological structure of the planets must take into account their history of being cratered. For example, Robert B. Hargraves of Princeton University and others have recently made progress in interpreting the Precambrian geology of the earth, invoking purely internal processes to account for the breakup of a hypothetical homogeneous crust underlying a global sea. This crust is pictured as forming between four and 4.5 billion years ago, and its breakup is needed to explain early continental evolution. An external process with the same result is cratering. On the moon basins hundreds of kilometers across excavated materials to a depth of at least several kilometers and piled them elsewhere earlier than four billion years ago. The same must have happened on the earth, giving rise to a large-scale heterogeneity in the earth's crust.

Many if not all of the planets began their history with a heavily cratered surface. Evidence comes from the ancient, light-colored highland areas of the moon, which are a maze of overlapping craters. According to data from lunar rocks, the highland crust formed between 4.4 and 4.5 billion years ago when heat (generated by an unknown source) melted the moon's outer layers, eventually creating a surface of igneous rock. This was during the period of intense bombardment, so that the primitive crust was quickly pulverized by overlapping craters. During the first few hundred million years after the planets formed, enough craters evidently accumulated on them to cover not just a few percent of their surface but 100 percent of it. Such a surface is said to be saturated with craters. Since on a saturated surface each new crater would on the average destroy the equivalent of one old crater the same size, the surface is also said to be in cratering equilibrium.

The geological consequences of saturation are important. The inverse-square power-law spectrum of crater sizes on the moon has the property that if there are enough primary craters between one kilometer and 10 kilometers in diameter to cover an entire planet, there are also enough craters between 100 and 1,000 kilometers across to cover it. The impact of a meteorite large enough to create a crater 100 kilometers across would have excavated the planet's surface to a depth of several kilometers. Thus whereas a planet 10 percent saturated with multikilometer craters would have a surface consisting of primeval rock strata interrupted by craters, a planet totally saturated with craters would have a surface consisting of pulverized debris several kilometers deep. Such a surface layer is called a megaregolith.

In all probability such a megaregolith accumulated on the moon during its first few hundred million years. The total cratering was probably enough to saturate the surface not once but many times over, since presumably the entire moon was built up by infalling bodies. This conclusion is supported by the fact that the lunar craters are in various states of degradation, ranging from fresh-looking craters with sharp rims to battered remnant ones overlain by many other craters and their ejecta. All the original igneous rock on the moon would have been excavated, pulverized and redistributed by the cratering. A vertical section through the crust would show a megaregolith consisting of overlapping layers of compacted ejecta that were heated and partly welded together and then perhaps broken up again. Such a formation would explain why the astronauts in the lunar highlands found not the well-preserved primordial rocks that some geologists expected but brecciated rocks consisting of materials that had been fragmented and recemented. It might also explain some of the seismic properties of the moon, which rule out the existence of layers of solid rock near much of the lunar surface.

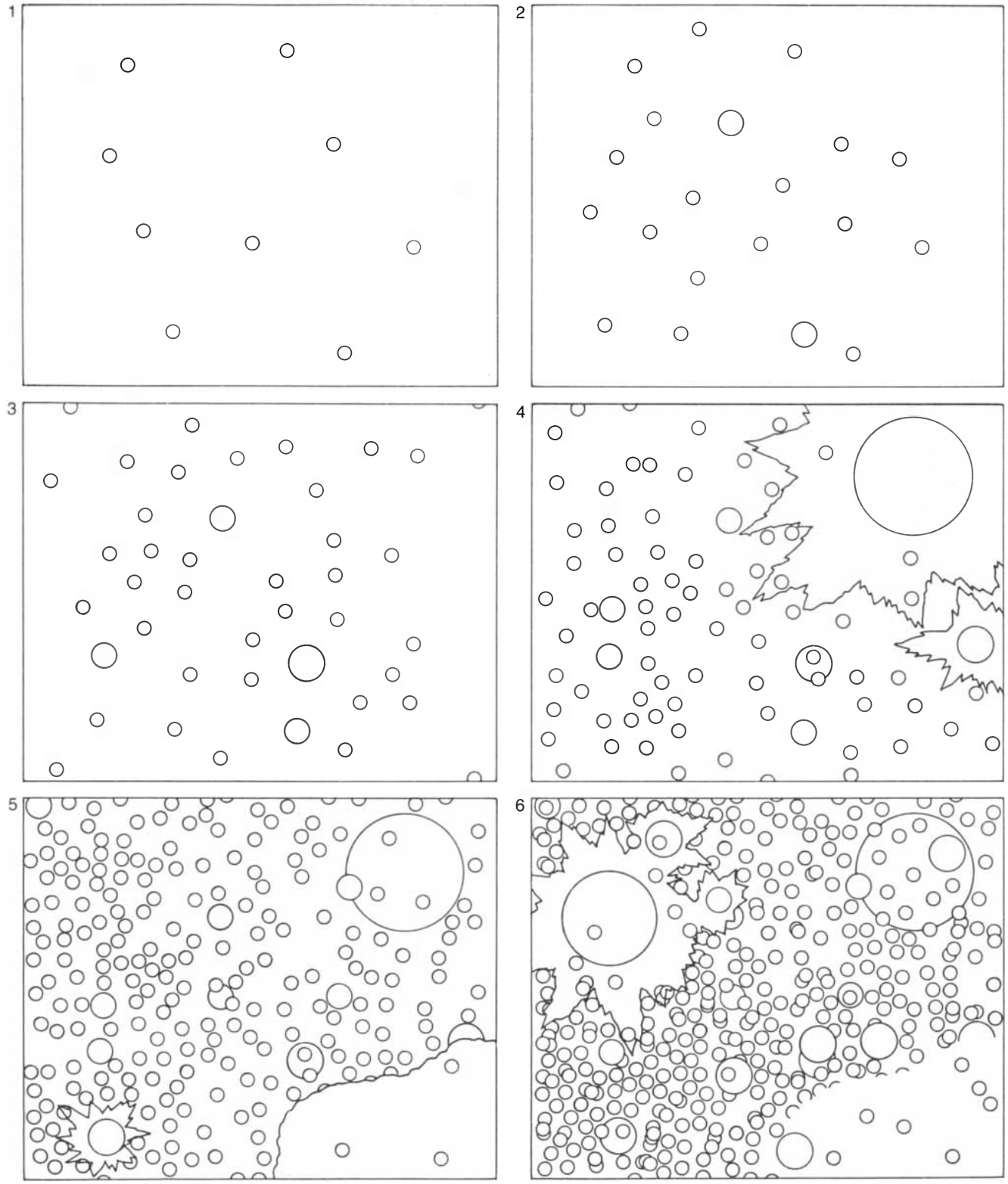


CRATERS ON VENUS are revealed through the cloudy atmosphere of the planet by a radar image obtained by Richard M. Goldstein and his colleagues at the Jet Propulsion Laboratory. Although this type of image makes craters appear rugged, their relief is actually quite low. Largest crater in the middle is 160 kilometers in diameter but its rim is only 500 meters high.

If there had been no lava flows, no tectonic activity, no atmosphere and no oceans on any body in the inner solar system, the surface of all its planets would look like the crater-saturated lunar highlands. On the larger planets, however, volcanic and erosive processes as well as continued cratering have erased the old craters. The number, size and condition of the craters on the various inner planets are different, suggesting different erosive conditions, different ages and perhaps even different meteoritic environments. Each planet has a cratering "signature" that yields clues to its history.

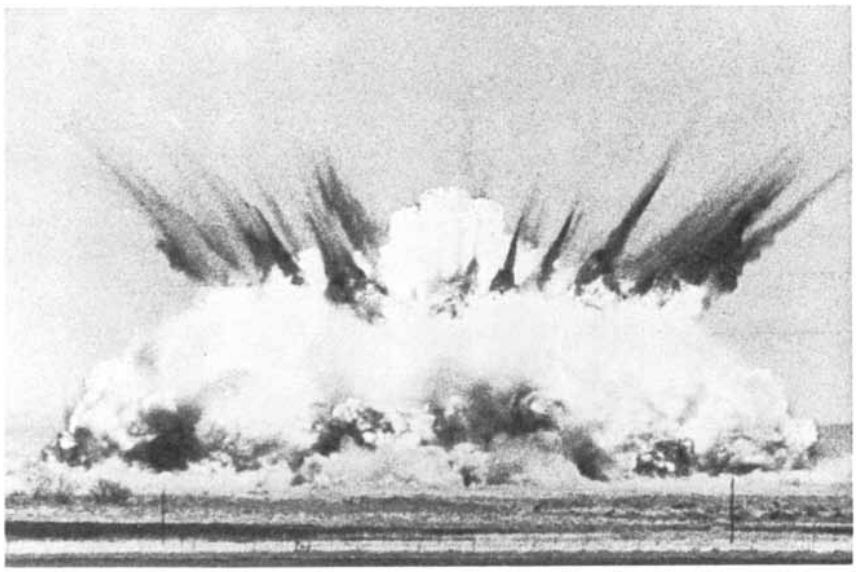
The two moons of Mars, Phobos and Deimos, are the only satellites in the inner solar system apart from the earth's moon. The density of craters on their surface is similar to that in the lunar highlands. The two satellites are believed to be too small to show any internal or surface activity that would erode craters, although there are linear grooves on Phobos that remain unexplained. (They may be impact-generated fractures.) Phobos and Deimos prob-

© 1976 SCIENTIFIC AMERICAN, INC

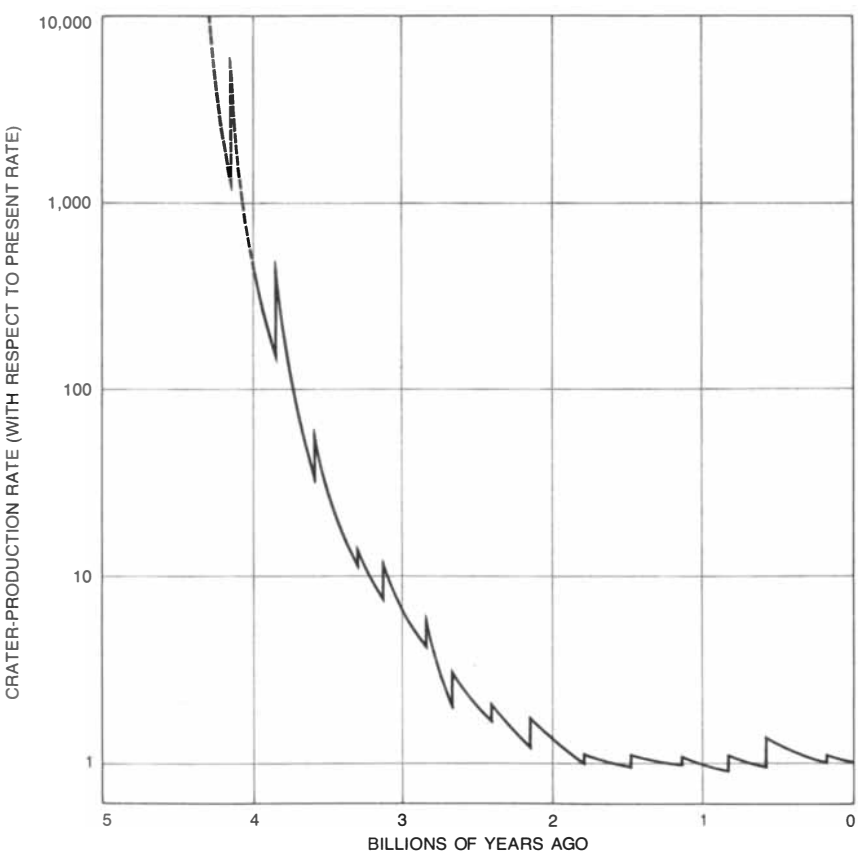


NUMBER AND SIZE OF CRATERS yield information on the bodies that formed them. In this schematic drawing the surface of a hypothetical planet is shown at six different epochs in its history starting a short time after the formation of a fresh surface (*1*). When actual craters are counted and measured in an undisturbed area on a planet, the cumulative number of craters of each size has been found to agree with the number predicted if the craters were made by impacting bodies having the same distribution of sizes as the bodies in the asteroid belt. Thus it can be inferred that the population of objects that made the craters was much like the population of asteroids today. The distribution of craters also provides information on the age of the different geological provinces of a planet or a satellite. In young provinces the craters are distinct from one another on a craterless background (*first three panels*); in old provinces new craters form on older ones and tend to obliterate them (*last three panels*). At times, however, a new province can be created when a blanket of ejecta from a gigantic impact covers old craters (*4*) or a flow of lava wipes the slate clean (*5*). After a certain point the oldest terrain becomes nearly saturated with craters and approaches a cratering equilibrium in which every new crater will on the average destroy an older crater of the same size. Nevertheless, inhomogeneities in the cratering record introduced by past processes of obliteration can still be detected (*6*).

© 1976 SCIENTIFIC AMERICAN, INC



TEST EXPLOSION of 100 tons of TNT in a Colorado desert illustrated some characteristics of crater formation. The material in high-speed jets at angles between the horizontal and the vertical was thrown clear of the crater; it is this kind of material that forms the long, bright rays surrounding some of the large new lunar craters such as Copernicus. This particular explosion excavated a crater 128 feet across and 23 feet deep. Blocks of material thrown out of primary crater made smaller secondary craters as far as 360 feet away from the center of the explosion.



RATE OF CRATERING HAS DECLINED over the past 4.5 billion years, according to the number of craters on the moon in different provinces whose ages have been established by analyses of lunar rock samples. The moon was intensely bombarded early in its history when the larger bodies in the solar system were sweeping up the last of the debris left over from the system's formation. Spikes in the declining curve mark succeeding times (of unusually intense bombardment) of new material being injected into inner solar system either when larger bodies collided and broke up or when material was deflected from outer solar system by the gravitational field of the giant planets. Precise distribution and magnitude of spikes is not known.

ably bear the scars of billions of years of impacts, perhaps mostly of errant bodies from the asteroid belt just beyond the orbit of Mars. They may themselves be fragments of a larger body that was broken up by some catastrophic cratering event.

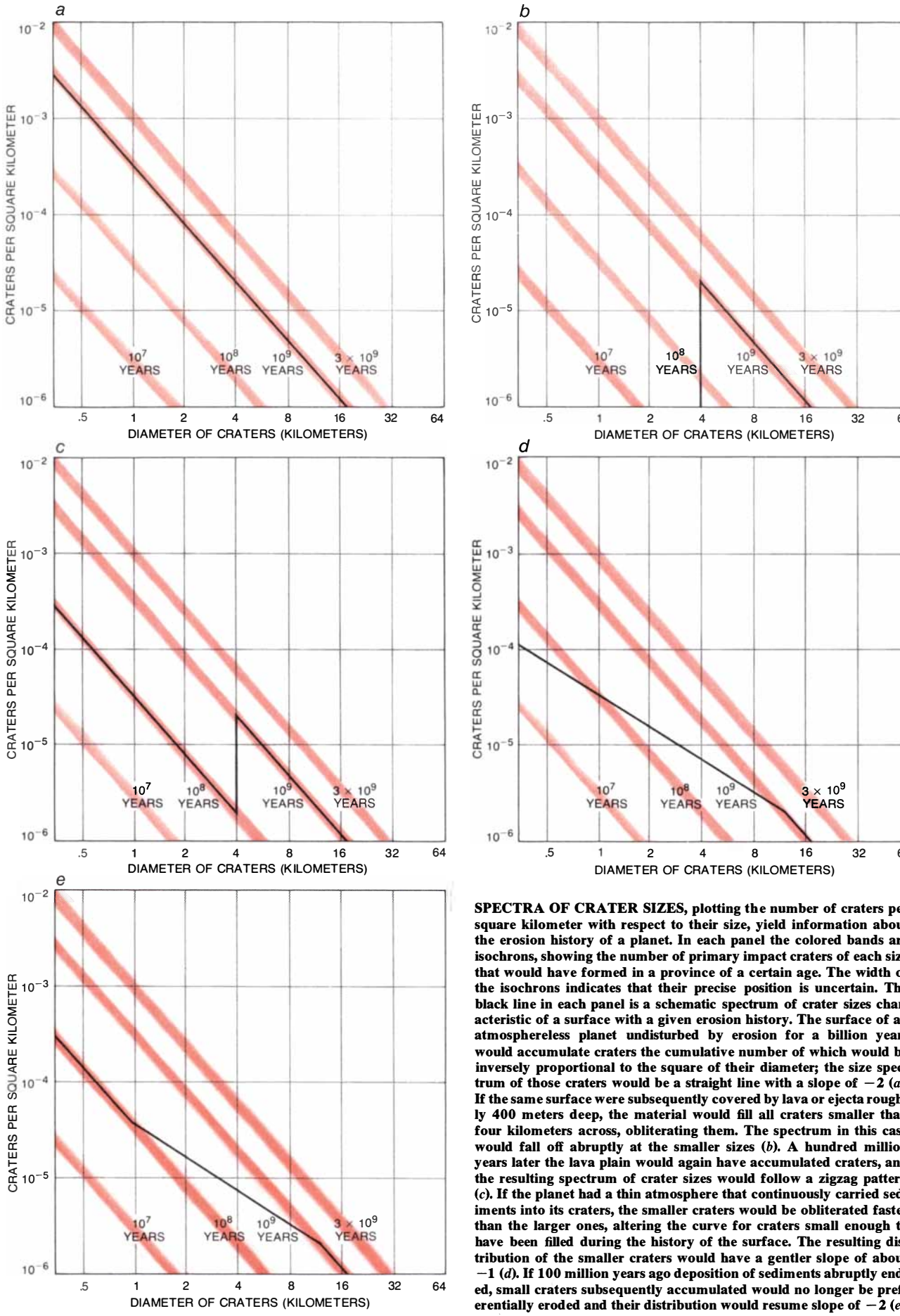
Mars itself presents a very different picture. The planet has been under intensive study since *Mariner 9* photographically mapped its entire surface in 1971 and 1972. The most striking finding is that the planet is roughly divided into two hemispheres, one sparsely cratered and the other heavily cratered. In the sparsely cratered region there are smooth plains—apparently young lava flows—dotted with scattered, fresh-looking craters. Analysis of the number and size of the craters reveals that the conditions on the plains of Mars are similar to those on the maria of the moon. The spectrum of sizes of the Martian craters follows the familiar power-law distribution with an exponent of about -2, which matches the law found for the spectrum of the masses of asteroids. Therefore the impacts of asteroids seem to have been recorded and preserved in the youngest and least disturbed areas on Mars.

If we widen our scope to include all of Mars, however, we easily find evidence of disturbance. Photographs show that there are many Martian areas, particularly near the poles, where thick layers of sediment, perhaps windblown, have been deposited. In other areas the layers of sediment are being stripped away, revealing older craters underneath. On the floor of several craters sand dunes are accumulating. Winding channels or arroyos seem to indicate that the surface of the planet was once eroded by running water.

Crater studies have been a key means of obtaining clues to the processes of erosion and deposition on Mars. If one counts craters in the ancient Martian highlands, where many craters appear to be eroded or filled with sediment, one finds that their size spectrum differs markedly from the simple power law of -2 characteristic of the size spectrum of craters in most regions on the moon or on the Martian plains. When the size spectrum of the craters in the Martian highlands is plotted on log-log paper, it shows a curious kink, as if the curve were made up of two segments following a power law with an exponent of -2 joined by an intervening segment following a power law with an exponent of about -1.

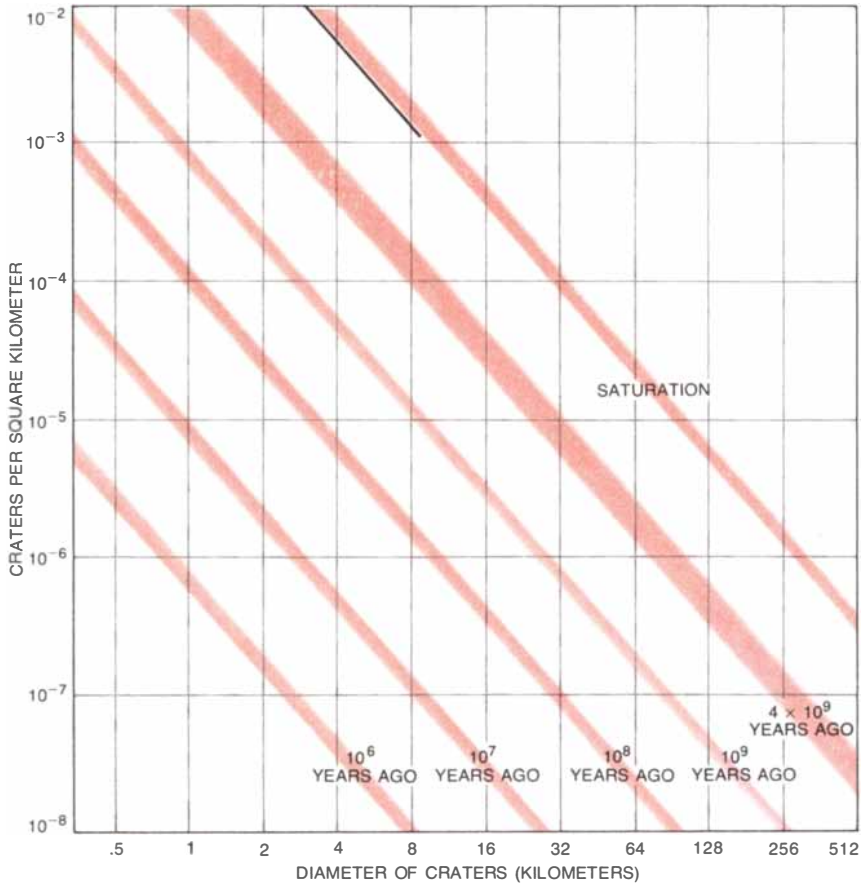
Such a kink might be explained if Mars had had a specific kind of erosion history. What appears to be indicated is a long period of erosion and deposition that ended fairly abruptly in relatively recent Martian history. Such deposition would have filled small, shallow craters, giving them lifetimes shorter than those

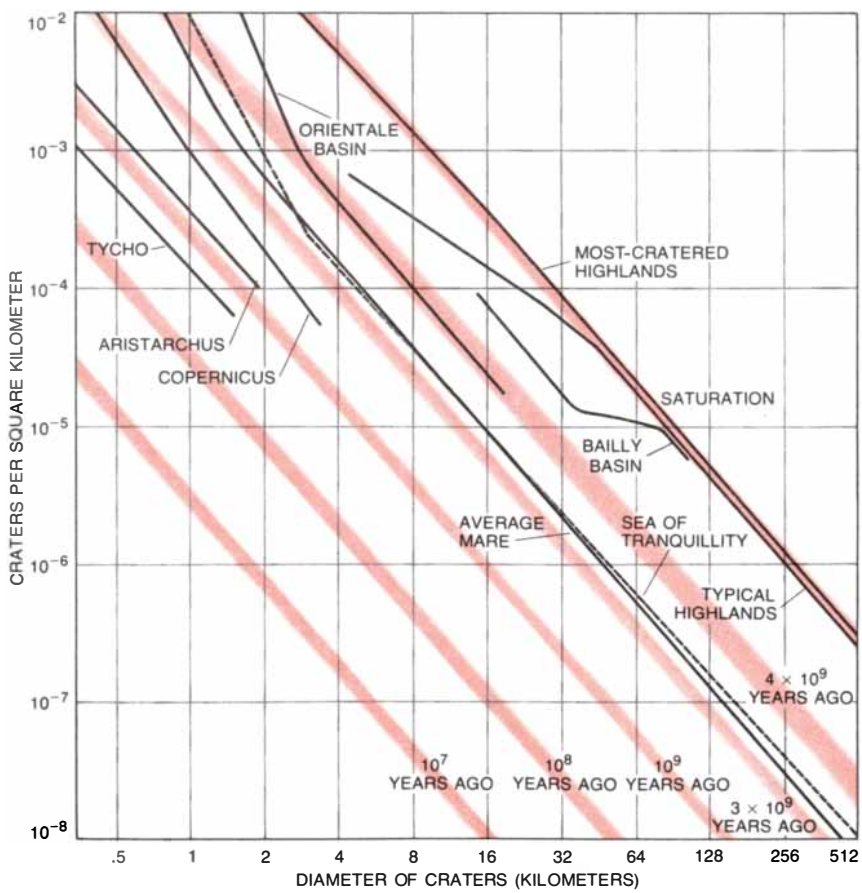
© 1976 SCIENTIFIC AMERICAN, INC



SPECTRA OF CRATER SIZES, plotting the number of craters per square kilometer with respect to their size, yield information about the erosion history of a planet. In each panel the colored bands are isochrons, showing the number of primary impact craters of each size that would have formed in a province of a certain age. The width of the isochrons indicates that their precise position is uncertain. The black line in each panel is a schematic spectrum of crater sizes characteristic of a surface with a given erosion history. The surface of an atmosphereless planet undisturbed by erosion for a billion years would accumulate craters the cumulative number of which would be inversely proportional to the square of their diameter; the size spectrum of those craters would be a straight line with a slope of −2 (*a*). If the same surface were subsequently covered by lava or ejecta roughly 400 meters deep, the material would fill all craters smaller than four kilometers across, obliterating them. The spectrum in this case would fall off abruptly at the smaller sizes (*b*). A hundred million years later the lava plain would again have accumulated craters, and the resulting spectrum of crater sizes would follow a zigzag pattern (*c*). If the planet had a thin atmosphere that continuously carried sediments into its craters, the smaller craters would be obliterated faster than the larger ones, altering the curve for craters small enough to have been filled during the history of the surface. The resulting distribution of the smaller craters would have a gentler slope of about −1 (*d*). If 100 million years ago deposition of sediments abruptly ended, small craters subsequently accumulated would no longer be preferentially eroded and their distribution would resume slope of −2 (*e*).

© 1976 SCIENTIFIC AMERICAN, INC





of larger craters. The larger the crater was, the greater its chance of survival would have been. The steady-state size spectrum of craters under such conditions would have a slope of about -1, as it does on Mars. In order to explain the observations the lengthy period of erosion and deposition must have ended suddenly, because continued cratering has reestablished the slope of -2 at the small sizes.

In other words, the Martian craters seem to tell a tale of intense erosion and deposition at some time in the past, a period with a rather abrupt ending. Numerous investigators, including Raymond E. Arvidson of Washington University, Clark R. Chapman of the Planetary Science Institute, Kenneth J. Jones of Brown University, Laurence A. Soderblom of the U.S. Geological Survey and me, have studied Martian craters and have affirmed this general picture. The different states of degradation of Martian craters, which range from fresh-looking bowl-shaped craters to dust-filled flat-floored ones, are compatible with a period of intense erosion. There is some disagreement, however, about whether there was only one period of erosion or many periods and whether most of the erosion was ancient or relatively recent. Very ancient erosion would agree with evidence obtained recently by the Viking spacecraft that in primordial times the atmosphere of Mars was more massive and therefore denser than it is now. Recent episodes of erosion, if they occurred, might imply that the Martian climate oscillates like the climate of the earth, with its ice ages and intervening warm periods. Michael C. Malin of Cal Tech has used crater counts to estimate that the Martian channels, which may be relics of the erosion period, are billions of

SPECTRA OF CRATER SIZES on the satellites Phobos and Deimos (*top*) and the moon (*bottom*) reveal that the smaller bodies of the inner solar system have been little modified by erosion. The isochrons (*color*) represent the predicted number of primary craters in provinces of different ages. The band labeled "Saturation" indicates the spectrum of crater sizes expected on a body the surface of which is saturated with craters. The solid black lines plot the number of craters observed in different provinces on the Martian satellites and the moon. Phobos and Deimos appear to be nearly saturated, so that their surface is very ancient. The moon, however, clearly has provinces of different ages. The curve labeled "Typical highlands" shows that some small craters in some of the highlands have probably been obscured by lava. The unusually steep curves at the top of the spectrum indicate that secondary impact craters and volcanic craters are probably included in the counts of primary craters in certain areas. The zigzag curve for Bailly Basin may indicate that smaller craters in the basin were once obliterated by ejecta.

© 1976 SCIENTIFIC AMERICAN, INC

years old. If they are, it would favor the idea that erosion was associated with a primordial dense atmosphere.

On the earth, of course, erosion is much more intense than it is on any of the other inner planets. The earth is so active that in geological terms craters last a very short time. Most of the surviving terrestrial craters are heavily eroded, and many of them have been recognized by geologists only in recent years. One geological province where the search for craters has been particularly fruitful is the Canadian shield, the region around Hudson Bay with stable formations about a billion years old. In that billion years, a third the span of the lunar maria, many craters were excavated by the impact of meteorites. The smaller were rapidly eroded. The larger, some tens of kilometers in diameter, were planed down by glaciers and partly filled with sediments. Some are being exhumed by the erosion of old sediments, and their remains have been identified in increasing numbers by Canadian geologists such as C. S. Beals and Michael R. Dence and their colleagues. Similar features have been found in younger geological provinces, including nonmountainous parts of the U.S.

If the number of craters per square kilometer in terrestrial cratered provinces is divided by the mean ages of the provinces, one can find the rate at which the craters have been produced over the past billion years. The crater-production rate on the Canadian shield agrees roughly with the rate derived for the past three billion years from the lunar plains. A few investigators, comparing the two rates in detail, have suggested that the average rate over the past billion years has been somewhat higher than the rate over the entire three billion years. If that is so, it would suggest that the rates vary with time as asteroids or comets supply new meteorite fragments to the inner solar system.

Even the relatively stable Canadian shield clearly reveals the effects of erosion on the population of craters, which allows one to confirm the validity of the technique of inferring erosion processes from peculiarities in the spectrum of crater sizes. Small Canadian-shield craters have been so quickly worn away that the size spectrum is strongly deficient in small craters. Thus an observer on another planet, armed only with photographs from an artificial-satellite mission, would be able to conclude that most of the earth's surface formed in recent geologic times (on the basis of the paucity of impact craters in general) and that processes of erosion and deposition are actively destroying small geological structures today (on the basis of the deficiency of small craters in particular). For example, crater statistics derived from various regions on the earth indicate that typically a terrestrial crater 100 meters in diameter has a life of a few thousand years, a crater one kilometer in diameter a life of a million years, a crater 10 kilometers in diameter a life of a few hundred million years and a crater 100 kilometers in diameter a life of a few billion years.

Venus appears to present a picture intermediate between that of Mars and the earth. Richard M. Goldstein and his associates at the Jet Propulsion Laboratory of Cal Tech have obtained radar images of the surface of Venus hidden by the cloud layers, and the images show craters. Only a small fraction of the planet has been mapped, but the preliminary crater counts give two impressions. First, craters larger than 100 or 200 kilometers in diameter are nearly as abundant on Venus as they are in the older regions of Mars, although they are less abundant on Venus than they are in the highlands of the moon. They are about 10 times more abundant on Venus than they are on the Canadian shield of the earth. Second, the population of craters less than 100 kilometers in diameter is strongly depleted, probably by erosion. The erosion hypothesis is supported by radar findings that the craters are quite shallow and may be partly filled with dust.

One can infer that the surface of Venus is neither as old as the surface of the moon nor as geologically active as the surface of the earth. The present surface of Venus must have formed at an intermediate time in the past, and erosion may be removing small craters today. This inference is consistent with the pictures sent back by the Russian Venus landers in 1975. The pictures show sharp, angular rocks at one landing site and smooth, rounded rocks at another, implying that rocky material has been both created and eroded in the recent geological past. The implication that Venus, which is nearly the same size as the earth, has preserved more ancient large craters and has had perhaps less tectonic activity calls for explanation by general theories of planetary evolution.

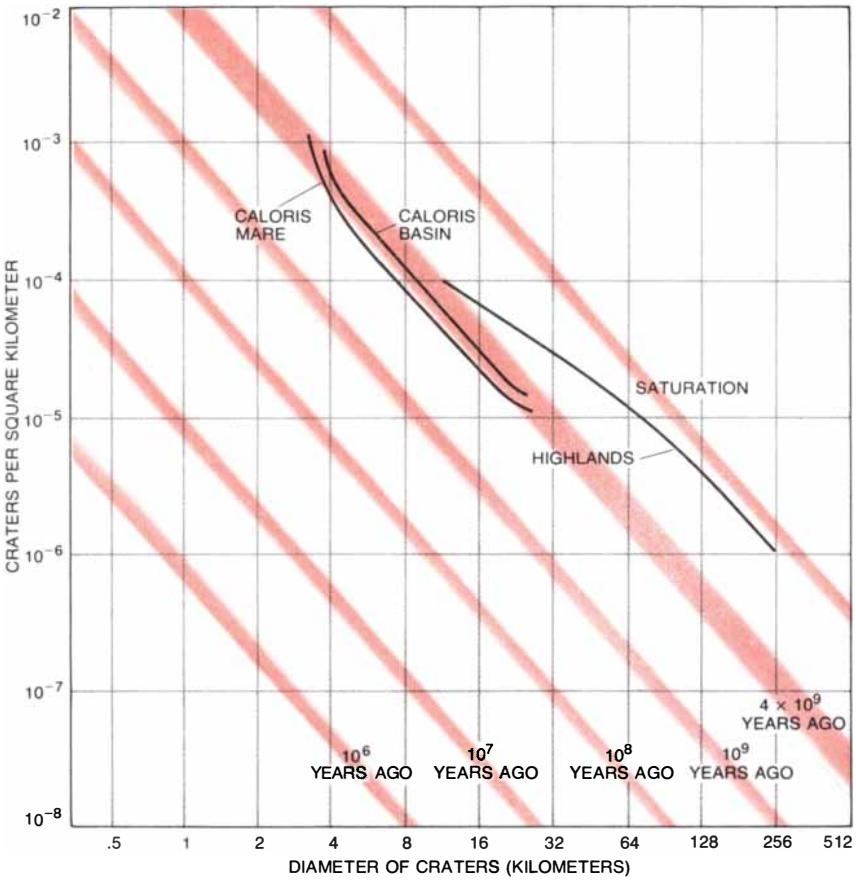
Mercury, the last of the inner planets, at first seems to be almost identical with the moon. The pictures from *Mariner 10* show heavily cratered highlands and dark smooth plains resembling the lunar maria. The craters on Mercury range from pits at the limits of resolution (about a kilometer) to the Caloris Basin 1,300 kilometers across, surrounded by fracture rings as much as 2,000 kilometers across. Closer examination reveals, however, that the planet differs from the moon in three important respects. First, the surface of Mercury has compression faults hundreds of kilometers long that seem to imply that the entire crust of the planet has shrunk in circumference by about two kilometers. Second, even in some of the heavily cratered regions the surface is not saturated with craters; on the contrary, the craters seem to have pocked a relatively smooth background plain. Third, Mercury shows a slight deficiency of small craters with respect to the number that would be expected if the craters had accumulated at the same rate as they have on the moon.

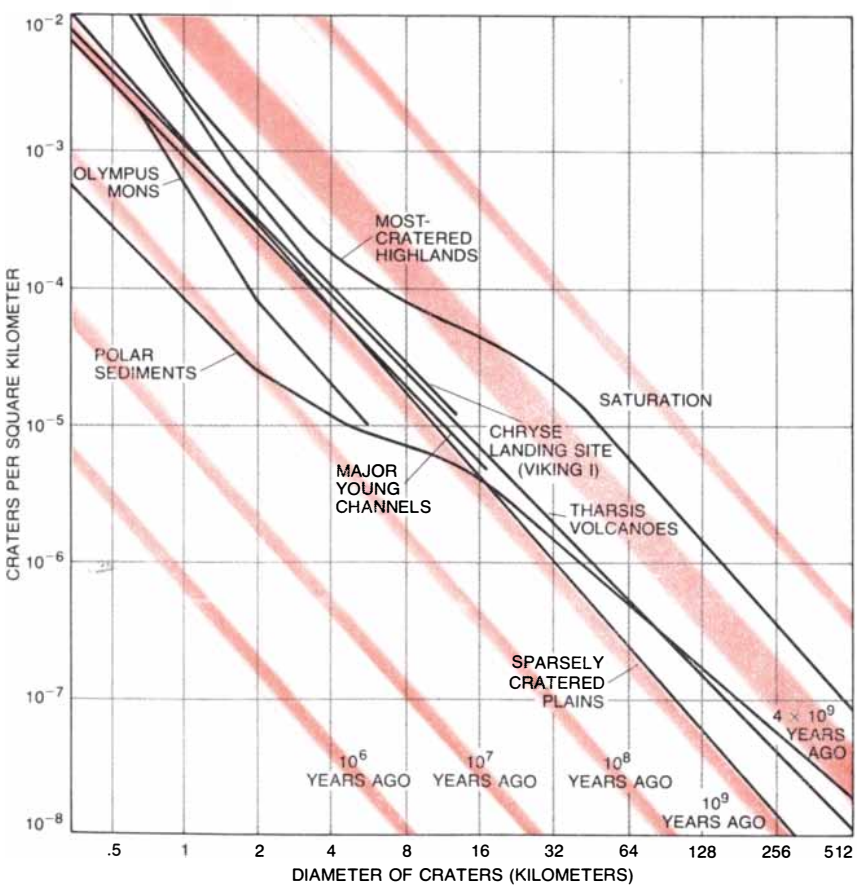
One interpretation of these observations has to do with the fact that the decay of radioactive elements within a planet causes it to heat up with the passage of time. It may be that because Mercury is larger than the moon it heated up more slowly, and that its crust cooled later in the period of intense bombardment. Therefore the final configuration of Mercury's surface, whether it was dominated by craters or by a smooth melted crust, would have been the result of competition between the rate at which the craters were formed and the rate at which the crust solidified to the point where it could retain craters.

Suppose Mercury still had a molten crust at the end of the period of intense early bombardment, say 4.1 billion years ago. Small early craters would not have been preserved. Larger ones would have relaxed into shallow, isostatically compensated ghosts of their original selves. If the crust had solidified soon thereafter, craters would then have begun to be recorded on the smooth background, but there would not have been enough of them to saturate the surface. Therefore the faults that fractured the surface when the crust cooled and shrank would not have been obliterated by cratering. Residual melting or thin lava flows in some areas might, however, have obliterated some of the smaller early craters. The three ways in which Mercury differs from the moon would thus be explained. Robert G. Strom of the University of Arizona and other investigators have discerned traces of these processes on the moon, where they are less apparent because the lunar crust solidified earlier and therefore was more affected by impacts and cratering.

The addition of Mercury to the suite of studied planetary surfaces has forced planetary geologists into interpretive discussions of the earliest cratering history of planets, and such discussions are straining the limits of the evidence. There are many explanations other than the one I have given. For example, some analysts of the data from *Mariner 10* argue that the same episode of cataclysmic cratering hypothesized for the moon affected Mercury, coming after the smooth crust on Mercury had formed and giving rise to most of the visible craters in a very short interval. Other investigators, comparing Mercury and other planets, have proposed that the deficiency in the number of small craters arises not because the small craters were obliterated but because they were never formed. They contend that the lack of small craters reflects actual differences in the size distribution of the

© 1976 SCIENTIFIC AMERICAN, INC





impacting meteorites at different eras and places in the solar system. Even the traditional interpretations of cratering equilibrium and the saturation of a planetary surface have been questioned. All these doubts remain to be resolved.

The cratering of the inner planets and their satellites and what the craters tell us about the geological conditions on the surface of those bodies can be summarized in the order of the bodies' mass. Phobos and Deimos are tiny worlds that are completely cratered. The moon is heavily cratered, with only 15 percent of its most ancient surface obliterated by the recent lava flows represented by the maria. Mercury is much like the moon, but its internal thermal activity lasted longer and left more traces. Half of the old surface of Mars has been modified by volcanism, and riftlike canyons and fractures suggest that the planet has experienced some rudimentary tectonic activity. Venus has scattered large craters, but radar evidence for a riftlike canyon and crater filling suggest that it is currently being modified by tectonic and erosion activity. The earth is so active that it no longer has any craters as old as four billion years and only a few highly eroded craters as old as one billion years. There seems to be a trend to the effect that the more massive a planet is, the more internal energy it generates and the longer it maintains processes such as volcanism and tectonic activity. It may be that only the earth has full-fledged plate tectonics, with thermal currents in the mantle driving plates into each other and crumpling the crust into mountain ranges unlike those on any other planet. Although the craters on the earth appear to be little more than superficial scars, they confirm a fundamental sequence of planetary evolution that might be less evident if they were lacking.

There is an additional impact-related effect that could be of major importance in determining the character of a planet. This effect has to do with the statistically small number of very large bodies striking each planet in primordial times. Any

MERCURY AND MARS crater-size spectra indicate that the surface of these larger bodies has been more highly modified with time. Crater counts for sample areas on Mercury (*top*) indicate that most of Mercury's surface is old. A deficiency of small craters in Mercury's highlands may indicate that the surface of the planet was in a plastic state when it was intensely bombarded, in which case the smaller craters would have disappeared. The curves in the spectrum for Mars (*bottom*) show that the oldest cratered provinces experienced an intense period of erosion and deposition that ended a few billion years ago. The Martian volcanoes and channels are younger. Polar regions appear to have experienced a later period of erosion and deposition that may have ended as recently as 100 million years ago.

© 1976 SCIENTIFIC AMERICAN, INC

population of meteorites that craters a planet will consist of many small bodies, a few medium-size bodies and one largest body. This is in the nature of the fragmentation processes that govern the size distribution of asteroids and meteorites. The largest impact basins on the moon, Mercury and Mars are some 1,000 kilometers across, and the bodies that made them must have been about 100 kilometers in diameter. Yet as we have seen, the observed craters on the planets may be only the last of a series that extends back to the beginnings of the solar system. The earliest craters may not have survived at all.

There is good reason to believe, based on extrapolations of the density of craters on the moon together with theoretically estimated distributions of the sizes of the impacting bodies and their orbits, that the largest bodies that struck each planet were 1,000 kilometers or more in diameter. An asteroid 1,000 kilometers across striking a primordial planet could have given rise to a fundamental asymmetry in the planet, perhaps by knocking the crust off one side. Although traces of this cataclysmic event could have been obliterated by cratering or tectonic processes, the underlying asymmetry could have been preserved in the planet's geology. Such a collision could explain the hemispheric asymmetry of the moon, whose crust on the side facing the earth is thinner than that on the side facing away. The same kind of collision might be involved in the asymmetry of Mars, where one hemisphere has many ancient craters and the other has been almost entirely modified by volcanism. It might finally be involved in the early breakup of the earth's crust into thick and thin units.

The impact of large asteroids may well explain the fact that the character and evolution of the planets are less uniform than one would expect them to be if the planets had built up in a purely evolutionary way by the accretion of myriads of small particles. The largest interplanetary bodies probably carried so much energy and momentum that, depending on the direction from which they approached a planet, they could have tilted it, speeded up its spin, slowed down its spin, destroyed a satellite or perhaps even have left rings of material around it after breaking up under gravitational forces.

So far I have described how cratering can yield clues both to the external environment in which the planets formed and to the surface and internal processes that have modified the planets. The study of craters can also contribute to determining how old the surface of a planet is. On any given planet the relative ages are clear, since heavily cratered areas are older than sparsely cratered ones. It is also possible to estimate absolute ages. If one knows the

KUWAIT UNIVERSITY

ACADEMIC POSTS AVAILABLE FOR 1977/78

IN THE BOTANY DEPARTMENT in the following fields: taxonomy of flowering plants (for two years only) Plant Pathology-Microbial Genetics.

QUALIFICATIONS:
- Must hold Ph.D. degree or its equivalent in the respective specialization.
- Must currently hold academic post in an accredited University or research center.

CONTACT the Kuwait Embassy Cultural Division for further information and application forms:
4340 Connecticut Ave. N.W.
Suite 500
Washington, D.C. 20008
Tel. 202-244-4709

If outside the United States, please contact your nearest Kuwait Embassy.

PROFESSIONAL DISCOUNTS

Texas Instruments electronic calculator

TI-30	$21.95	SR-51-II	56.95
1600	21.95	SR-56	84.50
1650	26.95	SR-52	199.95
2550-II	28.50	PC-100	162.50
SR-40	32.50	TI-5050M	89.95
Bus Anst.	32.50	5040PD	119.95
5100	49.95		
Libraries	26.00-39.95		
Little Prof. (2 min.)	$16.95		

HEWLETT hp PACKARD
(No credit cards)

HP-21	$71.00	HP-27	$154.00
HP-22	109.95	HP-25C	175.00
HP-25	127.60	HP-67	395.00
		HP-97	664.95

NOVUS ≈ NATIONAL SEMICONDUCTOR

Quiz Kid II	$17.95	Statistician PR 6035	$39.50
Math 4510 (2 min.)	17.88	NS-852	24.95
Scientist 4520	29.95	NS-4515 PR	39.50
Scientist PR4525	54.95	NS-4640	54.95
Financier 6020	26.45	NS-4650	44.95
Statistician 6030	26.45	NS-4660	54.95
Financier PR6025	39.50	Adversary TV Game	74.95

In Calif. call (213) 370-5795 or ORDERS ONLY CALL (800) 421-0367 (other than CA.) Send money order. Pers. ck (2 wks. to clear); in CA. add 6% sales tax. Add $3.50 min. shipping charges US only; $7.00 for SR-52, PC-100, HP-67, -97. Canada min. shipping charges $6.50. WE SHIP UPS AIR. Send mail orders to DEPT. SA-A.

tk enterprises
BETTER BUSINESS PRODUCTS
16611 Hawthorne Blvd., Lawndale, Ca. 90260
(213) 370-5795 **(800) 421-0367**

The Tao of Physics

An Exploration of the Parallels Between Modern Physics and Eastern Mysticism

In an easy-to-read but penetrating style, Dr. Fritjof Capra explores the parallels between modern physics and Eastern mysticism. A clear presentation, supplemented by diagrams and photographs, explains the theories of atomic and subatomic physics, of relativity theory, and of astrophysics, up to and including the most recent research, and relates the world view emerging from these theories to the mystical traditions of Hinduism, Buddhism, Taoism, Zen and the I Ching. A book of interest to both the scientist and the layman.

Available at all fine bookstores or direct from the publisher (add 50¢ shipping):

Cloth 12.50 Paper 5.95

Shambhala Publications, Inc.
1123 Spruce Street
Boulder, Colorado 80302

Distributed to the book trade by Random House

FREE catalog over 2000 unique tools, handy kits, precision instruments, technical supplies.

Our 23rd year of service to the World's finest craftsmen and technicians.

National Camera 2000 West Union Ave. Dept. JAA Englewood, Colorado, 80110 Phone (303) 789-1893

Send a FREE copy of the NC Flasher

name

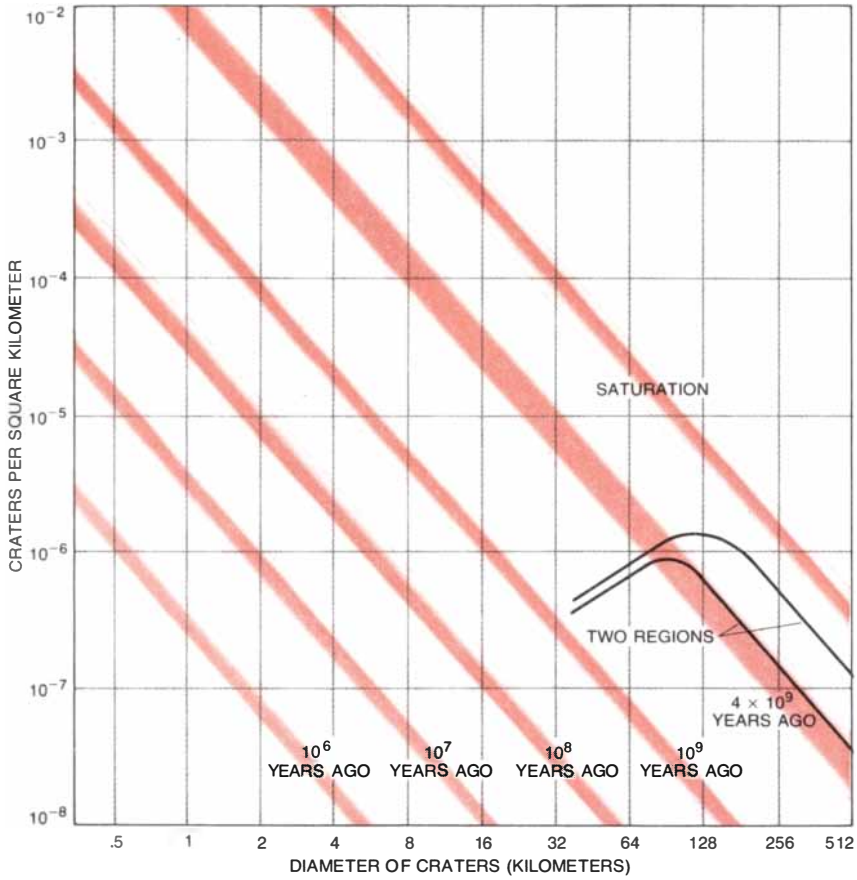
address

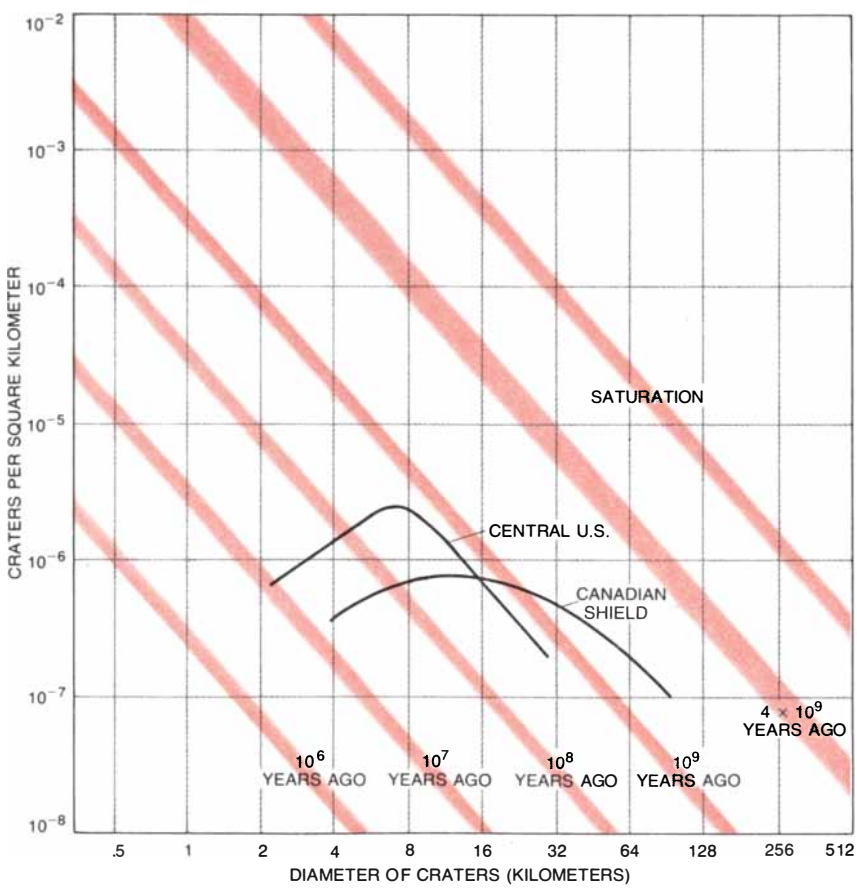
city

state zip

National Camera 2000 West Union Ave. Dept. JAA Englewood, Colorado, 80110 Phone (303) 789-1893

© 1976 SCIENTIFIC AMERICAN, INC





rate of crater production during each billion years of a planet's history, and if one counts the number of craters now found per square kilometer on the surface of different provinces with different erosion histories, one can directly calculate the absolute age in years of those provinces by dividing the number of craters by the crater-production rate. This crater-retention age differs in character from the age determined by measuring the decay of radioactive isotopes in surface rocks. The radiometric ages refer to specific events in the history of certain rocks, such as origin by solidification or metamorphism. The crater-retention ages refer to how long a surface can retain a topographic structure of any given size. The crater-retention age depends on the diameter of the crater, since a small crater is likely to be eroded or obliterated faster than a large one. In short, crater-retention ages measure levels of geological activity.

If one knew the exact rate of crater production for a planet, one could calculate the absolute crater-retention age for each province on the planet, since heavily cratered provinces are older than sparsely cratered ones. The trouble is that the crater-production rate may have been different in different parts of the solar system. The mere fact that we have absolute measurements of the crater-production rate on the earth and on the moon does not automatically yield the crater-production rate for the other planets. Therefore we lack definitive absolute chronologies for Mercury, Venus, Mars, Phobos and Deimos.

The best way to approach the problem of absolute dating is to compute for the known asteroids and comets the rate at which they collide with various planets today. That rate turns out to be highly dependent on the orbit of those bodies. For example, an asteroid the orbit of which extends from the asteroid belt to a point just inside the orbit of Mars may have a high probability of eventually hitting Mars and making a crater on it. There is also a chance that the asteroid will just miss Mars and be deflected into an orbit crossing that of the earth, so

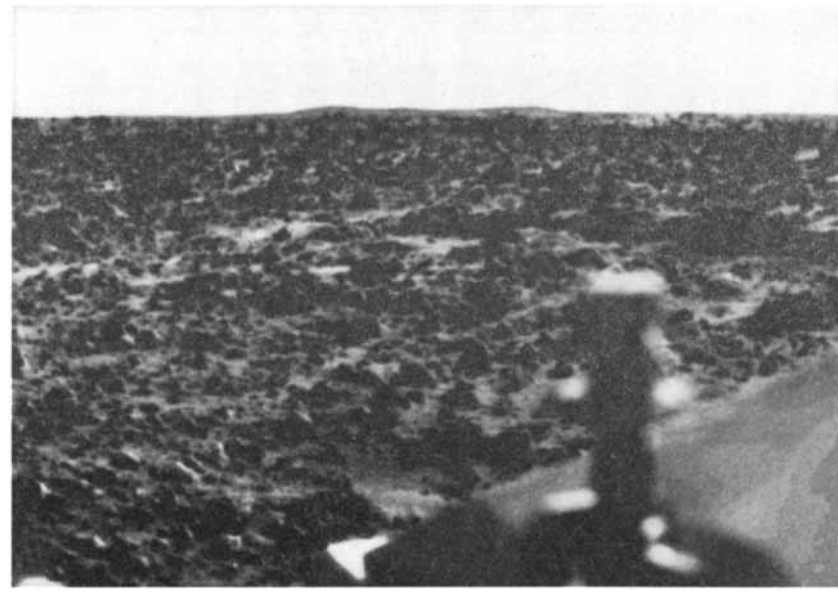
VENUS AND EARTH crater-size spectra show that the two largest inner planets are geophysically the most active. The spectrum for Venus (*top*) indicates that craters 32 kilometers across may have survived only the past billion years, whereas craters larger than 128 kilometers across could date back four billion years. Craters smaller than 32 kilometers across are not shown because the radar images of Venus' surface do not have sufficient resolution for them to be detected. On the earth size spectrum of craters (*bottom*) shows that number of craters that have survived is a function of their diameter. Even very large craters have not survived from earliest days of the earth's history. On the earth erosion conditions are so intense that craters are obliterated much faster than they are formed.

© 1976 SCIENTIFIC AMERICAN, INC

that it eventually will make a crater on the earth or on the moon. In other words, the body's initial orbit has a strong influence on the probability of its colliding with one planet or another. Wetherill and others have taken this probabilistic approach, and they have found that the crater-production rate at any one of the inner planets or satellites is likely to be within a factor of 10 of the rate at any other. The rates may even agree within a factor of three. Similar results have recently been reported from a survey of asteroids by Eugene M. Shoemaker and Eleanor F. Helin of Cal Tech.

To a terrestrial geologist used to discriminating a unit of rock 60 million years old from one 70 million years old an error in age by a factor of three may sound like little improvement over no dates at all. One must remember that a planetary geologist is faced with the problem of discriminating a unit of rock a million years old from one a billion years old, a range of a factor of 1,000. Moreover, since no planet is believed to be older than 4.6 billion years, and since the cratering rates earlier than four billion years ago were extremely high and declined quickly over a short period of time, uncertainties in dating a unit by its craters are reduced; the most highly cratered units are probably between four and 4.5 billion years old.

Apart from the earth-moon system we have no fossil stratigraphic sequences and no radiometric rock ages, only pictures of cratered surfaces. Accordingly it would be an exciting advance to be able to exploit craters to correlate the histories of the various planets and gain an absolute dating system for the entire solar system, even if the system suffered from errors of a factor of two or three. Such a system would enable us to compare the rate at which each of the planets evolved. Moreover, it might be an aid to answering dozens of fascinating questions. For example, did all the planets form marelike lava plains before three billion years ago as the moon did? Or were periods of volcanic activity dispersed in time? Was the sequence of events governed purely by the mass of the planet and the thermal conditions within it? Did the intensity of the cratering decline at the same time throughout the solar system? Is there evidence for a short burst of catastrophic cratering that extended throughout the entire inner solar system at some particular time, as some observers have hypothesized? Are the periods of erosion and possibly of a warmer climate on Mars synchronized with periods of warmer climate on the earth? Are climatic fluctuations on all the planets synchronized? If they are, they might be related to fluctuations in the intensity of radiation from the sun, a discovery that would be of intense interest to stellar astronomers and to evolutionary biologists.



GROUND PHOTOGRAPH ON MARS was made from *Viking 1* lander. On horizon is a ridge that looks remarkably like the profile of Meteor Crater in Arizona. It may be an impact crater.

The evidence we have in hand is not enough to answer these questions definitively. The best estimates of crater-production rates and accurate crater counts made from photographs, however, suggest the following chronology. All the inner planets and satellites were heavily cratered between four and 4.5 billion years ago. Lava flows were common on the moon and Mercury three to four billion years ago, but they have not occurred since. On the moon the youngest large craters visible to the unaided eye from the earth, such as the bright-rayed crater Tycho, may be only a few hundred million years old. The cratered regions of Venus also appear to show large impact structures that are several billion years old; smaller craters 30 kilometers across or less may date back only a billion years, since older ones would have been eroded.

Mars presents a different picture. The cratered provinces appear to retain large ancient craters that may be about four billion years old, like those in the cratered highlands of the moon and Mercury. Martian craters smaller than about four kilometers across have a short life because of erosion. Most of them are probably younger than three billion years. The less cratered hemisphere of Mars was apparently later resurfaced by lava from intense volcanic activity. At about that time, possibly one to two billion years ago, water may have flowed in some of the large Martian arroyos. Then as recently as a billion years ago or even a few hundred million years ago huge volcanoes such as Olympus Mons gave rise to lava flows that have not been carved by channels. After the period of volcanic activity the most active erosion ceased, although the Martian winds continued to transport dust and deposit sediments. Some of the youngest major geological units on Mars are laminations of sediment at the poles, which appear to have a crater-retention age as low as 100 million years.

It appears that all the planets were intensely cratered during the first half billion years after they were formed, and that then volcanism became the dominant process of their geology. The smaller bodies, such as the moon and Mercury, cooled most rapidly and on them the volcanism ceased. On the larger planets internal sources of heat remained important. Volcanism, tectonic activity and atmospheric erosion have partly destroyed the primitive cratered crust on Mars and Venus, and have almost completely destroyed it on the earth.

A definitive chronology for the solar system will have to await the return of datable rock samples from many provinces on each of the planets. Such samples would enable investigators to determine the history of the solar system with errors of only a few percent. Meanwhile one can obtain a preliminary chronology of the solar system both by photographically searching for asteroids to extend our knowledge of the number of interplanetary bodies and by analyzing their orbits to deduce their dynamical history. Such studies will help to provide better crater-production rates for all parts of the solar system. Armed with that knowledge we may come to better understand the early history of the solar system and the development not only of the inner planets and their satellites but also of the icy bodies of the outer solar system.

© 1976 SCIENTIFIC AMERICAN, INC

The Samaritans

Once the inhabitants of a rich and powerful kingdom, these followers of Moses now constitute perhaps the world's smallest ethnic minority. The forces that keep them isolated have also helped them to survive

by Shemaryahu Talmon

The name Samaritan is known to readers of the New Testament mainly from two texts. The more important of the two as far as doctrine is concerned is an account of Jesus' warm reception in Samaria in spite of the fact that because of religious differences the Jews normally had "no dealings with the Samaritans" (John 4:9). Far more familiar to most readers, however, is the parable Jesus related when he was asked to define the obligations of neighborliness.

A Jerusalemite bound for Jericho, he said, was robbed and left naked and near death on the road. First a passing priest from the Temple at Jerusalem and then a Levite, one of the hereditary Temple servants, shunned the victim. The calling of both passers-by of course required that they be compassionate toward the distressed and particularly toward a fellow believer. It was only when a man of Samaria, traveling far from his own land, came on the half-dead Jew that the Jew was succored, his wounds were bound up and he was carried to the shelter of an inn. There the Samaritan both paid for the Jew's immediate care and guaranteed the cost of his complete recuperation. "Which now of these three, thinkest thou," Jesus asked his questioner, "was neighbour unto him that fell among the thieves?" (Luke 10:36).

From these New Testament references one may deduce that in the time of Jesus the Samaritans were alien enough to Jewry to make the contrast between the callousness of the two supposedly pious Jews and the compassion of the Samaritan particularly pointed. A good deal more is known about the Samaritans. For example, they, like the Jews, accepted as holy writ the first five books of the Old Testament: the Pentateuch (Genesis, Exodus, Leviticus, Numbers and Deuteronomy). In fact, they still do. Although the kingdom of Samaria vanished long ago, the Samaritans still survive today as perhaps the smallest ethnic minority in the world.

Before the Six-Day War of 1967 the Samaritans, few though they were, had for years been divided into two communities. The people of the older and historically the more important of the two lived and continue to live in the city of Nablus, built in Roman times near the ruins of the biblical city of Shechem. Nablus (the word is an Arabic corruption of the Greek Neapolis, or New City) lies in a valley between the twin mountains Ebal and Gerizim. The latter mountain is held by the Samaritans to be the true site of Jehovah's sanctuary, which the Jews locate at the Temple in Jerusalem. The Nablus community had been reduced to approximately 150 early in this century; today it numbers about 250.

The Samaritans of the other group, also numbering about 250, live in suburban Holon to the south of Tel Aviv. They occupy a special quarter called in modern Hebrew Shikkun ha-Shomronim. Thus the total Samaritan population is about 500. One might think the survival of such a small ethnoreligious entity in the modern world would arouse the interest of anthropologists and sociologists alike, since it presents an opportunity to study social processes in a setting so restricted as to approximate laboratory conditions. A host of subjects for investigation come to mind: the consequences of inbreeding in such a state of reproductive isolation, the problems of acculturation, the struggle of a religious minority for the preservation of its identity—all could be investigated among the Samaritans as under a microscope. Yet comprehensive studies of the Samaritan community have still to be undertaken.

A close look at the community is also of benefit to those who are, as I am, students of theology and comparative religion. The Samaritans' beliefs represent an extreme in religious isolationism, and their personal and communal life is patterned by tradition to an extraordinary degree. An offshoot of biblical Judaism that seems to have remained in a state of arrested development for some 2,000 years, Samaritan religious tradition affords a kind of telescopic glimpse of the past: the ancient Judaism of prerabbinical times, its tenets and its way of life. Samaritan ritual and belief became virtually petrified, so to speak, at the beginning of the Christian Era; since then little innovation in thought, in literature or in social organization is known to have arisen within the community.

Before the 1967 war the Nablus and Holon components of the Samaritan community differed from each other in several respects because of their different settings. The Samaritans of Nablus, under Islamic rule (Arab, Turk and then once again Arab), lived in virtual seclusion. This was to no small extent imposed on them by the surrounding Islamic majority. To borrow the terminology of the sociologist Max Weber, they were in the truest sense a pariah people, geographically restricted to a special quarter of Nablus that was a small-scale replica of a typical medieval European Jewish ghetto.

At the same time the ghetto life of the Nablus Samaritans was not entirely a product of inimical forces; as with medieval Jewry, it was to some extent self-willed. Its isolation guarded the community against alien inroads and strengthened internal cohesion. The physical concentration of the entire group within a comparatively small area enabled its members to maintain their social and religious identity for century after century even though their number was small. Isolation also fostered communal institutions: a *kinshah,* or synagogue, a school and an internal system of jurisprudence. These institutions in turn enhanced the Samaritans' ability to preserve their exclusive religion in an unadulterated form. For example, whereas the Samaritans have used Arabic as their spoken language ever since their conquest in A.D. 632, their languages of religious worship remain Hebrew and Aramaic.

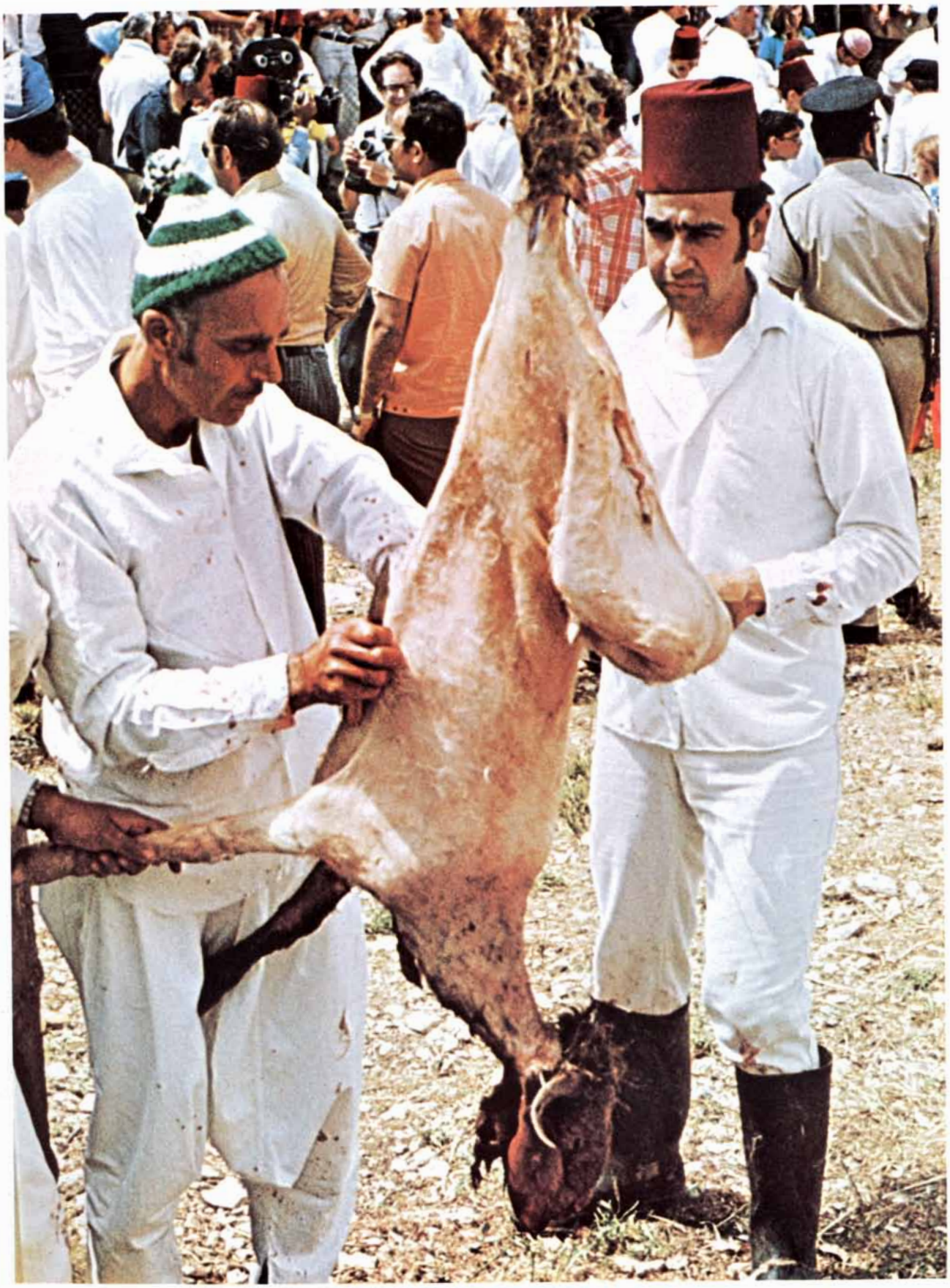
The Samaritans of Holon, unlike those of Nablus, have maintained comparatively close relations with their

© 1976 SCIENTIFIC AMERICAN, INC

SAMARITAN PILGRIMS, gathered on the summit of Mount Gerizim at dawn, are dressed in white ceremonial robes in preparation for a day of prayer leading up to the Passover sacrifice. Usually a sunset ritual, this takes place by day when Passover falls on the Sabbath.

SAMARITAN PRIEST, Berit ben Tabiah, holds up the scroll of the Torah as the pilgrims pray. The red tarboosh worn by most of the men seen here is the distinctive headgear of Nablus Samaritans; a white wrapping shows that the wearer belongs to the priestly family.

© 1976 SCIENTIFIC AMERICAN, INC



SLAUGHTERED LAMB, its fleece stripped off, is hung and gutted before being spitted for cooking. One Passover lamb is slaughtered for each Samaritan extended family; in 1975, eight years after regular contact between the two communities resumed, the number was 19.

© 1976 SCIENTIFIC AMERICAN, INC

neighbors over the past half century. The history of the Holon community goes back to the early 1920's, when a few Samaritan families left Nablus to settle in an area between the then predominantly Arab coastal city of Jaffa and the newly founded Jewish port of Tel Aviv. The immigrants numbered some 40 to 50, principally members of two extended families, the Sadakah and the Marhib, and the community remained nearly static until 1948. After the establishment of the State of Israel in that year other Samaritan families chose to leave Nablus (which lay within the region on the west bank of the Jordan River that was annexed the same year by the Hashemite Kingdom of Jordan) and settled in Holon.

The Samaritans of Holon soon became full-fledged citizens of the new state and integrated themselves into Israeli political, economic and social life. Although the immigrants are succeeding in the preservation of their religious identity, they are also gradually adjusting to the ethos of a modern westernized society. The younger generation of Samaritans is in many respects being progressively acculturated, but even it so far has resisted religious assimilation.

A visit to the Samaritan communities, either at Holon or at Nablus, shows that the Samaritans differ little in external appearances from their neighbors, Jewish or Arab. In dress, in daily habits of life and in language each community has in the main adopted the customs of the surrounding population. On taking a closer look at the Samaritans' dwellings, however, one begins to discover significant differences among Samaritans, Jews and Arabs. One example is the Samaritan interpretation of the pentateuchal admonition regarding the house doorpost inscription, the *mezuzah*. The Samaritans do not follow the Jewish practice of affixing to the doorpost a small receptacle that holds a piece of parchment bearing two passages from the Pentateuch. Instead they place near the doorpost a stone tablet; engraved on the tablet is a shortened text of the Ten Commandments. Sometimes the Samaritans even forgo the engraved tablet and simply paint a few letters on the wall next to the door, written in Samaritan Hebrew script.

This direct and almost primitive way of observing a pentateuchal command exemplifies the religious division between the Samaritans and the Jews. Both religions derive their basic ritual laws from the Pentateuch, but whereas in Judaism the Pentateuch makes up only one part of a sanctified three-part canon of Scripture, for the Samaritans the five books of Moses alone have a claim to sanctity. The Pentateuch is their entire Bible, and they have never accepted the rest of the Old Testament. The Samaritans, to be sure, have a book of Joshua, the traditional account of Israel's seizure of the Promised Land, but their book is a comparatively late compilation that presents a Samaritan interpretation of the history of Israel, beginning with the conquest of Canaan and carrying the record down to the early centuries of the Christian Era. It differs radically from the book of Joshua in the Old Testament.

In the light of the Samaritan's attitude toward the Pentateuch it is hardly surprising that they did not and still do not subscribe to the oral Jewish law as it developed in postbiblical times and ultimately became codified in rabbinical literature. The Samaritans recognize neither the Mishnah (the redaction of Jewish oral law produced about A.D. 200) nor the Talmud (a later and larger redaction that includes the Mishnah, traditionally produced in Babylonia about A.D. 500), and they ignore the later Jewish legal literature that has been derived from those sources. To the Samaritans the life of the community and the individual alike is based, at least in principle, solely on the Pentateuch. In actual practice a Samaritan system of oral exegesis, a Halakah, has arisen that differs considerably from the Jewish ritual code. The differences are apparent in everyday life and become particularly striking in such matters as the observance of the Sabbath and of various festivals.

To cite an example, the Samaritans' literal adherence to the pentateuchal text means that they have never developed anything comparable to the Jewish ritual use of phylacteries: small leather boxes containing texts from Exodus and Deuteronomy, worn during prayer. They do not interpret the pentateuchal command to "bind" such words "for a sign upon your hand" and put them "in your heart" and "between your eyes" (Deuteronomy 11:18) as grounds for any specific ritual. Rather, they take the command to mean that one should always keep God's laws in one's heart and observe them.

This apparent anticultist attitude should not be construed as indicating that Samaritan religion is antiritualistic or even nonritualistic. The opposite is the case: the Samaritans' insistence on strict adherence to the precise wording of the Pentateuch and their rejection of numerous rabbinical interpretations, often of an accommodating nature, make for a ritual rigidity unequaled in rabbinic Judaism.

The point can be illustrated by a short description of a Sabbath spent among the Samaritans. In conformity with the lunar calendar by which the Jewish year is calculated the Samaritan Sabbath begins on Friday evening. To observe the day of rest according to prescribed customs a Samaritan will make it a point to return to his quarter of Nablus or Holon before sundown; only in his own group can the Sabbath be properly sanctified. The observance shows a stringency unparalleled even among the most conservative Jewish communities.

Taking literally the pentateuchal command to keep complete rest and not to kindle any fire, the Samaritans remain in darkness within their houses from Friday night to Saturday night. No food may be cooked, and there are no accepted ways of keeping food warm such as have been instituted in Jewish observance. A Samaritan will leave his house on the Sabbath only to worship in the *kinshah*.

Both the service and the liturgy of the Samaritans are quite unlike those of the Jewish synagogue today and are particularly unlike those of European Jews. The Samaritans appear to have preserved salient features of ancient Jewish ritual as it is described in early rabbinical literature (and also as it is to some degree preserved in non-European Jewish communities). For example, most prayers are chanted by the entire congregation in unison. Others are intoned in the form of a responsory: one member chants a paragraph and then the congregation repeats it. The role of prayer leader is not confined, as it is in European Jewish practice, to a specially appointed and specially trained cantor but is assumed by each member of the congregation in turn. The women of the community do not actively take part in the service, but they are not physically isolated from the male congregation.

Most of the prayers in the Samaritan service are in Aramaic; the passages from the Pentateuch are of course in Hebrew. In the main the prayer texts are compilations created by successive generations of medieval Samaritan liturgical poets. The prayers resemble one another to a large degree; in the main they are variations on a few basic themes. In the Hebrew portion of the liturgy even someone who speaks modern Hebrew would have difficulty following the text; the Samaritan pronunciation of Hebrew differs considerably from that of European and non-European Jews.

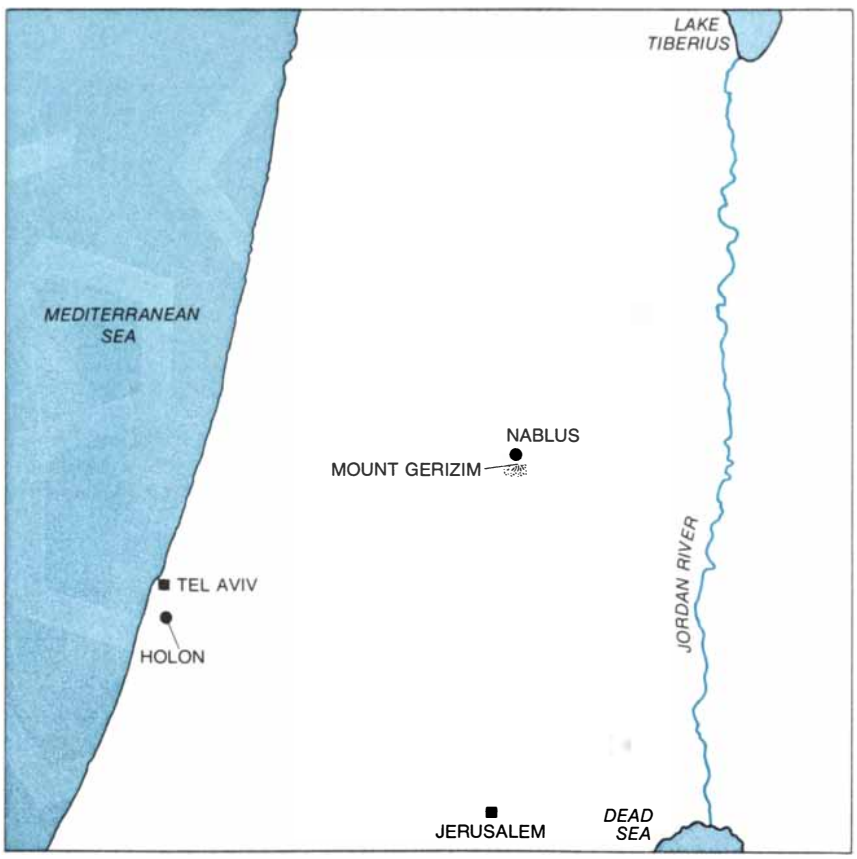
The high point of the Sabbath service is reached when the congregation stands and the weekly text from the Pentateuch is read. In its external appearance the Samaritans' Torah scroll (the five books in one continuous manuscript) resembles the scrolls in contemporary non-European Jewish communities. Made of parchment, it is kept in a cylindrical case and is never removed. The Samaritan Torah is written in the ancient Hebrew alphabet, the prototype of which is known to us from early Hebrew, Moabite and Phoenician inscriptions. It is the same script used to record

© 1976 SCIENTIFIC AMERICAN, INC

the Pentateuch on some of the fragments found in the library at Qumran, known popularly as the Dead Sea scrolls [see "The New Covenanters of Qumran," by Shemaryahu Talmon; SCIENTIFIC AMERICAN, November, 1971]. The Samaritan forms of the letters evidently represent the stage of development attained by the ancient alphabet during the time of the Maccabees, the second century B.C.

The Samaritans attach much importance to the fact that they have retained the ancient script. They offer this fact as proof that their version of the Pentateuch, which differs in many particulars from the Jewish text (in its Masoretic redaction of the seventh to 10th centuries), is a more accurate reflection of the original. Actually most of the differences between the two are of interest only to scholars, being largely confined to variations in spelling or pronunciation. Some of the variants, however, reflect historical and religious issues that divided the Samaritans from ancient Jewry. For example, the Ten Commandments as they are known to Jew and Christian alike are in the Samaritan version compressed into nine. They have been increased to the traditional 10 only by the introduction of an extra commandment (made up of passages drawn from Deuteronomy 11 and 27) declaring the holiness of Mount Gerizim.

In Nablus the Samaritan place of worship in ancient Shechem was left undisturbed by the Arab rulers of the region until the Crusaders arrived at the end of the 11th century. After the Crusaders took the town the old *kinshah* was destroyed, and a church was built on the site. When the Moslems reasserted their authority over Shechem in 1167, the Crusaders' church became a mosque. These vicissitudes of history can still be traced. The mosque of today includes a Norman tower, obviously part of the Crusaders' church. A stone tablet, which in the past had been cemented into the east wall of the tower, is engraved with the opening word of each of the Samaritans' version of the Ten Commandments. The words are written in the same ancient Hebrew script used in the Samaritans' Torah, and the elongated, triangular form of one of the letters, the *ayin,* strikingly resembles the tall and narrow shield the Crusaders carried. The resemblance must have caught the imagination of some medieval Christian mason among the Crusaders: on close scrutiny one discovers that a slender Christian cross has been incised within the *ayin* triangle, transforming the letter into a replica of the Crusaders' shield.



MODERN SAMARITAN COMMUNITIES are in Holon, south of Tel Aviv, and in Nablus, on the flank of Mount Gerizim, the Neapolis of the Roman period. Nablus stands near the ruins of ancient Shechem, once a major Samaritan city. Mount Gerizim remains a focus of ritual.

The *kinshah* in Holon was at first only a room set apart in a private house. The late president of Israel, Itzhak Ben-Zvi, was a patron of the Samaritans, and through his endeavors a special building was erected in the 1960's to serve as a place of worship. This was the first Samaritan *kinshah* built outside the Nablus community in at least 200 years; by the middle of the 18th century all the other Samaritan settlements that once had flourished in the Near East had been abandoned, and their remaining members had resettled at Nablus in order to be close to holy Mount Gerizim.

To this day the mountain remains the Samaritans' holy land. It was there rather than in Jerusalem, they say, that the Temple should have been built, and early in the Hellenistic period there a Samaritan temple probably did stand. Recent excavations at Gerizim have exposed evidence of what seem to be successive sanctuaries topped by the ruins of a temple constructed during the rule of the Roman emperor Hadrian (A.D. 117–138). The underlying structures, in the opinion of the excavator, Robert J. Bull of Drew University, may well represent earlier Samaritan sacred buildings. The site appears to have been abandoned by the Samaritans, possibly before the advent of Christianity and certainly by Hadrian's time. The Samaritans subsequently sanctified another part of Mount Gerizim, but no building stands there now, and it is unlikely that one ever did. In the sacred precinct is a short trench leading toward an earth altar where to this day animal sacrifices are offered at Passover.

As Passover approaches it is customary for the Samaritans from both communities to congregate at Nablus, preparing to make a three-day pilgrimage to the summit of Mount Gerizim that ends on the holy day. In the past the pilgrims camped in tents near the sacred precinct; today permanent quarters have been built there, and the pilgrims occupy them during the Passover festival. It is in the course of this annual spring rite that further salient features of the Samaritan community become apparent. For example, the Samaritans maintain not only that they are the sole true remnant of the 10 tribes of Israel but also that they are the true Israel. In explaining this claim they say that the word Samaritan is derived not from the city name Samaria, the capital of ancient Ephraim, but from the Hebrew word *shamar,* meaning to guard or observe. That is why they call themselves the Shamrayin, implying they alone are the true guardians of divine law.

The Samaritans consider themselves representatives of the house of Joseph; with the exception of the priestly families, who maintain that they are descended from the tribe of Levi, all Sa-

© 1976 SCIENTIFIC AMERICAN, INC

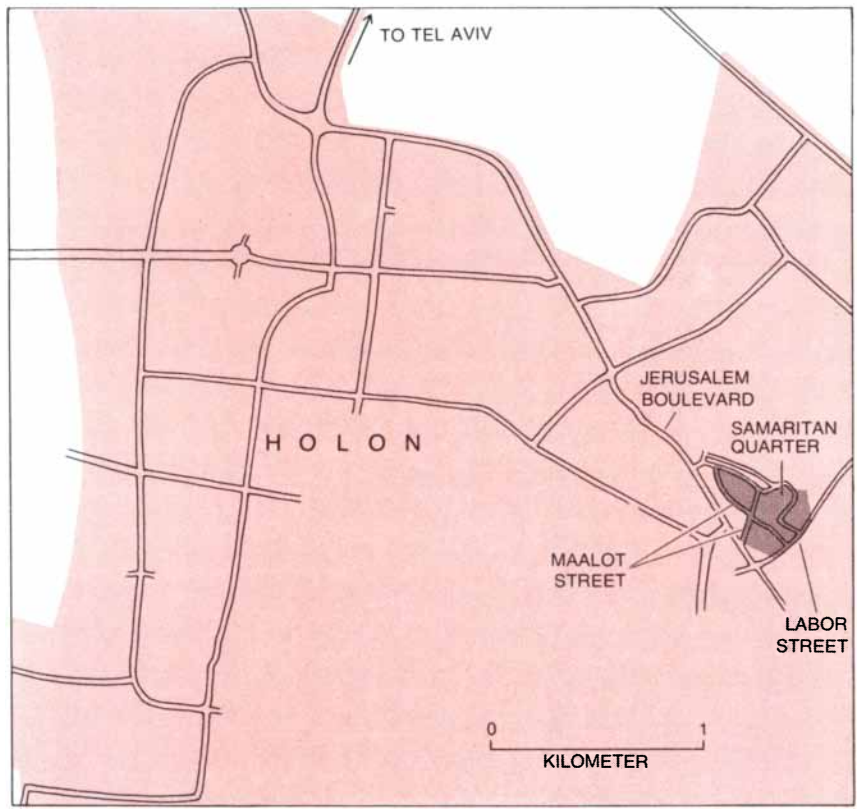
maritan families trace their lineage either to the tribe of Ephraim or to the tribe of Manasseh. Until the beginning of the 17th century there had existed in the Samaritan community an uninterrupted chain of high priests who were by tradition the direct descendants of Eleazar, son of the high priest Aaron, brother of Moses. The line of Aaron came to an end early in the 17th century with a high priest who died without having a son. One of the Samaritan families that traced its descent from the tribe of Levi was then invested with the high priesthood; the family has maintained its primacy to this day.

The priests have fulfilled in the Samaritan tradition the functions of spiritual and communal leadership that in Judaism became the province of the sage and the rabbi. During Passover their role becomes particularly prominent. Within the entire orbit of Judeo-Christian culture only the Samaritans have actively retained the tradition of animal sacrifice. Thus the elaborate ceremony that is still performed on Mount Gerizim on the eve of Passover is the last remaining example of the ritual slaughter characteristic of Temple practices in Jerusalem at the time of Christ.

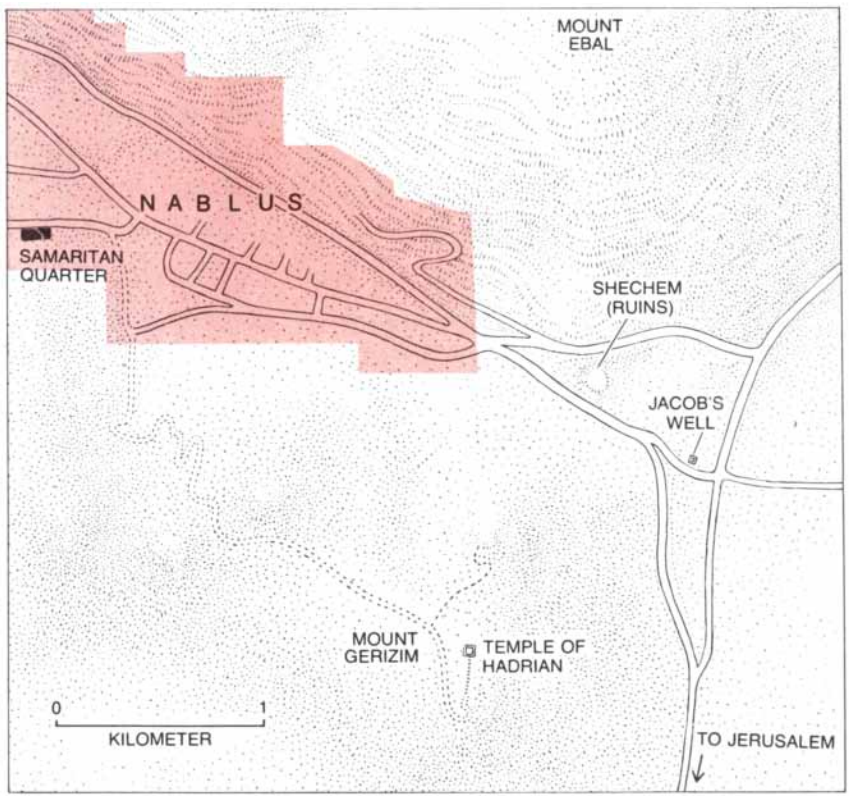
The Samaritan ritual is conducted in strict accordance with the prescriptions recorded in the Pentateuch. At the beginning of the month of Nissan the head of each Samaritan patriarchal unit selects a lamb on behalf of the extended family; thereafter the animal is guarded until the day of the sacrifice. In the early decades of this century seven lambs were sacrificed, each representing one Samaritan patriarchal unit. (Today one of these extended families is almost extinct.) Since then, however, the Samaritan community has grown to its present 500, and the seven former family units have subdivided. As a result the number of Passover lambs sacrificed in 1975 was 19, one for each of the new family subdivisions.

The Passover service begins shortly before sundown; the leading members of the Samaritan community, dressed in flowing white robes, chant a long series of prayers. As the chanting progresses the priests prepare the lambs for sacrifice. The animals are led to the trench near the earth altar in the sacred precinct. Just before sundown, as is prescribed in the Pentateuch, a priest selected for the task cuts the lambs' throats, moving quickly from one animal to the next. Boiling water is immediately poured over the carcasses, and the lambs are stripped of their fleece, gutted and impaled on wood spits. Meanwhile a slow fire has been burning at the bottom of a deep pit; the spitted lambs are now placed in this earth oven, the pit is covered and the meat is left to bake.

As the sacrifice cooks, more prayers

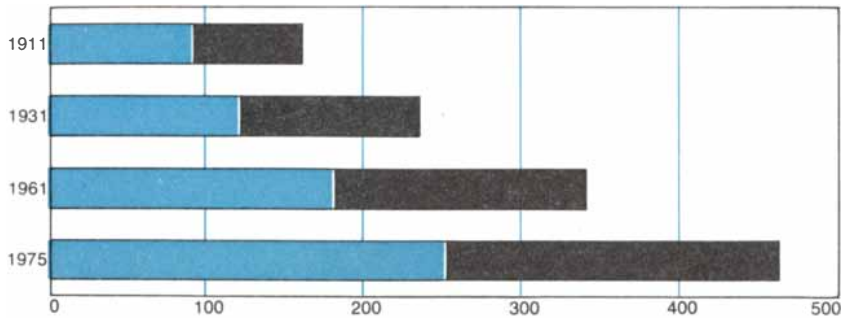


SAMARITAN QUARTER in Holon occupies a few square blocks in the eastern half of the town (*gray*). Known as Shikkun ha-Shomronim, the quarter has a population of some 250 Samaritans. The first Samaritans came from Nablus in the 1920's; others followed after 1948.

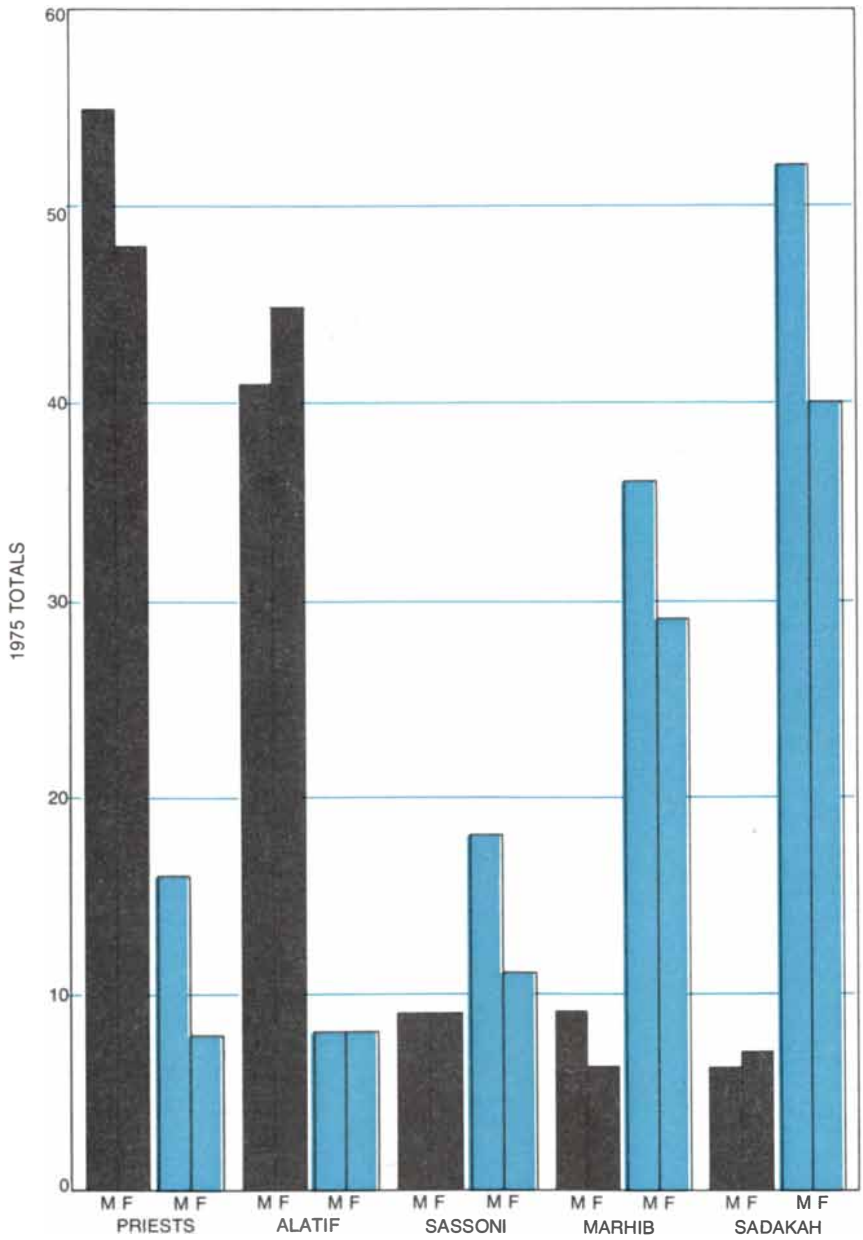


QUARTER IN NABLUS lies at the western end of the town (*black*), near the path leading to the summit of Mount Gerizim. The number of Samaritans in Nablus also approaches 250.

© 1976 SCIENTIFIC AMERICAN, INC



SAMARITAN POPULATION, at a low ebb early in this century, had tripled by 1975; it is now about 500. Males (*color*), however, continue to outnumber females. This is a threat to the Samaritan tradition of exclusively endogamous marriages, often between first cousins.



FIVE MAJOR FAMILIES are not equally represented in the two communities: Nablus (*black*) and Holon (*color*). Few of priestly lineage are found in Holon; until recently there were none, and the great majority still remain in Nablus in order to be near Mount Gerizim. Similarly, the Sadakah and Marhib lineages constitute a majority of the Samaritan immigrants to Holon.

are said. Then all members of the community go indoors and remove their white robes. They emerge clad in rough garments and heavy shoes, with staffs in their hands and bundles on their backs, ready to reenact the Exodus. Exactly at midnight the earth oven is opened. Each family group claims its lamb and everyone in the group tears a piece of baked meat from the carcass, eating quickly in order to symbolize the hasty departure of the children of Israel from Egypt. In compliance with pentateuchal law not a single bone of a sacrificed lamb is broken; all the bones and any leftover meat are thrown on the altar for burning to ensure that no meat remains by morning. After everyone has eaten, the pilgrims circle the holy precinct and its surroundings in a procession that symbolizes the departure from Egypt and the long journey through the wilderness to the Promised Land. After that a full meal is served; this late-night supper more closely resembles the traditional Jewish Seder, the festive repast on the eve of Passover.

During the 20 years when the Jordanian and the Israeli branches of the Samaritan community were separated the Passover ceremony on Mount Gerizim was the only occasion when the community was united. Since Samaritans almost never marry outside their community, the yearly meeting became a time for matchmaking. The Nablus Samaritans often "exported" wives to the Holon Samaritans, who could not easily find spouses within their group, which at that time was smaller. Young people could meet; parents, in consultation with the high priest, could discuss the particulars of a prospective marriage and decide which partner should follow the other, to Israel or to Jordan.

Today Samaritans from Holon and Nablus are free to meet all year long. In the wake of this development a feeling of renaissance has taken hold of many in the community. Never faltering in their belief that they are the true Israel and that the day will come when God will turn his favor toward them, the Samaritans interpret the reunion of the community as a good omen.

The new situation may indeed have a decisive impact on the social structure of the community. Since ritual functions were closely bound up with Mount Gerizim, the Samaritans who settled in Israel in 1948 included no representatives of the priestly lineages. Traditionally the spiritual leaders of the community, the priests had also been responsible for all aspects of community life. They were not only priests but also teachers and the only community scribes. (All the holy books of the Samaritans are handwritten.) The families resettled in Holon suddenly found themselves deprived of these services. The holy books they brought with them were soon too worn

© 1976 SCIENTIFIC AMERICAN, INC

for further use, and yet new copies of the Pentateuch and new prayer books were required for Sabbath services at the *kinshah* and for the instruction of the young.

Wishing to avoid any appearance of encroaching on priestly prerogatives, the laity in Holon refrained from making the much-needed new copies themselves. It soon became evident that the situation was a serious handicap to community education. Finally photocopying and then printing solved the problem. Lay leaders emerged to assume the social role of the absent priests and also to represent the Holon Samaritans in dealings with the outside world. As a result a new kind of communal leader appeared in Holon: one who had acquired his position not by inheritance, as a priest does, but by his and his family's prominence outside the special fields of liturgy and teaching.

The Holon Samaritans, at first fewer in number than the Nablus Samaritans, have at the same time been heavily influenced by the westernized Israeli population that surrounds them. One result was that the Holon group soon outdistanced the Nablus majority intellectually, economically and organizationally. The Holon Samaritans, who had graduated from Israeli high schools and some of whom had gone on to become teachers or Israeli government officials, were more influential than the less educated priestly leaders in Nablus. In fact, they became rivals of the hereditary priesthood and contenders with them for social supremacy. If this development had continued unchecked, it might have caused a permanent social rift between the two Samaritan groups. Their reunion appears to have alleviated the tension and at least for the time being has forestalled serious internal conflict.

So far the Samaritans have succeeded in maintaining their identity by persistently adhering to an endogamous marriage system that is practiced not only within the limits of the community but also often within the limits of an extended family lineage. Batsheva Bonné of the Tel Aviv University School of Medicine has done considerable demographic research among the Samaritans. Her studies show a strong community preference for first-cousin marriages, a tendency that seems also to persist among the Samaritans at Holon. As a result the inbreeding coefficient calculated for the generation of Samaritans born of marriages contracted after 1933 is the highest on record in Israel. This is in sharp contrast to the findings of surveys conducted among several immigrant ethnic groups in Israel, and among non-European Jewish immigrants in particular. Those findings indicate a trend toward declining endogamous marriage.

Bonné's most recent research indi-

IMSAI Presents the Intelligent Breadboard System

IMSAI, one of the world's largest manufacturers of microcomputer systems – the high-performance IMSAI 8080, has achieved another breakthrough – The IMSAI Intelligent Breadboard System. What you can do with the IMSAI Intelligent Breadboard is enough to qualify you and your IMSAI 8080 computer as a microprocessor R&D department. Our Intelligent Breadboard is a complete system for development of discrete logic circuits and microprocessor software circuits. It connects directly to the IMSAI 8080, communicating with the 8080's bus and 48 TTL level I/O lines.

Use the Intelligent Breadboard to implement circuits in hardware and transfer them to software in a step-by-step fashion. Because the computer and Intelligent Breadboard communicate, hardware-software tradeoffs may be studied and new circuit designs tested extensively. It is a flexible, easy-to-use instructional tool ready to serve your development ideas under conditions of total control.

Circuits can be built using the bus logic of the 8080 while remaining outside the computer chassis. Think of that for ease of construction and analysis! There are dozens of uses for this new and unique laboratory tool. Examples:

- As a learning tool to teach replacement of discrete devices in digital logic circuits with microprocessors and computer software.
- A sophisticated classroom laboratory for logic design courses.
- To simulate hardware circuits in software.
- To develop new I/O interfaces and memory systems that connect directly to, yet are physically outside, the computer. And much, much more.

Best news of all, pricewise the IMSAI Intelligent Breadboard, like the IMSAI 8080, is as affordable as it is useful: $435 kit; $625 assembled. Immediate delivery. Call or write for a complete illustrated brochure describing the IMSAI 8080, options, peripherals, software, prices and specifications, plus everything you want to know about the IMSAI Intelligent Breadboard.

Call for the name of your nearest IMSAI dealer. Dealer inquiries invited. Why not the very best?

IMSAI

Manufacturing Corporation

14860 Wicks Blvd., San Leandro, CA. 94577, (415) 483-2093 TWX 910-366-7287

© 1976 SCIENTIFIC AMERICAN, INC

cates that since 1963 the rate of Samaritan inbreeding has increased still further and has been accompanied by a rise in the incidence of genetic defects. For example, preliminary statistics reveal that some 27 percent of Samaritans suffer from color blindness; the large majority of the color-blind are members of the extended Sadakah family living in Holon. The extended Marhib family shows a high level of respiratory deficiency, indicated by chronic shortness of breath. The incidence of deaf-mutism is high in the Alatif family and the priestly family. In view of the persistent preference for first-cousin marriages Bonné's survey suggests an urgent need for young Samaritans, particularly young priests, to seek genetic counseling before choosing a spouse.

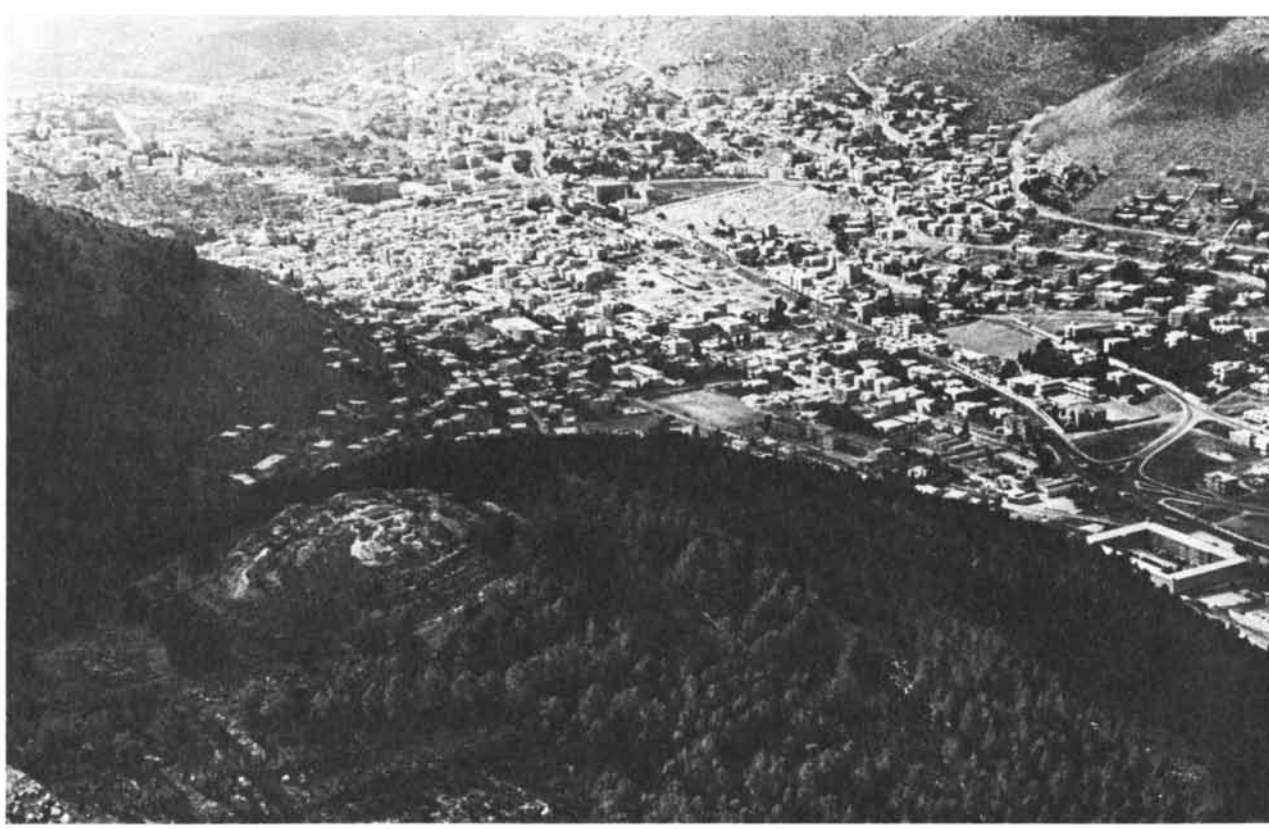
At the same time isolated instances of intermarriage between Samaritans and Jews are known. In the early 1930's the number of Samaritan women in Holon had declined to the extent that many young Samaritan men could not find wives. At that time at least three Jewish women married into the community; reportedly two of them were non-European immigrants and one was a European.

Today the fact of uninterrupted contact with Israeli society, together with Samaritan recognition of the grave threat presented to the integrity of their community by the adverse genetic effects of inbreeding, have produced a change of attitude toward endogamy. Young Samaritans born, bred and educated in Israel are now more willing to marry Jewish girls in spite of some continuing community opposition to such matches. In all the intermarriages that have taken place in recent years a Jewish woman has joined the Samaritan community. To the best of my knowledge there has been no instance of a Jewish man joining the Samaritan community either by marriage or by conversion.

Even more than in Judaism the Samaritan religion is inseparably linked to ethnic extraction. Being born a Samaritan by definition means adhering to the Samaritan religion, and only a Samaritan by birth can be a true Samaritan by belief. There is no way out of the community: once a Samaritan, always a Samaritan. By the same token there is no way in; there is no recognized means of absorbing a male outsider.

The history of the Samaritans is a long and tragic one. During the last centuries before the Christian Era the Samaritans were a vigorous nation, with a population that was probably more than a million. At that time they held their own against their Jewish cousins and even tried to resist the Roman conquerors of Palestine. Since then, however, their lot has been endless persecution. They were killed by the thousands by Jews, Christians and Moslems in succession, and those who survived lived an existence of political, economic and cultural suppression. These adverse external conditions reinforced the stultifying impact of the Samaritans' self-willed religious conservatism. For a millennium and a half they produced few new ideas or new forms in the realms of thought and literature and contented themselves with reproducing a series of variations on themes established in antiquity. The conservatism of their culture weakened the Samaritans' power to resist their oppressors; they simply lacked the means of adjusting to novel situations. In a way the Samaritans illustrate what might have happened to Judaism if the early leaders of the Pharisees—the rabbis—had not evolved ways of interpreting biblical law in the light of changing social conditions.

The survival of the Samaritans nonetheless embodies a reassuring lesson. The fact that a Samaritan minority still exists would seem to indicate that there is room in a society for diverse interpretations of common religious heritages. This tiny remnant of what was once a sovereign nation is a clear illustration of how the interaction of ethnic and religious factors can safeguard the integrity of a social group that would otherwise have vanished almost without trace.



MOUNT GERIZIM lies in the foreground of this aerial photograph; Nablus is in the background. At the top of the mountain at the left are temple ruins being excavated by Robert J. Bull of Drew University. The latest of the buildings is a temple erected during the reign of the Roman emperor Hadrian (A.D. 117–138). Under it is a fieldstone wall that may be a part of the Samaritan temple on the mountain.

© 1976 SCIENTIFIC AMERICAN, INC

APF

New programmable you should consider...

More preprogrammed functions than any other hand-held programmable scientific calculator on the market.

PROGRAMMING CAPABILITY	APF	SR-56	HP-25	51-II
Program steps	72	100	49	None
Conditional branching instructions	Yes	Yes	Yes	No
Unconditional branching	Yes	Yes	Yes	No
Looping	Yes	Yes	Yes	No
Editing: Step	Yes	Yes	Yes	No
Backstep	Yes	Yes	Yes	No
Insert, delete	Yes	Yes	Yes	No
NOP	Yes	Yes	Yes	No
Single step execution	Yes	Yes	Yes	No
OPERATING CHARACTERISTICS				
Logic system	ALG	AOS	RPN	AOS
Max. number of pending operations	4	7	3	5
Parentheses levels	4	9	0	9
Memories	10	10	8	3
Store & recall	Yes	Yes	Yes	Yes
Clear memory	Yes	Yes	Yes	Yes
Sum to memory	Yes	Yes	Yes	Yes
Subtract from memory	Yes	Yes	Yes	Yes
Multiply into memory	Yes	Yes	Yes	Yes
Divide into memory	Yes	Yes	Yes	Yes
Exchange display with memory	Yes	Yes	No	Yes
Perform same calculation on all memories simultaneously	Yes	No	No	No
Exchange x with y	Yes	No	Yes	Yes
Fixed decimal option	Yes	Yes	Yes	Yes
Calculating digits	10	12	10	12
Angular mode Deg/Rad/Grads	Yes	Yes	Yes	Yes
Digits displayed (mantissa & exponent)	8+2	10+2	8+2	8+2
Eng. Notation	Yes	No	Yes	Yes
Sci. Notation	Yes	Yes	Yes	Yes
Change Sign	Yes	Yes	Yes	Yes
Automatic Constant	Yes	No	No	Yes
CALCULATING CHARACTERISTICS				
Log. ln x, 10^x, e^x, y^x, $\sqrt{x}$, $\bar{x}$, $\hat{\sqrt{y}}$ π, x^2	Yes	Yes	Yes	Yes
X!	Yes	No	No	Yes
Int X (integer part)	Yes	Yes	Yes	No
Fractioned part	Yes	Yes	Yes	No
Trig functions & inverses	Yes	Yes	Yes	Yes
Hyperbolic function & inverses	Yes	No	No	Yes
Deg/MIN/sec to decimal deg & inverses	Yes	No	Yes	Yes
Polar to rectangular conversion & inv.	Yes	Yes	Yes	Yes
Mean, variance & standard deviation	Yes	Yes	Yes	Yes
Linear regression	Yes	No	No	Yes
Trend line analysis	Yes	No	No	Yes
Slope and intercept	Yes	No	No	Yes
Spherical to cartesian conversion & inv.	Yes	No	No	No
%	Yes	No	Yes	Yes
Δ %	Yes	No	No	Yes
Permutations & Combinations	Yes	No	No	No
English-Metric Conversions & inverse F-C, In-Cm, Gal-Ltr, Lb-Kg	Yes	No	No	Yes

TK Enterprises, world-leading calculator specialist, is proud to present a newly advanced programmable scientific calculator.

The **Mark 90** is designed primarily for scientists and engineers, yet its price puts it well within the reach of students and schools. Its power makes it invaluable for any commercial or industrial manager.

The Mark 90 calculator uses the algebraic method of entering problems. Its programming feature, 10 addressable memories and 4 levels of parenthesis will allow you to work the most complicated mathematical equations with full confidence of accuracy.

It operates on either self-contained rechargeable battery pack, or from an A.C. adaptor. A.C. Adaptor/recharger and Carrying Case are included. Dimensions: L 6" x H 3" x W 1⅛."

TRY IT FOR 10 DAYS. The APF Mark 90 carries a one year warranty on repair or replacement. This remarkable programmable calculator is only $69.95 plus $3.50 shipping & insurance. SEND COUPON BELOW TODAY OR CALL TOLL FREE (800) 421-0367...before news of APF Mark 90's demand exceeds supply!

T. K. ENTERPRISES: DEPT. SA-A, 16611 Hawthorne Blvd., Lawndale, CA. 90260—(213) 370-5795.
CALL TOLL FREE (800) 421-0367

tk enterprises
BETTER BUSINESS PRODUCTS

DEPT. SA-A
16611 Hawthorne Blvd.
Lawndale, CA. 90260
(213) 370-5795

Yes...I'd like to try the APF Mark 90 programmable calculator for 10 days.
Please ship me the APF Mark 90 complete with accessories for $69.96 plus $3.50 for shipping and insurance. I understand that if I'm not completely satisfied, I may return the unit and accessories within 10 days for full refund (or, at my option, receive full credit toward the purchase of any other calculator carried by T. K Enterprises)

NAME _____
ADDRESS _____
CITY, STATE, ZIP _____
PHONE _____

☐ **Cash Payment:** Check or money order enclosed (Please be sure to add any applicable state and local taxes.) To speed up delivery send money order

☐ **Bill My Credit Card:**
☐ Master Charge ☐ BankAmericard
Bank No _____
Account No _____
Expiration Date _____
Signature _____

© 1976 SCIENTIFIC AMERICAN, INC

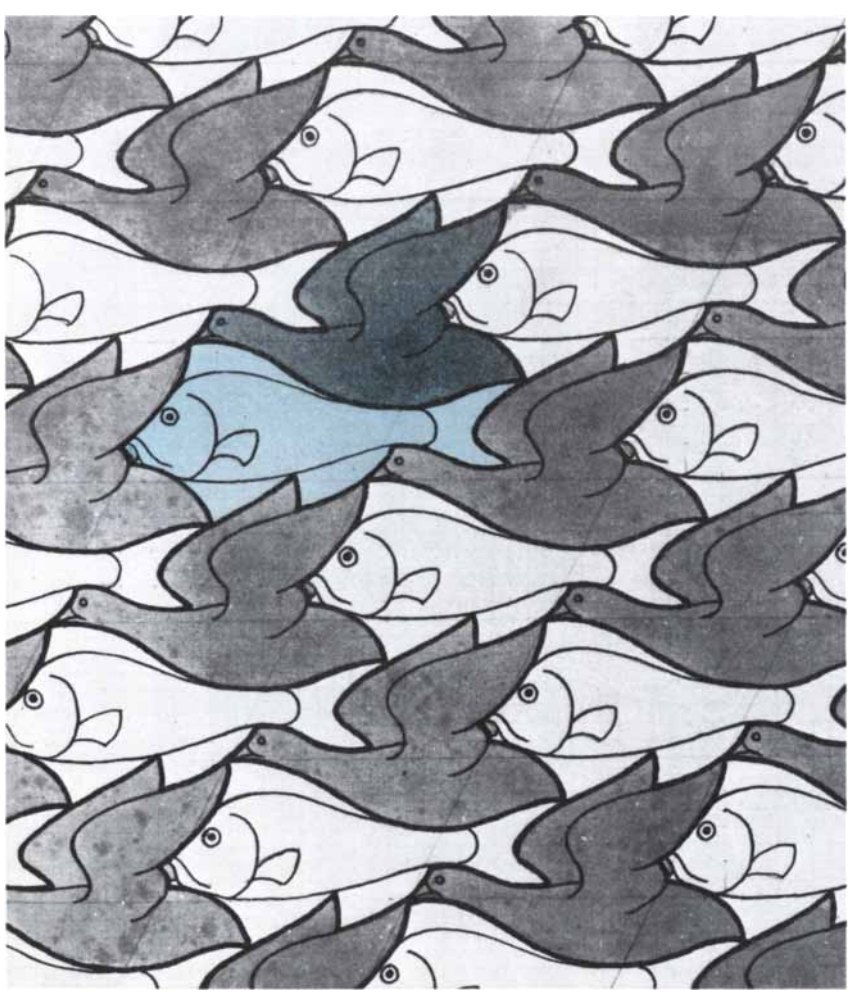
MATHEMATICAL GAMES

Extraordinary nonperiodic tiling that enriches the theory of tiles

by Martin Gardner

In August, 1975, at the end of a two-part article on tiling the plane with congruent convex polygons, I promised a later article on nonperiodic tiling. This column fulfills that promise and presents for the first time a remarkable nonperiodic tiling discovered by Roger Penrose, a British mathematical physicist. First let me give some definitions and background.

A periodic tiling is one on which you can outline a region that tiles the plane by translation, that is, by shifting the position of the region without rotating or reflecting it. M. C. Escher, the Dutch artist, is famous for his many pictures of periodic tilings with shapes that resemble living things. The illustration below is typical. The colored area outlines a fundamental region that tiles by translation. Think of the plane as being covered with transparent paper on which each tile is outlined. Only if the tiling is periodic can you shift the paper, without rotation, to a new position where all outlines again exactly fit.



A periodic tessellation by M. C. Escher (1938)

An infinity of shapes—for instance the regular hexagon—tile only periodically. An infinity of other shapes tile both periodically and nonperiodically. A checkerboard is easily converted to a nonperiodic tiling by identical isosceles right triangles or by quadrilaterals. Simply bisect each square as is shown at the left in the top illustration on the opposite page, altering the orientations to prevent periodicity.

Isosceles triangles also tile in the radial fashion shown in the center of the illustration. Although the tiling is highly ordered, it is obviously not periodic. As Michael Goldberg pointed out in a 1955 paper titled "Central Tessellations," such a tiling can be sliced in half, and then the half planes can be shifted one step or more to make a spiral form of nonperiodic tiling, as is shown at the right in the illustration. The triangle can be distorted in an infinity of ways by replacing its two equal sides with congruent lines as is shown in the middle illustration on the opposite page. If the new sides have straight edges, the result is a polygon of 5, 7, 9, 11... edges that tiles spirally. The bottom illustration on the opposite page shows a striking pattern obtained in this way from a nine-sided polygon. It was first found by Heinz Voderberg in a complicated procedure. Goldberg's method of obtaining it makes it almost trivial.

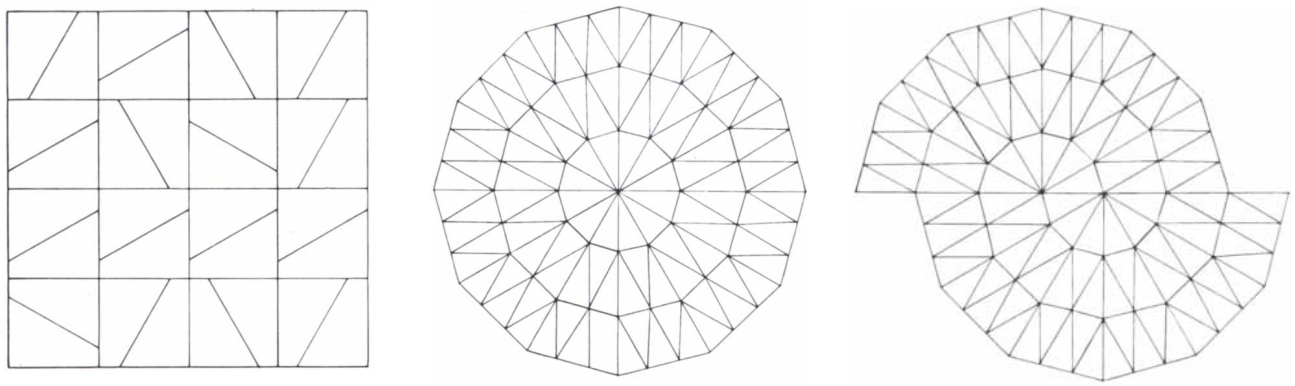
In all known cases of nonperiodic tiling by congruent figures the figure also tiles perodically. The right part of the middle illustration on the opposite page shows how two of the Voderberg enneagons go together to make an octagon that tiles periodically in an obvious way.

Another kind of nonperiodic tiling is obtained by tiles that group together to form larger replicas of themselves. Solomon W. Golomb calls them "reptiles." (See Chapter 19 of my book *Unexpected Hanging*.) The bottom illustration on page 112 shows how a shape called the "sphinx" tiles nonperiodically by giving rise to ever larger sphinxes. Again, two sphinxes (with one sphinx rotated 180 degrees) tile periodically in an obvious way.

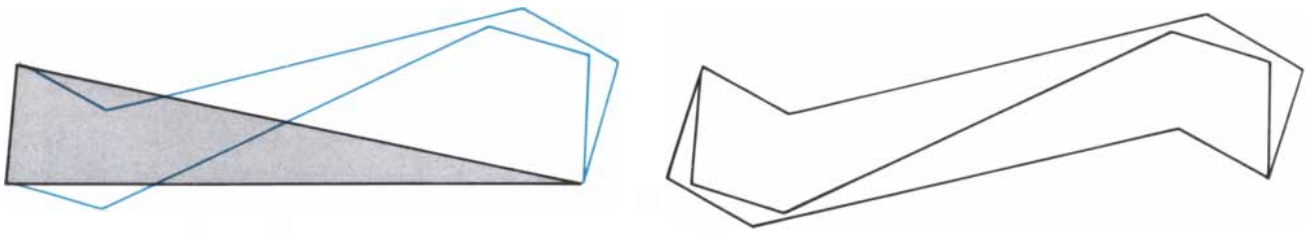
Are there sets of tiles, having two or more *different* shapes, that tile only nonperiodically? By "only" we mean that neither a single shape or subset nor the entire set tiles periodically but that by using all of them a nonperiodic tiling is possible. Rotating and reflecting tiles are allowed.

For many decades experts believed no such set exists, but the supposition proved to be untrue. In 1961 Hao Wang became interested in tiling the plane with sets of unit squares whose edges were colored in various ways. They are called Wang dominoes, and Wang wrote

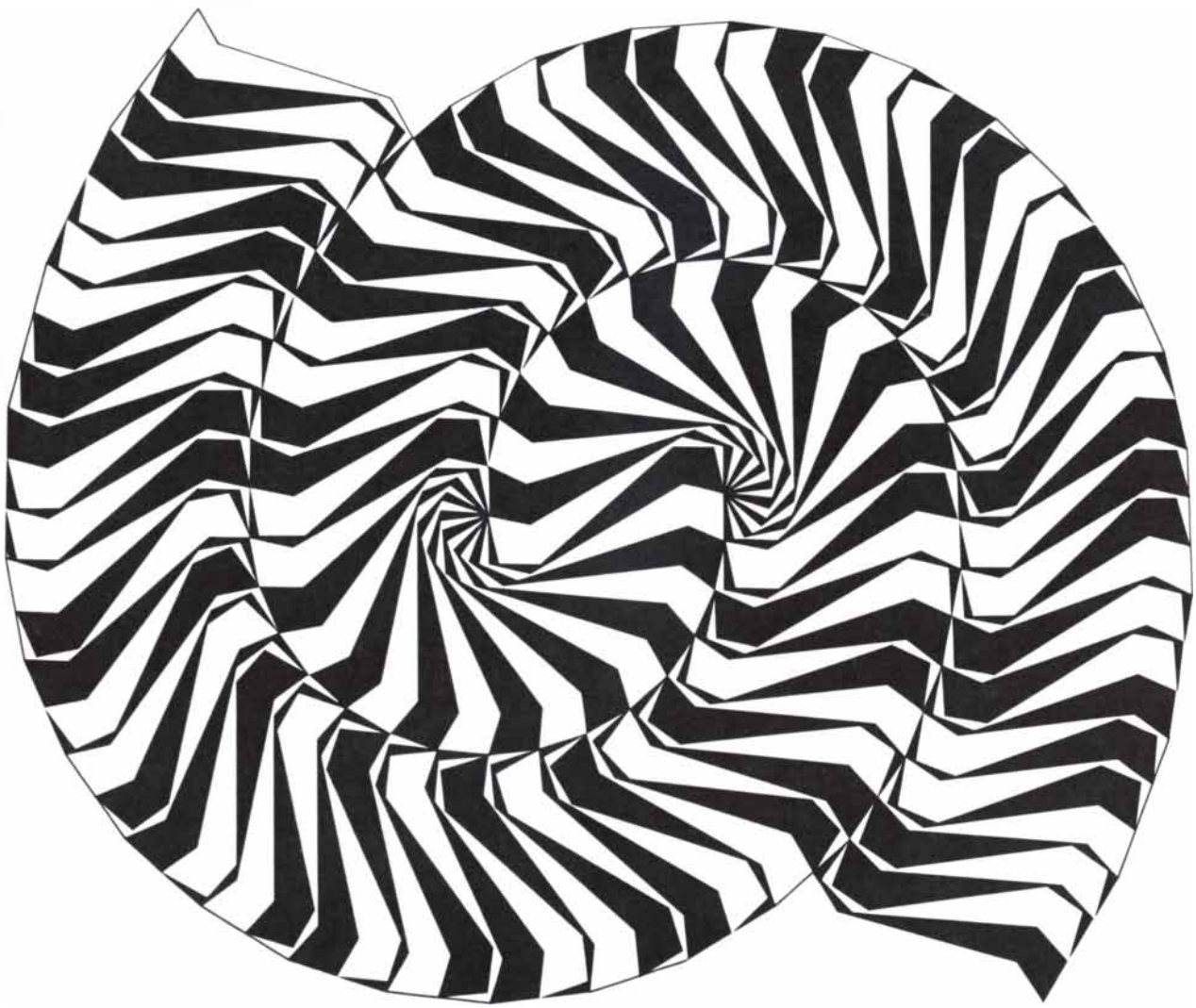
© 1976 SCIENTIFIC AMERICAN, INC



Nonperiodic tiling with congruent shapes

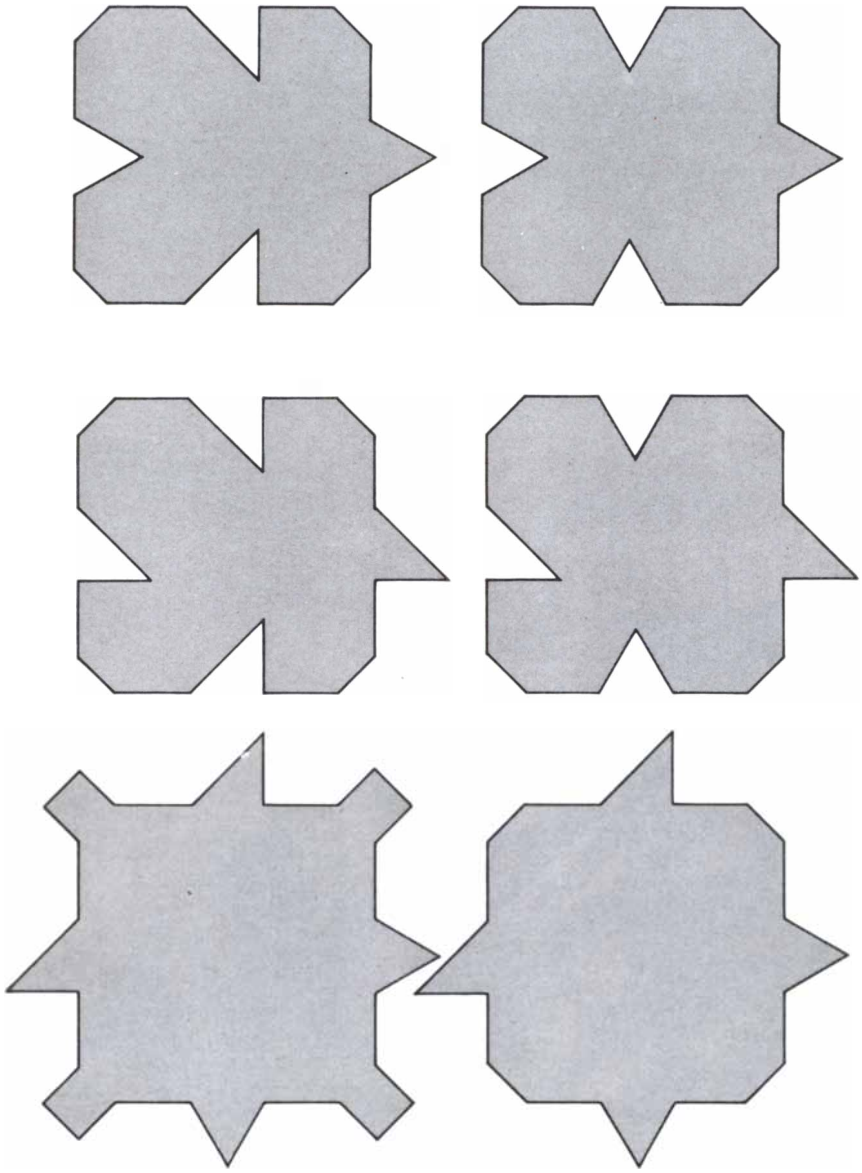


An enneagon **(color at left)** *and a pair of enneagons* **(right)** *forming an octagon that tiles periodically*

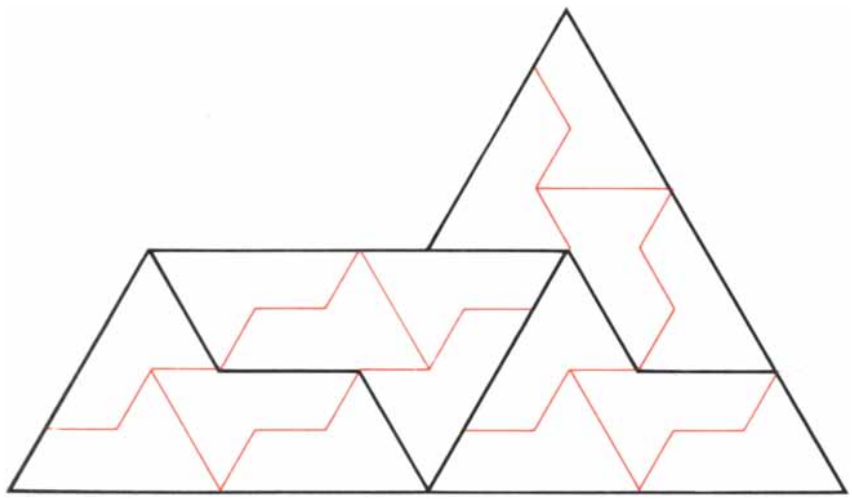


A spiral tiling by Heinz Voderberg

© 1976 SCIENTIFIC AMERICAN, INC



Raphael M. Robinson's six tiles that force a nonperiodic tiling



Three generations of sphinxes in a nonperiodic tiling

a splendid article about them for this magazine [see "Games, Logic and Computers," by Hao Wang; SCIENTIFIC AMERICAN, November, 1965]. Wang's problem was to find a procedure for deciding whether any given set of dominoes will tile by placing them so that abutting edges are the same color. Rotations and reflections are not allowed. The problem is important because it relates to decision questions in symbolic logic. Wang showed that if and only if there is a decision procedure, then any set of dominoes that tiles the plane nonperiodically will also tile periodically. He conjectured that such a procedure exists.

In 1964 Robert Berger, in his thesis for a doctorate from Harvard University in applied mathematics, showed that Wang's conjecture is false. There is no general procedure. Therefore there is a set of Wang dominoes that tiles only nonperiodically. Berger constructed such a set, using more than 20,000 dominoes. Later he found a much smaller set of 104. Last year Raphael M. Robinson reduced the set to 24.

It is easy to change such a set of Wang dominoes into polygonal tiles that tile only nonperiodically. You simply put projections and slots on the edges to make jigsaw pieces that fit in the manner formerly prescribed by colors. An edge formerly one color fits only another formerly the same color, and a similar relation obtains for the other colors. By allowing such tiles to rotate and reflect Robinson constructed six tiles [*see top illustration at left*] that "force nonperiodicity" in the sense explained above.

At the University of Oxford, where he is Rouse Ball Professor of Mathematics, Penrose searched for still smaller sets. Although most of his work is in relativity theory and quantum mechanics, he continues the active interest in recreational mathematics he shared with his geneticist father, the late L. S. Penrose. (They are the inventors of the famous "Penrose staircase" that goes round and round without getting higher; Escher depicted it in his lithograph "Ascending and Descending.") In 1973 Penrose found a set of six tiles that force nonperiodicity. Soon he found a way to reduce them to four, and in 1974 he lowered them to two.

Because the tiles lend themselves to commercial puzzles, Penrose was reluctant to disclose them until he had applied for patents in the United Kingdom, the U.S. and Japan. Now that these patents are pending, I have his permission to write about the tiles. I am equally indebted to John Horton Conway for many of the results of his study of the Penrose tiles.

The shapes of a pair of Penrose tiles can vary, but the most interesting pair have shapes that Conway calls "darts"

© 1976 SCIENTIFIC AMERICAN, INC

Chevy Chevette

43 MPG / 31 MPG

EPA HIGHWAY EPA CITY ESTIMATE

Smile: You've just read the highest EPA estimates for any car built in America. Ever.

Chevy Chevette with available 1.6-litre engine and standard manual transmission.

EPA figures are estimates. The mileage you get will vary depending on your type of driving, driving habits, car's condition and available equipment. In California, EPA figures are lower.

And keep on smiling, there's more.

Chevette has more head room than many mid-size cars.

More horsepower this year.

A hatchback that opens up over 26 cu. ft. of carrying space.

One of the tightest turning circles of any car in the world.

A diagnostic connector for quick electronic service checks.

A service manual written for do-it-yourselfers.

A unitized body that's corrosion-protected.

And over 6,000 Chevy dealers everywhere.

If 43 and 31 are your kind of figures, stop by a Chevy dealer soon.

He has some more numbers you'll like too: Chevette's price.

It'll drive you happy.

GM
MARK OF EXCELLENCE

© 1976 SCIENTIFIC AMERICAN, INC

Now, Dr. Linus Pauling presents new findings that support his arguments even more forcefully than before:
Vitamin C has significant protective activity against many viral diseases, including influenza.

LINUS PAULING

1948 PRESIDENTIAL MEDAL FOR MERIT
1954 NOBEL PRIZE IN CHEMISTRY
1962 NOBEL PEACE PRIZE
1975 NATIONAL MEDAL OF SCIENCE

Vitamin C the Common Cold and the Flu

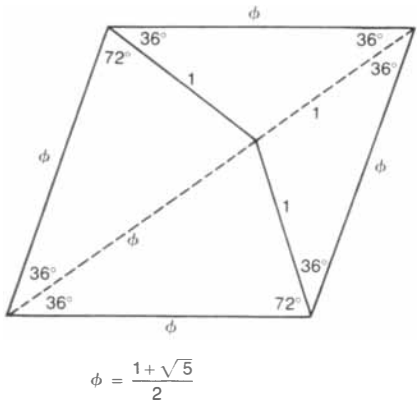
"In 1976-1977, when another swine-flu epidemic is expected, it is especially important that everyone know that he can protect himself to a considerable extent against the disease and its consequences by the use of this important nutrient, vitamin C."

Available now at bookstores everywhere.

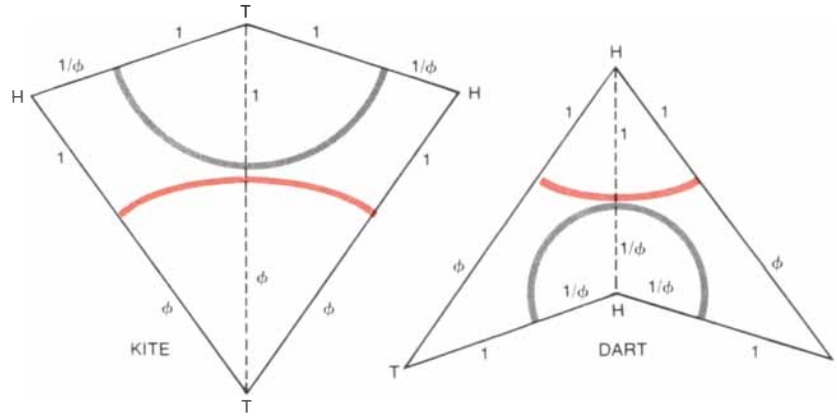
hardcover: $8.95, softcover: $3.45

W. H. FREEMAN AND COMPANY
660 Market Street, San Francisco, CA 94104
58 Kings Road, Reading, England RG1 3AA

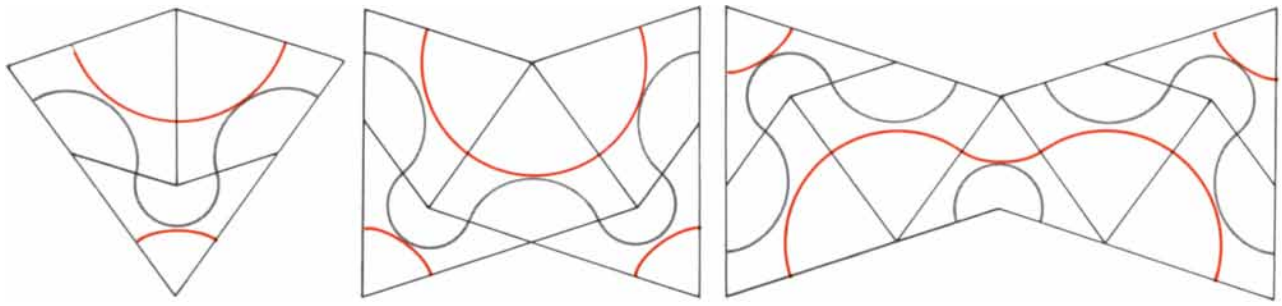
© 1976 SCIENTIFIC AMERICAN, INC



Construction of dart and kite



A coloring of dart and kite to force nonperiodicity



ACE (FOOL'S KITE) SHORT BOW TIE LONG BOW TIE

Aces and bow ties that speed constructions

and "kites." The illustration at the top left on this page shows how they are derived from a rhombus with angles of 72 and 108 degrees. Divide the long diagonal in the familiar golden ratio of $(1 + \sqrt{5})/2 = 1.61803398\ldots$, then join the point to the obtuse corners. That is all. Let phi stand for the golden ratio. Each line segment is either 1 or phi as indicated.

The rhombus of course tiles periodically, but we are not allowed to join the pieces in this manner. Forbidden ways of joining sides of equal length can be enforced by bumps and dents, but there are simpler ways. For example, we can label the corners *H* and *T* (heads and tails) as is shown in the illustration at the top right on this page, and then give the rule that in fitting edges only corners of the same letter may meet. Dots of two colors could be placed in the corners to aid in conforming to this rule, but a prettier method, proposed by Conway, is to draw circular arcs of two colors on each tile, as the illustration shows. Each arc cuts the sides as well as the axis of symmetry in the golden ratio. Our rule is that abutting edges must join arcs of the same color.

To appreciate the full beauty and mystery of Penrose tiling one should make at least 100 kites and 60 darts. The pieces need be colored on one side only. The areas of the two shapes are in the golden ratio. This proportion also applies to the number of pieces you need of each type. You might think you need more of the smaller darts, but it is the other way around. You need 1.618... as many kites as darts. In an infinite tiling this proportion is exact.

A good plan is to draw as many darts and kites as you can on one sheet, with a ratio of about five kites to three darts, using a thin line for the curves. The sheet can be photocopied many times. The curves can then be colored with, say, red and green felt-tip pens. Conway has found that it speeds constructions and keeps patterns stabler if you make many copies of the three larger shapes in the lower illustration on this page. As you expand a pattern you can continually replace darts and kites with aces and bow ties. Actually an infinity of arbitrarily large *pairs* of shapes, made up of darts and kites, will serve for tiling any infinite pattern.

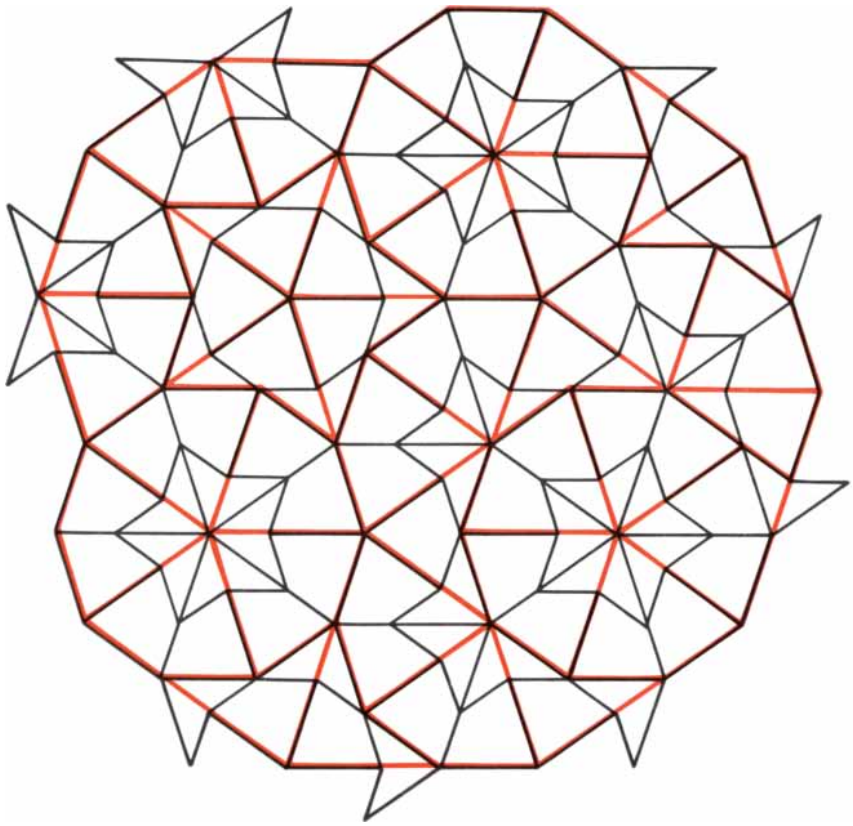
A Penrose pattern is made by starting with darts and kites around one vertex and then expanding radially. Each time you add a piece to an edge you must choose between a dart and a kite. Sometimes the choice is forced, sometimes it is not. Sometimes either piece fits, but later you may encounter a contradiction (a spot where no piece can be legally added) and be forced to go back and make the other choice. It is a good plan to go around a boundary, placing all the forced pieces first. They cannot lead to a contradiction. You can then experiment with unforced pieces. It is always possible to continue forever. The more you play with the pieces, the more you will become aware of "forcing rules" that increase efficiency. For example, a dart forces two kites in its concavity, creating the ubiquitous ace.

There are many ways to prove that the number of Penrose tilings is uncountable, just as the number of points on a line is. These proofs rest on a surprising phenomenon discovered by Penrose. Conway calls it "inflation" and "deflation." The top illustration on the next page shows the beginning of inflation. Imagine that every dart is cut in half and then all short edges of the original pieces are glued together. The result: a new tiling (shown in color) by larger darts and kites.

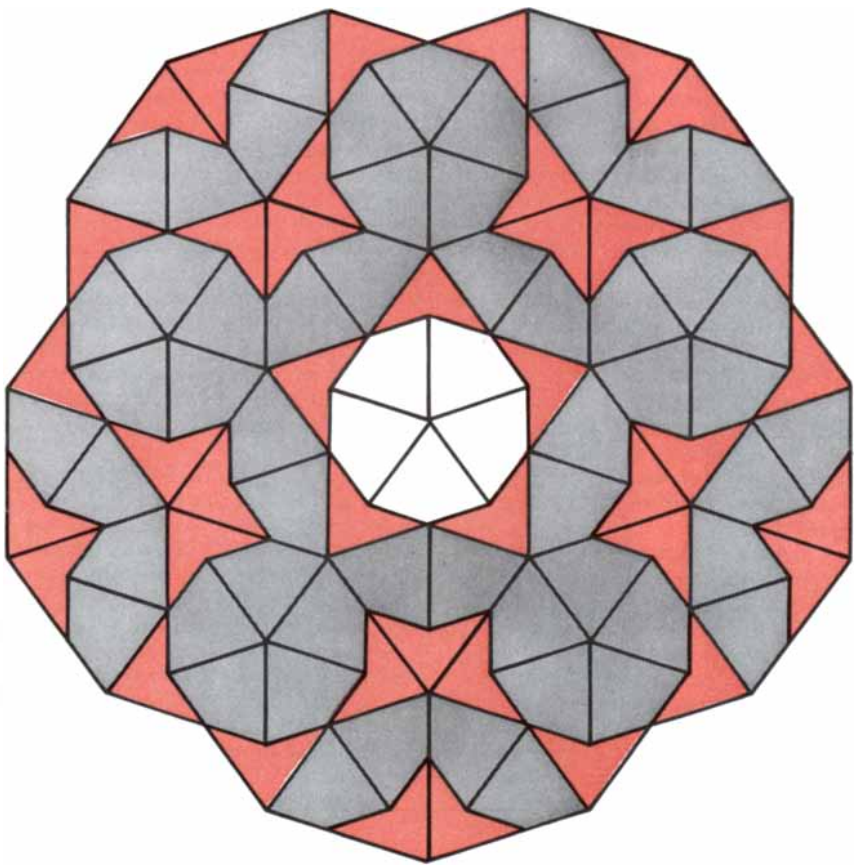
Inflation can be continued to infinity, with each new "generation" of pieces larger than the last. Note that the second-generation kite, although it is the same size and shape as a first-generation ace, is formed differently. For this reason the ace is also called a fool's kite. It should never be mistaken for a second-generation kite. Deflation is the same process carried the other way. On every Penrose tiling we can draw smaller and smaller generations of darts and kites. This pattern too goes to infinity.

Conway's proof of the uncountability of Penrose patterns (Penrose had earlier proved it in a different way) can be outlined as follows. On the kite label one side of the axis of symmetry *L*, the other

© 1976 SCIENTIFIC AMERICAN, INC



How a pattern is inflated



The infinite sun pattern

R (for left and right). Do the same on the dart, using *l* and *r*. Now pick a random point on the tiling. Record the letter that gives its location on the tile. Inflate the pattern one step, note the location of the same point in a second-generation tile and again record the letter. Continuing through higher inflations, you generate an infinite sequence of symbols that is a unique labeling of the original pattern seen, so to speak, from the selected point.

Pick another point on the original pattern. The procedure may give a sequence that starts differently, but it will reach a letter beyond which it agrees to infinity with the former sequence. If there is no such agreement beyond a certain point, the two sequences label distinct patterns. Not all possible sequences of the four symbols can be produced this way, but those that label different patterns can be shown to correspond in number with the number of points on a line.

We have omitted the colored curves on our pictures of tilings because they make it difficult to see the tiles. If you work with colored tiles, however, you will be struck by the beautiful designs created by these curves. Penrose and Conway independently proved that whenever a curve closes, it has a pentagonal symmetry, and the entire region within the curve has a fivefold symmetry. At the most a pattern can have two curves that do not close. In most patterns all curves close.

Although it is possible to construct Penrose patterns with a high degree of symmetry (an infinity of patterns have bilateral symmetry), most patterns, like the universe, are a mystifying mixture of order and unexpected deviations from order. As the patterns expand they seem to be always striving to repeat themselves but never quite managing it. G. K. Chesterton once suggested that an extraterrestrial being, observing how many features of a human body are duplicated on the left and the right, would reasonably deduce that we have a heart on each side. The world, he said, "looks just a little more mathematical and regular than it is; its exactitude is obvious, but its inexactitude is hidden; its wildness lies in wait." Everywhere there is a "silent swerving from accuracy by an inch that is the uncanny element in everything... a sort of secret treason in the universe." The passage is a nice description of Penrose's planar worlds.

There is something even more surprising about Penrose universes. In a curious finite sense, given by the "local isomorphism theorem," all Penrose patterns are alike. Penrose was able to show that every finite region in any pattern is contained somewhere inside every other pattern. Moreover, it appears infinitely many times in every pattern.

To understand how crazy this situation is, imagine that you are living on an

© 1976 SCIENTIFIC AMERICAN, INC

infinite plane tessellated by one tiling of the uncountable infinity of Penrose tilings. You can examine your pattern, piece by piece, in ever expanding areas. No matter how much of it you explore you can never determine which tiling you are on. It is no help to travel far out and examine disconnected regions, because all the regions belong to one large finite region that is exactly duplicated infinitely many times on all patterns. Of course, this is trivially true of any periodic tessellation, but Penrose universes are not periodic. They differ from one another in infinitely many ways, and yet it is only at the unobtainable limit that one can be distinguished from another.

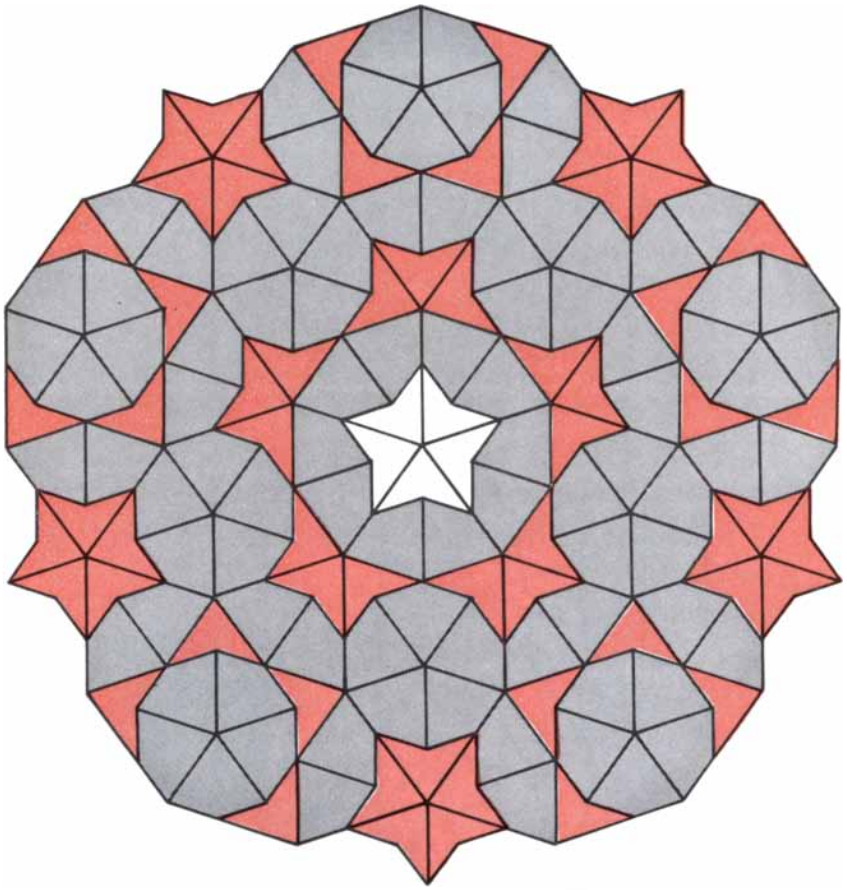
Suppose you have explored a circular region of diameter *d*. Call it the "town" where you live. Suddenly you are transported to a randomly chosen parallel Penrose world. How far are you from a region that exactly matches the streets of your home town? Conway answers with a truly remarkable theorem. The distance is never more than 2*d*! (This is an upper bound, not an average.) If you walk in the right direction, you need not go more than a distance of 2*d* to find yourself inside an exact copy of your home town. The theorem also applies to the universe in which you live. Every large circular pattern (there is an infinity of different ones) can be reached by walking a distance in some direction that is certainly less than twice the diameter of the pattern and more likely about the same distance as the diameter.

The theorem is quite unexpected. Consider an analogous isomorphism exhibited by a sequence of unpatterned digits such as pi. If you pick a finite sequence of 10 digits and then start from a random spot in pi, you are pretty sure to encounter the same sequence if you move far enough along pi, but the distance you must go has no known upper bound, and the expected distance is enormously longer than 10 digits. The longer the finite sequence is, the farther you can expect to walk to find it again. On a Penrose pattern you are always very close to a duplicate of home.

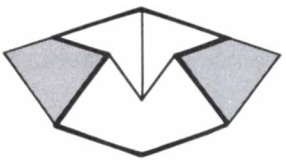
There are just seven ways that darts and kites will fit around a vertex. Let us consider first, using Conway's nomenclature, the two ways with pentagonal symmetry.

The sun (shown in white in the bottom illustration on the opposite page) does not force the placing of any other piece around it. If you add pieces so that pentagonal symmetry is always preserved, however, you will be forced to construct the beautiful pattern shown. It is uniquely determined to infinity.

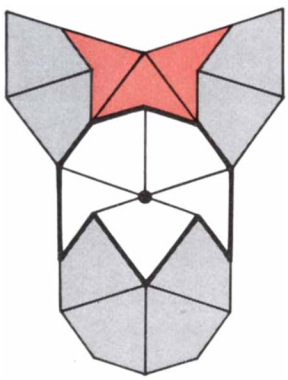
The star, shown in white in the top illustration at the right, forces the 10 gray kites around it. Enlarge this pattern, always preserving the fivefold symmetry, and you will create another flowery design that is infinite and unique. The star and sun patterns are the only



The infinite star pattern

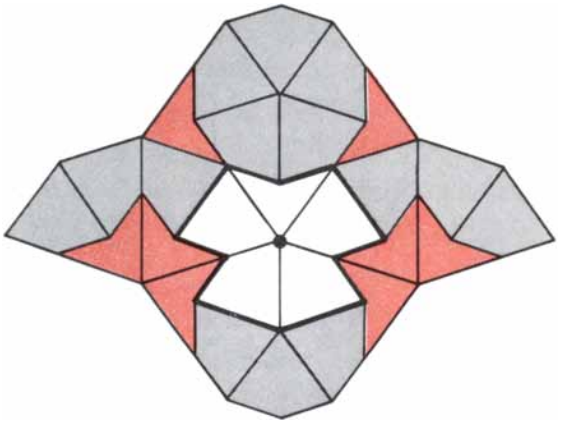


DEUCE



JACK

QUEEN



The "empires" of deuce, jack and queen

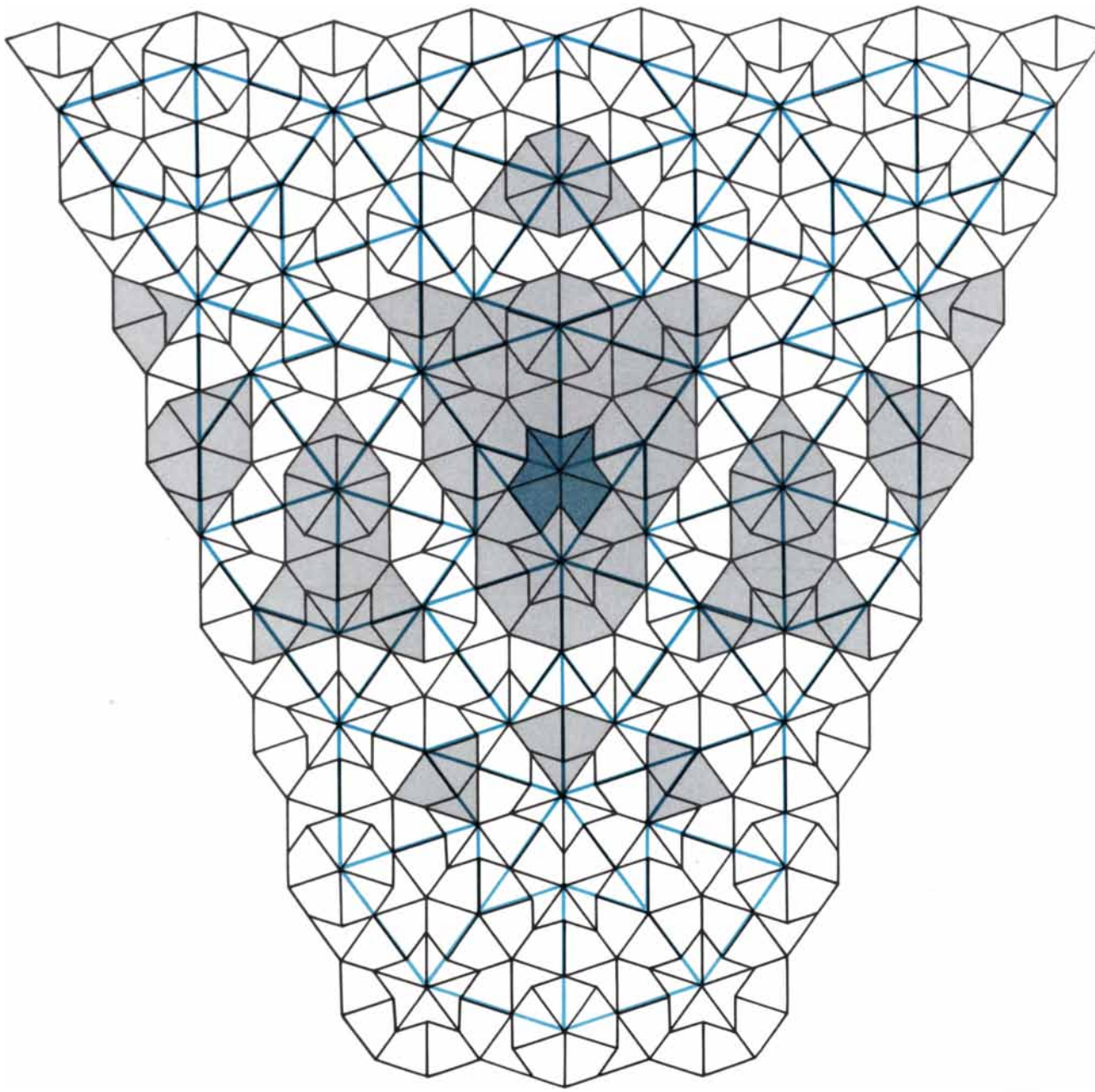
© 1976 SCIENTIFIC AMERICAN, INC

Penrose universes with perfect pentagonal symmetry, and there is a lovely sense in which they are equivalent. Inflate or deflate either of the patterns and you get the other.

The ace is a third way to tile around a vertex. It forces no more pieces. The deuce, the jack and the queen are shown in white in the bottom illustration on the preceding page, surrounded by the tiles they immediately force. As Penrose discovered (it was later found independently by Clive Bach), some of the seven vertex figures force the placing of tiles that are not joined to the immediately forced region. The illustration below shows in gray what is probably the major part of the king's "empire." (The king is the colored area.) All the gray tiles are forced by the king. (Two aces, just outside the left and right borders, are also forced but are not shown.)

This picture of the king's empire was drawn by a computer program written by Eric Regener of Concordia University in Montreal. His program deflates any Penrose pattern any number of steps. The colored lines show the domain immediately forced by the king. The black lines are a third-generation deflation in which the king and almost all of his empire are replicated. It is not known how much farther the empire extends, and the smaller empires of the jack and the queen have not been completely mapped.

The most extraordinary of all Penrose universes, essential for understanding the tiles, is the infinite cartwheel pattern, the center of which is shown in the illustration on the opposite page and on the cover. The regular decagon at the center, outlined in heavy black (each side is a pair of long and short edges), is what Conway calls a "cartwheel." Every point on any pattern is inside a cartwheel exactly like this one. By one-step inflation we see that every point will be



The king's empire

© 1976 SCIENTIFIC AMERICAN, INC

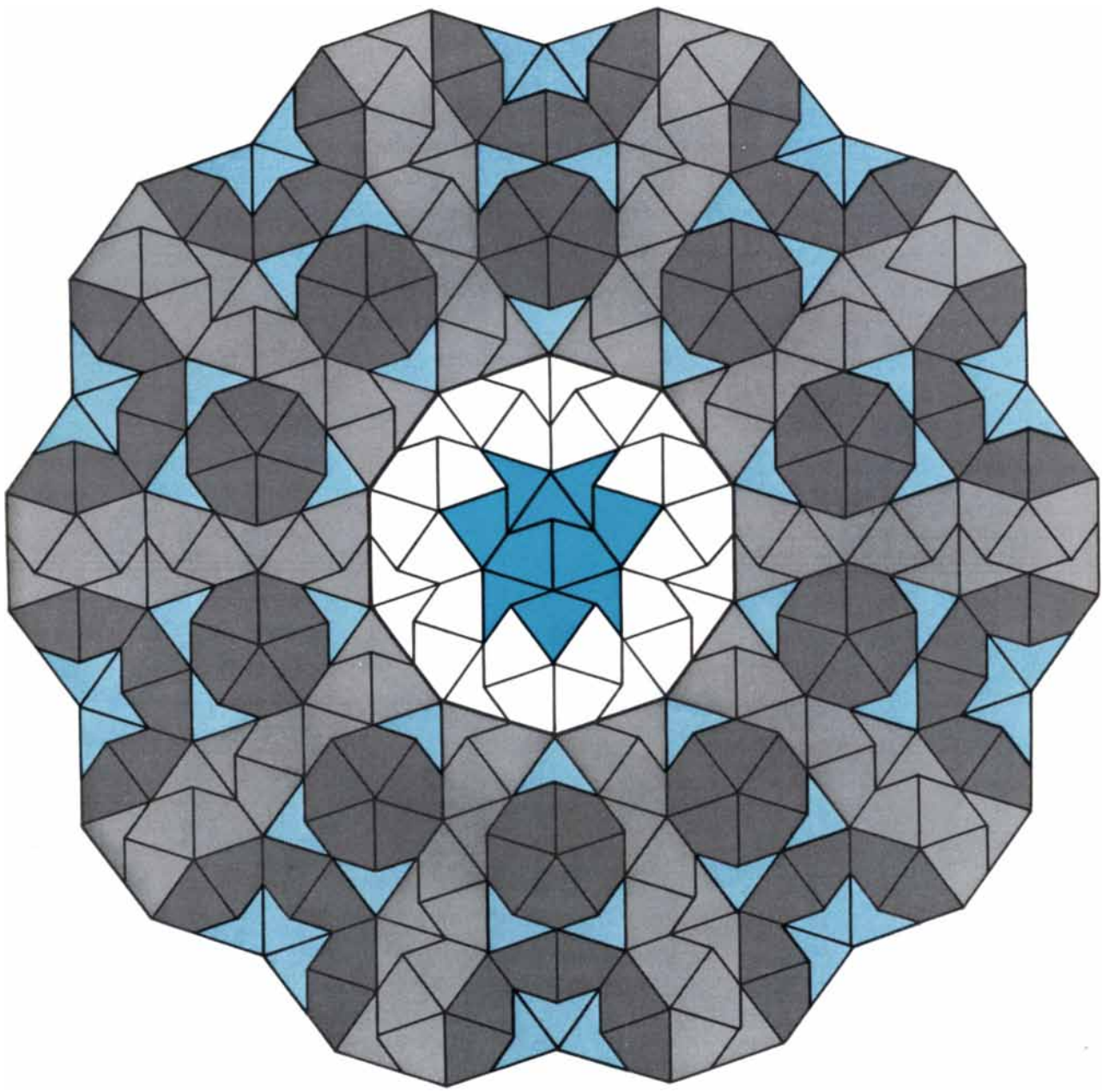
inside a larger cartwheel. Similarly, every point is inside a cartwheel of every generation, although the wheels need not be concentric.

Note the 10 light gray spokes that radiate to infinity. Conway calls them "worms." They are made of long and short bow ties, the long ones being in the golden ratio to the short ones. Every Penrose universe contains an infinite number of arbitrarily long worms. Inflate or deflate a worm and you get another worm along the same axis. Observe that two full worms extend across the central cartwheel in the infinite cartwheel pattern. (Inside it they are not gray.) The remaining spokes are half worms. Aside from spokes and the interior of the central cartwheel, the pattern has perfect tenfold symmetry. Between any two spokes we see an alternating display of increasingly large portions of the sun and star patterns.

Any spoke of the infinite cartwheel pattern can be turned side to side (or, what amounts to the same thing, each of its bow ties can be rotated end for end) and the spoke will still fit all surrounding tiles except for those inside the central cartwheel. There are 10 spokes; thus there are $2^{10} = 1{,}024$ combinations of states. After eliminating rotations and reflections, however, there are only 62 distinct combinations. Each combination leaves inside the cartwheel a region that Conway has named a "decapod."

Decapods are made up of 10 identical isosceles triangles with the shapes of half darts. The decapods with maximum symmetry are the buzzsaw and the starfish shown in the upper illustration on the next page. Like a worm, each triangle can be turned. As before, ignoring rotations and reflections, we get 62 decapods.

When the spokes are arranged the way they are in the infinite cartwheel pattern shown, a decapod called Bat-



The cartwheel pattern

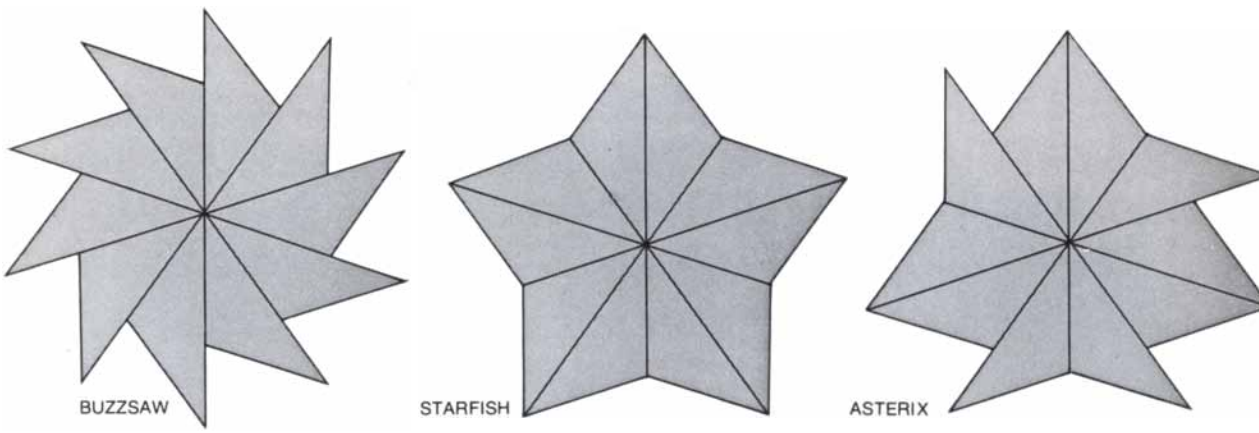
© 1976 SCIENTIFIC AMERICAN, INC

man is formed at the center. Batman (shown in dark color) is the only decapod that can legally be tiled. (No finite region can have more than one legal tiling.) Batman does not, however, force the infinite cartwheel pattern. It merely allows it. Indeed, no finite portion of a legal tiling can force an entire pattern, because the finite portion is contained in *every* tiling.

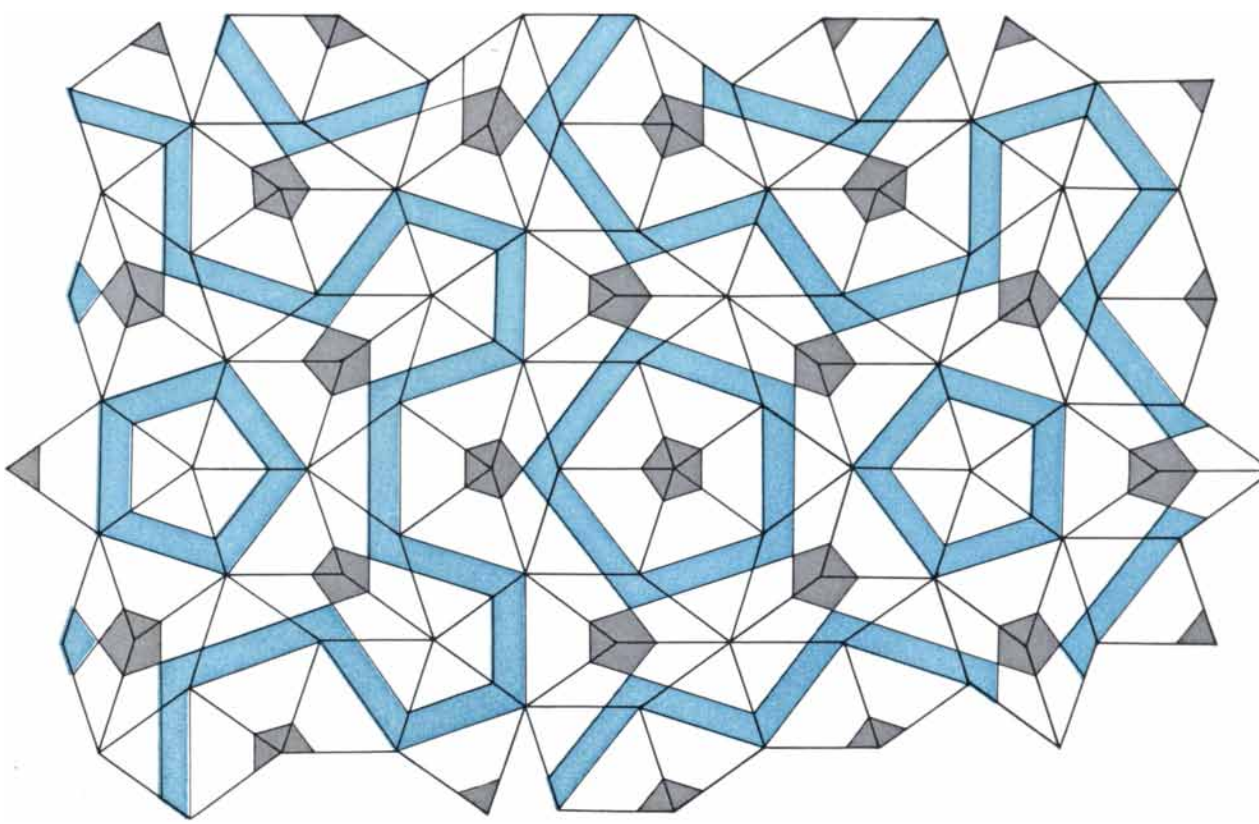
Note that the infinite cartwheel pattern is bilaterally symmetrical, its axis of symmetry going vertically through Batman. Inflate the pattern and it remains unchanged except for mirror reflection in a line perpendicular to the symmetry axis. The five darts in Batman and its two central kites are the only tiles in any Penrose universe that are not inside a region of fivefold symmetry. All other pieces in this pattern or any other one are in infinitely many regions of fivefold symmetry.

The other 61 decapods are produced inside the central cartwheel by the other 61 combinations of worm turns in the spokes. All are "holes" in the following sense. A hole is any finite empty region, surrounded by an infinite tiling, that cannot be legally tiled. You might suppose each decapod is the center of infinitely many tilings, but here Penrose's universes play another joke on us. Surprisingly, 60 decapods force a unique tiling that differs from the one shown only in the composition of the spokes. Only Batman and one other decapod, called Asterix after a French cartoon character, do not. Like Batman, Asterix allows an infinite cartwheel pattern, but it also allows patterns of other kinds.

Now for a startling conjecture. Conway believes, although he has not completed the proof, that every possible hole, of whatever size or shape, is equiv-



Three decapods



A nonperiodic tiling with Roger Penrose's rhombuses

© 1976 SCIENTIFIC AMERICAN, INC

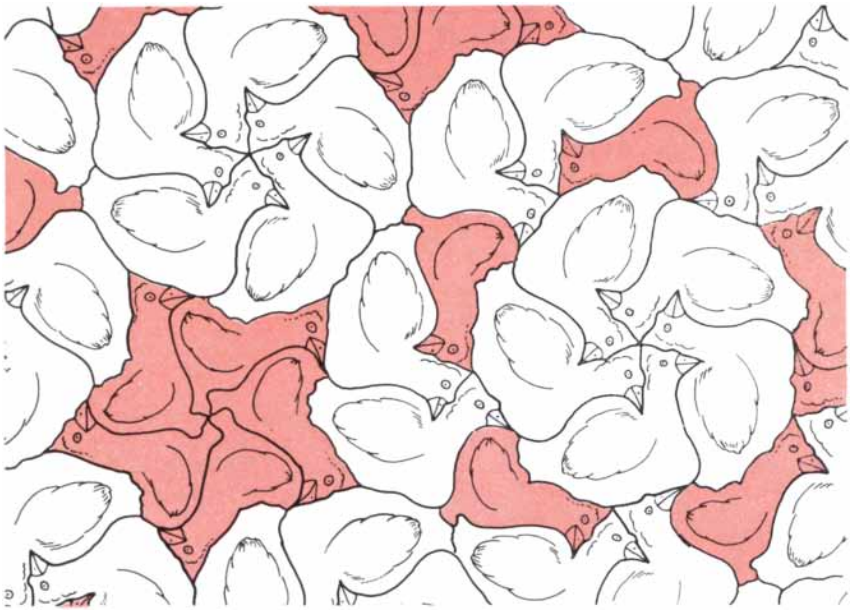
alent to a decapod hole in the following sense. By rearranging tiles around the hole, taking away or adding a finite number of pieces if necessary, you can transform every hole into a decapod. If this is true, any finite number of holes in a pattern can also be reduced to one decapod. We have only to remove enough tiles to join the holes into one big hole, then reduce the big hole until an untileable decapod results.

Think of a decapod as being a solid tile. Except for Batman and Asterix, each of the 62 decapods is like an imperfection that solidifies a crystal. It forces a unique infinite cartwheel pattern, spokes and all, that goes on forever. If Conway's conjecture holds, any "foreign piece" (Penrose's term) that forces a unique tiling, no matter how large the piece is, has an outline that transforms into one of 60 decapod holes.

Kites and darts can be changed to other shapes by the same technique described earlier for changing isosceles triangles into spiral-tiling polygons. It is the same technique that Escher employed for transforming polygonal tiles into animal shapes. The top illustration on this page shows how Penrose changed his darts and kites into chickens that tile only nonperiodically. Note that although the chickens are asymmetrical, it is never necessary to turn any of them over to tile the plane. Alas, Escher died before he could know of Penrose's tiles. How he would have reveled in their possibilities!

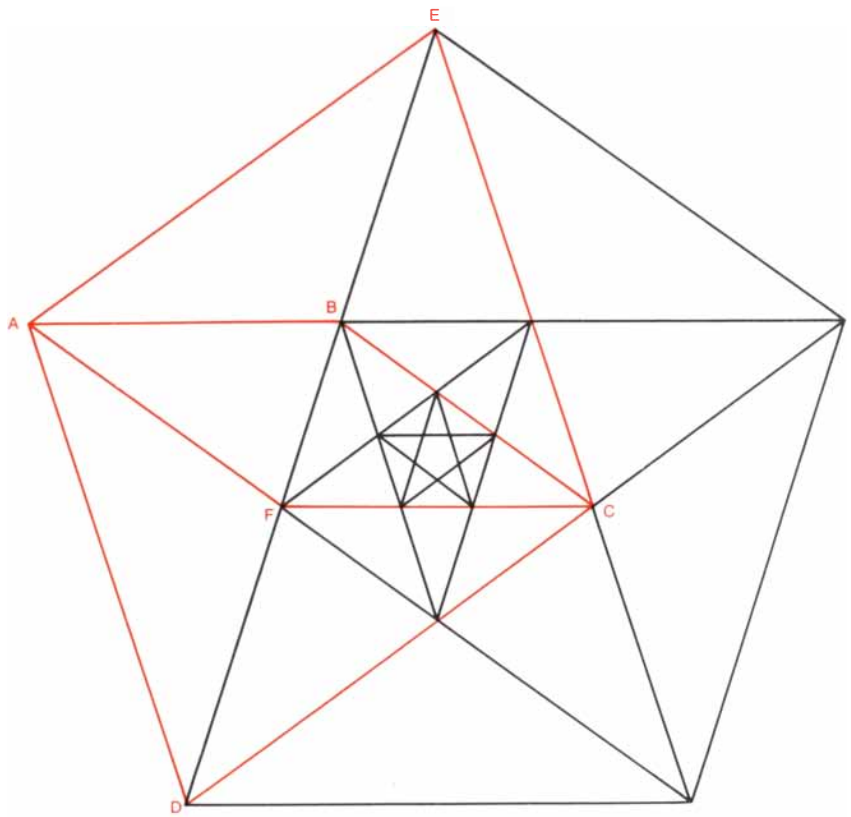
By dissecting darts and kites into smaller pieces and putting them together in other ways you can make other pairs of tiles with properties similar to those of darts and kites. Penrose found an unusually simple pair: the two rhombuses in the sample pattern in the bottom illustration on the opposite page. All edges are the same length. The larger piece has angles of 72 and 108 degrees and the smaller one has angles of 36 and 144 degrees. As before, both the areas and the number of pieces needed for each type are in the golden ratio. Tiling patterns inflate and deflate and tile the plane in an uncountable infinity of nonperiodic ways. The nonperiodicity can be forced by bumps and dents or by a coloring such as the one suggested by Penrose and shown in the illustration.

We see how closely the two sets of tiles are related to each other and to the golden ratio by examining the pentagram in the bottom illustration on this page. This was the mystic symbol of the ancient Greek Pythagorean brotherhood and the diagram with which Goethe's Faust trapped Mephistopheles. The construction can continue forever, outward and inward, and every line segment is in the golden ratio to the next smaller one. Note how all four Penrose tiles are embedded in the diagram. The kite is *ABCD*, the dart is *AECB*. The rhombuses, although they are not in the proper relative sizes, are *AECD* and *ABCF*. As Conway likes to put it, the two sets of tiles are based on the same underlying golden stuff.



Penrose's nonperiodic chickens

Are there pairs of tiles not based on the golden ratio that force nonperiodic tiling? Is there a single piece that tiles only nonperiodically? These questions define two of the most intriguing problems that remain to be solved in the theory of tiling.



The Pythagorean pentagram

© 1976 SCIENTIFIC AMERICAN, INC

BOOKS

The blowfly as a machine, the wild boy of Aveyron and modern merchant ships

by Philip Morrison

THE HUNGRY FLY: A PHYSIOLOGICAL STUDY OF THE BEHAVIOR ASSOCIATED WITH FEEDING, by V. G. Dethier. Harvard University Press ($30). The blowfly *Phormia regina* is a 30-milligram "jet jewel," three imperfect spheres in line, diaphanous wings, great eyes faceted in hexagonal elegance, body encased in iridescent semirigid tanned armor, "perfection in its parts." Professor Dethier, a biologist at the University of Massachusetts, has lived with the 60-million-year-old species for 30 years, asking questions. With a score of graduate students and many colleagues worldwide, he has built up an oversimplified but powerful model of blowfly feeding behavior (the greater task of blowfly life, reproduction, is set aside), which he outlines here in depth. He tells of old experiments and new; the history of experimentation throws the sharpest light on conclusions as well as on technique. Some of the experiments are wonderfully artless, such as letting the blowfly walk on a piece of paper moistened with a blue-dyed sugar solution; the fly's "lip marks" record its walk. Others are the fruit of the instrumental virtuosity of our times: scanning electron micrographs and fly electroencephalograms. Dethier looks at many other species from mouse to man for analogy or illuminating difference.

Blowflys are not only abstractly beautiful; they are intricate. Fly eyes are rivaled only by those of the cephalopods and the vertebrates; when the Spanish neuroanatomist Santiago Ramón y Cajal "first looked at the neural network in the eyes of flies, he exclaimed in amazement at their enormous complexity." They are of course superlatively adapted as motion detectors; try to catch a fly. They see the entire visible spectrum, they detect polarization and they find flicker at 20 times our own time resolution. For form perception we vertebrates have better electronics. Flies are attracted by spots of flickering light, by fly-size black objects, by the colors black and red. Their milligram of brain, with its 100,000 neurons, collects eye data and signals from the thousands of chemical, tactile and odor sensors fitted into the prickles and bristles of that fly armor, together with many internal signals, from the size of the well-fed insect's crop to the beat of the hidden clock that generally enforces a circadian rhythm.

First of all the fly must eat. It flies randomly, sensing air currents to fly upstream toward the source of desirable odors, while it monitors visual pattern to check its ground speed. Finally it flies in narrowing circles until some optical datum provides the cue for it to land. Once it has landed it walks across inert surfaces in short strolls, typically grooming itself before it takes a random turn to a new lap. If the surface smells (or tastes) unacceptable, the fly stops walking and soon flies off. On a delicious surface (all those legs are tasters) it follows remarkably regular patterns. The fly turns once a leg loses contact with the food. Or it executes a "dance," a stereotyped, if "noisy," motion within one small area, with its shape responding to the tilt of the surface, at least in the dark. Surely this foreshadows the signal dance of the social honeybee.

Once the food is located in this complex way, the fly drinks. (Not having a mouth, it only sucks liquids through its complex tubular proboscis.) More than half of the book deals with the task of eating. The fly must choose between acceptable foods and reject the harmful or unpleasant ones. (Foods in the wild are never pure.) It must know when to stop feeding as much as it must know when to seek food. It needs fresh water too; the blowfly can apparently detect pure water vapor as well as less common volatile compounds. A typical fly find, however, is the cloud of faintly alcoholic and yeasty vapor drifting from a pine tree, whose needles exude the sap a host of aphids has sucked out. The old droplets of honeydew are fermenting, and the fly is brought in from its random flight by the good smell. A walk, a minute's sucking, a walk, a repeat. After a few more droplets the fly ignores the next ones; it is replete. Then hours of inaction before a new flight. Day after summer day its life continues. Flies that are not in immediate need of food or water and are not mating are "going nowhere, doing nothing.... Most of the time is spent in pointless idleness" among flies and most other nonsocial animals.

What a fly takes for energy is sugars. "Next to water, sucrose is the most universally acceptable compound." The fly's own main blood sugar is an unusual dimer of glucose, trehalose. The story of what flies choose and avoid and how, of fly taste, of how ingestion is controlled, is the major tale here. Experimenters have flown flies around and around on a tether to test hunger; they have found delicate nervous circuitry that can be cut to induce a fly to eat without restraint (the obese flies are seen in the book), and they have grafted two flies together back to back, merging their bloodstreams to see if one was hungry after its partner was fed. (It wasn't.) The story is well elaborated. From our viewpoint the body of a mature fly is strange. Nothing of it grows or is replaced; an old fly is simply tattered and worn like an old coat. There is no new hair growth or repaired wounds; a fly responds to a wound like a radiator full of Neverleak, by a clot at the lesion. There are specific needs for protein (to make eggs and for other reproductive purposes), and there is a special overall response to shortening day length (a slowing down for overwinter survival). *Phormia* is a Temperate Zone animal. Its types of behavior can be fitted into a general pattern, and much subtle detail can be added in support—a bookful. The model is a big set of peripheral reflex loops that describe average behavior. To this must be added an internal source of variable actions, controlled by a hierarchy of internal firings that we do not yet see, plus an ebb and flow of overall "hormonal tides" and simpler reaction-rate dependences on chemistry and temperature. Surprisingly, no computer models of blowfly feeding are offered.

Flies can even be conditioned. In a subtle set of experiments Margaret C. Nelson was able to demonstrate that a blowfly could be trained to extend its proboscis for a sugar reward by a salt-water signal to its feet. In general, however, learning is weak in flies; they have neither social interactions nor nests nor fixed food sources. Why learn? One pine is like the next; maybe when learning is of no great value, the small brain is saved for a special-purpose computer. Flight, search, food choice and timing provide plenty of inbuilt needs for computing. Flies never become habituated to a wave of the human hand; they always respond by flight, until they are starved to immobility. They will, however, stop sucking sugar after taking too much. (It is not yet sure that the sensory apparatus is not simply showing fatigue.) All such doubts have been controlled out of the best conditioning experiments.

This handsome, expensive, compen-

© 1976 SCIENTIFIC AMERICAN, INC

"The most complete and most scholarly dictionary of the English language"

—The Christian Science Monitor

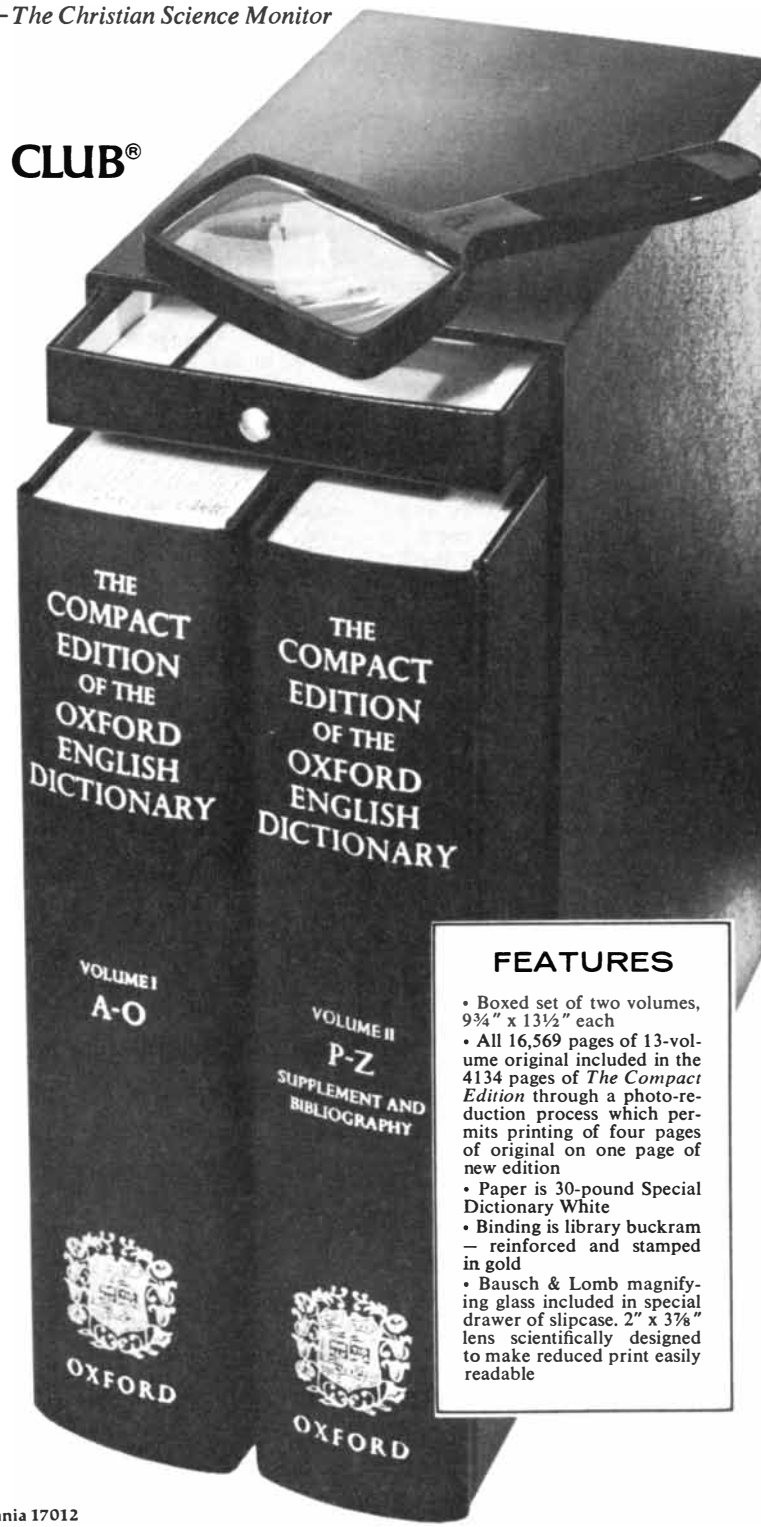
As an introduction to membership in the

BOOK-OF-THE-MONTH CLUB® yours for only

$17.50

Publisher's list price $95

THE SUGGESTED TRIAL: You simply agree to buy four Club choices within a year at substantial savings on most books you choose.



THE OXFORD ENGLISH DICTIONARY is generally regarded as the final arbiter of the English language. Until recently, it has been available only as a thirteen-volume set, currently priced at $395. Now, through the combination of an ingenious method of micrographic reproduction and a fine Bausch & Lomb optical lens, every single one of its 16,569 pages, 50 million words and close to 2 million illustrative quotations appears, in easily readable form, in *The Compact Edition.*

> Book critic Christopher Lehmann-Haupt of *The New York Times* has said of this edition: "It is something of a miracle....*The Compact Edition* is easier to work with than the original with its thirteen separate volumes."

Even more extraordinary, as a trial member of the Club you may obtain the two-volume set for only $17.50. And as long as you remain a member, you will receive the *Book-of-the-Month Club News,*® a literary magazine announcing the coming Selection and describing other important books, known as Alternates, most of which are available at substantial discounts—up to **40%** on more expensive volumes.

If you continue after your trial membership, you will earn at least one Book-Dividend® Credit for every Selection or Alternate you buy. These Credits enable you to acquire for your library a wide variety of books, called Book-Dividends, at astonishing savings—*at least 70% of publishers' list prices.*

All Book-of-the-Month Club books, including Book-Dividends, are identical to the publishers' editions in content, format, size and quality.

FEATURES

- Boxed set of two volumes, 9¾" x 13½" each
- All 16,569 pages of 13-volume original included in the 4134 pages of *The Compact Edition* through a photo-reduction process which permits printing of four pages of original on one page of new edition
- Paper is 30-pound Special Dictionary White
- Binding is library buckram — reinforced and stamped in gold
- Bausch & Lomb magnifying glass included in special drawer of slipcase. 2" x 3⅞" lens scientifically designed to make reduced print easily readable

BOOK-OF-THE-MONTH CLUB, INC., Camp Hill, Pennsylvania 17012

© 1976 SCIENTIFIC AMERICAN, INC

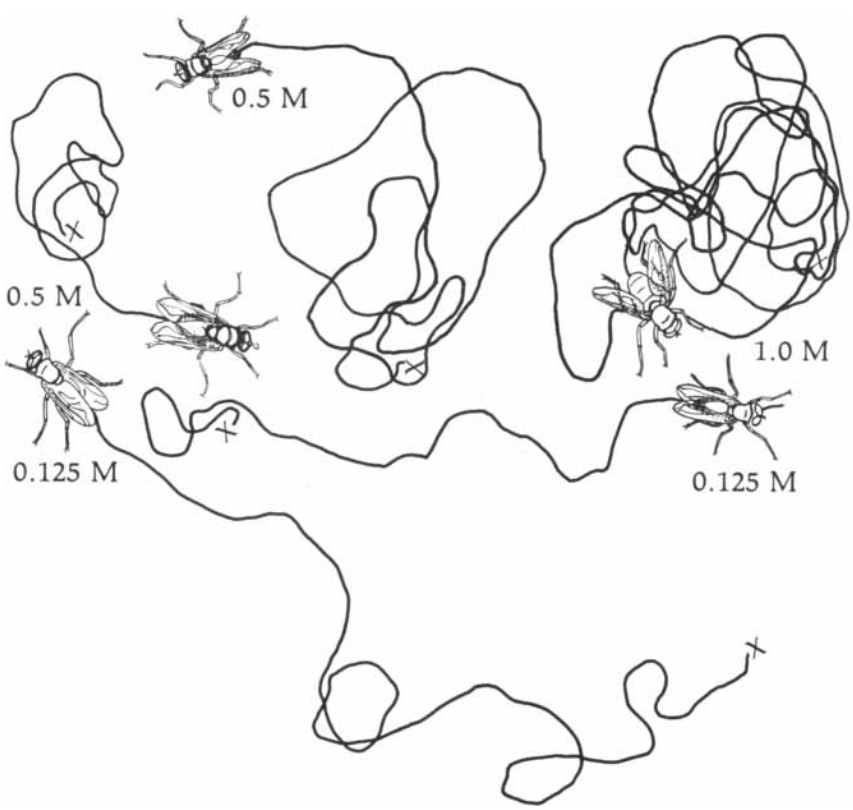
dious and learned volume is written with wit and clarity, although in many chapters it is highly technical. (There is, for example, a 30-page atlas of the blowfly brain.) The author could hardly have begun his book, as this reviewer could hardly have ended, without a citation from some metaphysical poet (here William Oldys): "Busy, curious, thirsty fly, / Drink with me, and drink as I; . . . / Make the most of life you may, / Life is short and wears away. . . . / Threescore summers when they're gone, / Will appear as short as one."

THE WILD BOY OF AVEYRON, by Harlan Lane. Harvard University Press ($15). The chief acts and ideas of the Enlightenment and its revolutionary aftermath have a local habitation: the Seine's Left Bank in Paris. It is the intellectual legacy of one event there that is chronicled in this lively, informed and arguable book. The author, a psychologist at Northeastern University, begins in the Luxembourg Gardens, where on a summer day in 1800 the wild child they had brought from the forest first met the young physician Jean-Marc-Gaspard Itard, a brilliant student of the psychiatrist Philippe Pinel (he who had just ordered the inmates of the asylums of Paris unchained). "I retraced their steps," Lane writes, in "the boiler-room 'archives' . . . , the dusty attics of the Sorbonne and School of Medicine (with the priceless view of Paris rooftops accorded only to students and cleaning ladies) [and] a dozen other places but especially the cavernous hushed reading rooms of the National Library: these were my joyful haunts." He brings from those quarters much solid material, forgotten accounts by various actors in the drama and a step-by-step unwinding of the origins and implications of Itard's five years of intimate work with the "disgustingly dirty" Victor, "who swayed back and forth ceaselessly like certain animals in a zoo." What we examine as much through debate as in careful narrative is the history of the concept of the perfectibility of humankind.

Itard saw hope for Victor. Pinel had given a darker prognosis, which he delivered before the Société des Observateurs de l'Homme. Lost for many years, the report was published in 1911; it was rediscovered by Lane, and he cites it at length. Pinel believed nothing whatever could be done for the boy he deemed to have been abandoned at the age of nine or 10 by his parents as being incapable of education. Itard, and following him Lane, reject the idea; they blame the boy's symptoms, including his inability to speak, on his enforced isolation. Lane, citing modern cases, rejects the diagnosis of retardation and of autism; the conclusion follows that "prolonged isolation deprived him of the crucial skill . . . of imitation," just as Itard had written that "the foremost faculty of human intelligence, that of imitation, is annihilated." Not every reader will be persuaded that Pinel was wrong; the issue seems to turn on the uncertain condition and age of the boy when he was left in the forest. All the comparisons seem to depend strongly on the developmental stage and the duration of isolation.

The theater of these ideas is much wider than the intimate two-person drama played out over five years. Victor lived in Paris on a state allowance until he died at about 40, still without speech, under the devoted care of his kind but uneducated governess. Itard went on to teach the deaf with striking success. It was the work of Étienne Bonnot de Condillac that gave him theoretical support. Itard saw education as being neither the imposition of an intrinsic logic that arose out of the subject matter of instruction nor a merely empirical set of practices and hints. For him, and for the hopeful philosophy of education that we still pursue, instruction is determined mostly by the person taught, by the careful observation of his maturing and his needs, which are not uninfluenced but are experimentally managed in their antecedents and their consequences. New needs are created, social and above all sensory, often by the invention of instructional devices. Itard used a black-painted board "on which everyday objects were placed and their outlines chalked," together with cutout letters, a red disk, a blue triangle, a black square, a drum, snuff, bells, a vase with nutmeats hot and cold—an arsenal of pedagogy mainly within the spirit of early schooling today. His later success with the deaf depended on his observation that many were not entirely devoid of hearing but could at least make out vowels quite well. He worked at extending their rudimentary discrimination of sounds and often saw great progress.

Itard was the founder of the oral education of the deaf, followed by such figures as Alexander Graham Bell. But Itard himself, who "never learned a sign during forty years among the deaf," became the champion of the opposing procedure, the language of signs. Since the 1760's, when deaf children brought to the home of the Abbé de l'Epée the signs they themselves had improvised to converse with their families, there has been a struggle between oral and sign instruction for the deaf. Today in the U.S. and Canada half a million people use the American Sign Language, which can be shown to have grown from Parisian roots. Thomas Gallaudet brought it to Hartford in 1816 with the deaf young Laurent Clerc, earlier a student and later a teacher at the Paris Institute for Deaf-Mutes; it was Gallaudet's language of sign that spread and grew into today's Ameslan. One great step remained: the sign code had to free itself



The effect of different concentrations of sugar on the form of the blowfly's "dance"

© 1976 SCIENTIFIC AMERICAN, INC

of the heavy burden of French grammar and inflection that had been added to encumber the "natural language" by the philosophical instructors at the institute. Like Chinese or even English, Ameslan has a structure of context and redundancy, with a swift grammar of position and gesture but no time for inflection or any apparatus that derives from Latinity.

Itard came at last to understand that, given the indispensable little community of the deaf, most of those who cannot hear consonants at all will profit by the sign language. They develop conceptually through that channel and they may later come to use their understanding to gain some control over oral speech. In 1827 Itard made a comparative test of two mute students. He found the student taught by signs to be superior, as has the latest careful study, made in the 1970's, of congenitally deaf children matched for I.Q., age and sex. The signing "twins" did significantly better in reading, writing, psychological adjustment, oral speech, school completion and college entrance. "But the Great Sign Controversy continues to this day."

Itard's student Édouard Séguin went on to teach not only the deaf but also the retarded. Plainly the narrower goals of such instruction made more evident the effectiveness of the method drawn from Condillac. First, Séguin taught, comes motor control, based on an analytical strategy of finding the elements of complex acts. For example, climbing a ladder, with the teacher climbing the other side of the ladder and holding the child's fingers against the rungs, proceeds by easy stages to independence. Sensory training comes in its turn: the sense of hot and cold, warm and cool, light and heavy. The drum and the bells follow, and with them music itself. The child is led to speech and finally to moral education, a domain for which Itard and Condillac had given little guidance.

The regime had its effect, as one American witness touchingly related in 1847. "During the past six months, I have watched... nearly one hundred fellow beings who were objects of loathing and disgust... these... I have seen properly clad, standing erect,... eating in an orderly manner at a common table,... gaining, by their own labor, the means of existence... and singing in unison songs of thanksgiving." Séguin was drawn by inner commitment into the French revolutionary storm of 1848; the Second Empire saw him an emigrant to the U.S., where he worked as a teacher and a consultant until his death in 1880.

The story of Victor is not over yet. Just before the turn of the century the courageous Maria Montessori became the first woman to receive a medical degree in Italy. She was posted to the asylums of Rome, where she found retarded children thrown indiscriminately

Realize your ideas at the working laboratory.

Draper Laboratory brings science into the public realm through design and prototype hardware development.

Our work is generally focused on inertial guidance and navigation systems for all environments and special applications such as gunsights and fire control systems.

Our opportunities lend themselves more to personal discussion. Therefore if your education and experience relate to these areas we welcome your response.

Guidance Algorithm Analysis. Navigation Algorithm Analysis. Orbital Mechanics. Digital Computer Simulation. On-board Targeting. Space Vehicle Guidance. Kalman Filtering. On-board Space Vehicle Navigation.

A/D Circuit Design. Radiation Environment, A/D Subsystems. A/D state-of-the-art Design and Analysis. Guidance and Navigation and Control Systems Design and Development. Pointer/Tracker Systems. Optical Instrumentation. Instrumentation Tribology/Bearings.

We offer salaries competitive with industry, a superior work environment in our new 450,000 sq. ft. facility adjacent to the MIT campus, an excellent benefits program that includes tuition assistance, and a scholarship aid program for children of staff members.

Reply by calling John J. McCarthy at (617) 258-2678 collect, or mail your resume to his attention.

CHARLES STARK
DRAPER LABORATORY
555 Technology Square
Cambridge, MA 02139

An equal opportunity and affirmative action employer.

PROGRAMMABLE
MODEL #4525
Scientist
$38.95
100 STEPS!!!

100 STEP LEARN MODE, KEYBOARD PROGRAMMING CAPABILITY.

• RPN logic • Rollable 4-level stack • 8-digit plus 2-digit exponent LED display • Scientific notation • Sine, cosine, tangent & inverse trigonometric functions • Common & natural logarithms & antilogarithms • Instant automatic calculation of powers and roots • Single-key square root calculations • Single-key Pi entry • Separate storage memory. • Square, square root and reciprocal calculations. Change sign & register exchange keys • Includes NiCad batteries.

Manufactured in USA
by NATIONAL SEMICONDUCTOR

One year unlimited warranty • 10 days money back guarantee.

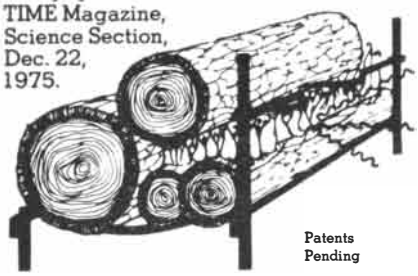
Mail Check to:
ILDAN Inc.
2901 SENTNEY AVE.
CULVER CITY, CA. 90230

FOR CHARGE CARD ORDERS
CALL: (213) 559-1044
Same day service if paid by money order

____4525 @ 38.95____
____AC Charger @ 4.95____
____Case @ 2.95____
____Calc. stand @ 2.50____
Cal. res. add 6% tx ____
Ship. hndl. & Ins. 3.75
TOTAL ____

"The Physicist's Fire"

The TEXAS FIREFRAME™ grate makes a *"hot, even, slow-burning fire; ... heat ... streams out into the room ... easy to light ... only a few sheets of newspaper."*
TIME Magazine, Science Section, Dec. 22, 1975.



Patents Pending

Unique, adjustable arms open up the red-hot core of the fire.
New era in home fire-making; save energy, wood, fuss, money and enjoy the fire with a TEXAS FIREFRAME grate.

Economy Model, 15 lbs. $29.95 • Heavy-duty, 30 lbs. $39.95 • Includes shipping and state tax • Both 25" front width, 21" back, 15" deep, 15" high • In handsome carton • NOVEL, PRACTICAL GIFT

Please send ____ Economy unit(s)
____ Heavy-duty unit(s)
Check for $____ enclosed.
Ship to: ____
Address ____
City ____ State ____ Zip ____

TEXAS FIREFRAME CO.
PO Box 3435 Austin, Tx. 78764

© 1976 SCIENTIFIC AMERICAN, INC

among the insane. She read Itard and Séguin ("the voice of a forerunner crying in the wilderness"), and within a decade she had opened her first Casa dei Bambini, where she offered to normal preschool children in her own way the hopeful opportunities Itard had set before poor Victor. The drum and the bell, the letter cutouts, the pairing of names and the objects they symbolize—these and much more of Montessori practice were first seen by the wild boy of Aveyron. Montessori's books and her original methods of instruction are now known worldwide. She died in Holland in 1952.

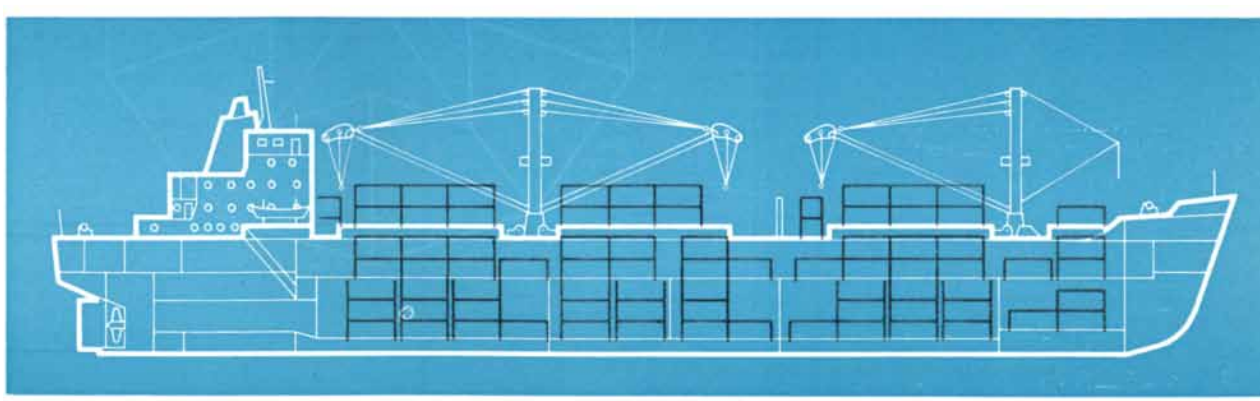
"Itard had set out to train an *enfant sauvage;* by his journey's end he had become the originator of instructional devices, the inventor of behavior modification . . . , creator of oral education of the deaf, and father of special education for . . . the handicapped." The "education of the senses" spread through Séguin to the universal view of education seen by Montessori. The end is not visible; even models as constraining as behavioral modification now appears to be are transient elements of a view of education based, like Itard's, on honest observation of the subtle needs of the learner and a profound respect for human potential. Excesses such as the shock from the Leyden jar and the frightening time Itard held Victor "out of the window, his head facing directly down toward the bottom of the chasm" are matched in our time by shock therapy, the demeaning reward of the tossed chocolate drop and the related pseudoconstancy of the I.Q. score. No, neither Itard nor the behavior modifiers of our own day are always successful or even plausible, yet there is a kernel of hope for a more perceptive pedagogy in the 100-year story of the wild child, if we read it in sympathy and wisdom.

MERCHANT SHIP TYPES, by R. Munro-Smith. Marine Media Management Limited, International Scholarly Book Services, Inc., Portland, Ore. ($35). Automobiles present themselves so persuasively to Americans that even the least concerned come to have a good deal of knowledge about them, their looks, use, trends, costs and so on. Many books serve the more interested, forming a literature one could call engineering appreciation, full of pictures, data and evaluation, but not at all aimed at automotive engineers. Here is just such a volume, not about roadways but about seaways. "Ships are the biggest moving objects ever made by man" proudly reads the caption to the frontispiece, which shows one of the huge crude-oil carriers under way. The volume is published for the British Institute of Marine Engineers; its author is an experienced and authoritative writer and lecturer on marine architecture. One would expect such a volume to originate away from these shores, since the ship is no longer salient in American life, even in the few most active ports. Those big tankers never enter our shallow harbors.

Yet shipping is a major economic activity and a big piece of world heavy industry. The world's ships could carry off to sea all the world's motor vehicles at once, although the steel invested in ships is only a modest fraction of that invested in automobiles. After all, a loaded ship floats with an overall density not so far from the density of water; except for fast and roomy passenger ships, the entire mass of a modern ship—its hull, equipment and engines—is only about a fourth of that of its cargo, fuel and expendable stores. Just about a third of all ships now afloat are steamships, with nearly all the shafts spun by turbine. The reciprocating steam engine "will soon become a museum piece." Almost all the rest are diesel-powered, although the seaborne gas turbine has much potential. Fuel costs favor the efficient diesel, but big ships need the smooth turbine's high power. In a good year—say in 1972—Japanese shipyards launch half of the world's new tonnage, in the spacious yards built after the war "on green-field sites." The big steel plates and sections are shot-blasted, cleaned and spray-painted, cut into accurate shapes by computer-controlled cutting torches, welded into subassemblies weighing hundreds of tons and built up into ships by giant cranes at a sloping seaside berth or in a dock that can be flooded. (Birth trauma is severe for both ships and mammals; the bigger the infant, the more difficult the launch.) The national flag under which more ships sail than any other is the Liberian, a legal "flag of convenience" (and tax benefits), most of whose proud vessels will never approach Liberia. Japan and Britain are more genuinely maritime runners-up (1973 figures from *Lloyd's Register of Shipping*).

The volume begins with a background chapter on design fundamentals, maritime practice, maritime economics, maritime law and then, chapter by chapter, there follow all types of civil ships, from proud liners (now almost all vacation cruise ships) to such auxiliaries as icebreakers, hydrofoils, tugs, trawlers and cable-layers. Photographs and line drawings of plans and profiles are here in plenty, with much tabulated data, both about particular ships and about such complexities as the approximate lengths of caught fish, the right storage temperatures for fruits and the size classifications of icebergs.

Tankers for crude and refined petroleum products, plus some chemicals, dominate all other types of vessel. It was in 1861 that the sailing brig *Elizabeth Watts* carried the first Pennsylvania crude oil out of Philadelphia to London in wood barrels. Today half of the world's total cargo capacity is in tankers, although they are on the average so big that they number only about a ninth of the world merchant fleet. This very cheap heavy haulage has its external costs: the risks of widespread spillage and pollution. The author alludes to but does not describe the newest tankers for cold liquid methane now being built south of Boston for Burmah Oil, the largest single merchant shipbuilding contract ever let and a litigious one. Those ships are a sight to see. The $100-million hull is a huge pod with hemispherical holes for five methane-holding peas, each one an insulated tank some 35 meters in diameter. The crane that hoists the peas into the pod is the biggest



Elevation of a Clyde general-purpose cargo vessel, from **Merchant Ship Types.** *Black rectangles are containers*

© 1976 SCIENTIFIC AMERICAN, INC

in the Western Hemisphere; only Japan has bigger ones. The concentrated energy store of the dense methane plainly has its own risks.

Bulk-cargo volume has much increased since the days of the tramp steamer, and specialized ships are now made for such trades, which often flow in only one direction. Iron ore, coal, bauxite, grain and sugar are the chief commodities of this type. Big ships with large hatches ply between ports that have specialized handling equipment to take the low-density stuff out fast with suction devices and the high-density with grabs. These are big, bland, simple-looking ships, with capacious holds and with double bottoms for ballast. There are ships designed for double duty, say iron ore out of icebound Labrador by summer and oil through the Tropics by winter. There are even several hundred sizable triple-threat ships dubbed OBO, able to handle by turns ore, bulk cargo and oil. The depressed state of world trade has becalmed not a few bulk carriers, one guesses. (Two float rustily in Boston backwater.)

The general cargo carrier is no less changed. The typical tramp has grown from some 2,000 tons' capacity to 14,000 and runs by diesel, but the busiest European and British ports are now served by ships designed to carry a unitized cargo, for ease of transshipment, quick turnaround, low damage and low longshore labor costs. There are four major innovations in this line. The most used is the container. Today the wheel is no more essential on the land than the box is on the sea, they say. These standardized boxes are loaded on and off the ship by big heavy-lift shore cranes and are hustled onto trucks or railroad cars with a formidable array of mobile lifts and straddlers. The ship at sea is now a shipful of boxes; there are about 400 such ships, with hatches so large that they eat up the deck, which remains only as a longitudinal stiffener. The bigger and faster container ships store a couple of hundred big boxes below decks and a three-layer pile of them above deck, as many as 600 boxes the size of a truck trailer.

This idea has been extended to ships that can carry not mere trailers but entire river barges, as many as 83 standardized steel barges carrying 375 tons each. Such a ship has a big gantry crane straddling its deck, rolling on tracks the length of the ship to handle the barges. Another system floats its barges, more than twice as big as the others, onto a 2,000-ton elevator available at water level through the open stern of the ship; it lifts the barges to deck level, where they are rolled into place. Roll-on, roll-off ships are fitted with ramps for driving loaded trucks right onto the ship, now a seagoing ferry. Trawlers too now drag their nets and trawls over a special stern, giving shelter to the net handlers,

QUESTAR 700

Called by Modern Photography "the best" the Questar 700 is unique in bringing to the telephoto market a lens that guarantees theoretical optical perfection and flatness of field from edge to edge, with no linear distortion. Precise focusing from infinity down to 10 feet is accomplished with a single turn of the focusing ring, and at the 10-foot distance the image magnification is 1:4, an unheard-of ratio in telephoto lenses. Only Questar can deliver such performance. With its standard screw-thread T-mount adapter system, the Questar 700 can be connected to virtually every 35 mm SLR camera. Write for our literature: Questar Corporation, Box Q700, New Hope, PA 18938.

Specifications

Number of Elements	4	Angle of Acceptance	3° 25'
Focal Length	700 mm.	Length	8 inches
Aperture Range	f8	Diameter	4⅜ inches
Clear Aperture	88.9 mm.	Weight	4 pounds
Minimum Focusing Distance	10 feet	Filter Provisions	35 mm.

© 1976 SCIENTIFIC AMERICAN, INC

Fun&Games

FAIRCHILD VIDEO CENTER

Game Console with digital remote controls contains Hockey & Tennis . . . $149.95

VideoCarts ..snap-in cartridges . 19.97 ea.

See Page 123 of December SA for details
More educational & game cartridges to come

VIDEO GAME CLEARANCE

Ranked by sophistication in category

ADVERSARY by Nat'l Semi . **$79.88**
6G, 2P, C, R, D, AC A

PONG by Atari **38.88**
1G, 2P, C, D, B, S

SUPERPONG 4G, C, D, B, S . . . **69.88**

MOS CHIP

Name of the Game I. 94.88
6G, 4P, C, R, AC, D, A

Optional Rifle 8.88

Name of the Game II 69.88
4G, 2P, C, R, AC, D A

G.I. CHIP

QUADTRONICS Q-476 . . . . . **59.88**
4G, 2P, C, R, S, 8D, AC, D

TELSTAR by Coleco **49.88**
3G, 2P, 3D, D, B

G.I. GAME CHIP. 20.00

CODE**
2P=2 players D=digital scoring
4G=4 games A=auto speed
C=color R=remote control S=slide control
AC=adapter incl. B=battery operated

National Semiconductor SUPER PROGRAMMABLE SALE

8-digit 100-step programmables with skip/delete features; can store multiple programs and constsnts in memory; include carrying case & Ni-Cd battery.

4515/4615 Mathmatician PR. . . 39.88
6035 Statistician PR 39.88
6025 Financier PR. 39.88
4525 Scientist PR 44.88

TEXAS INSTRUMENTS

SR-52 . . . . . $229.50
PC-100 Printer for 52, 56 . . . . 229.50
SR-52 Libraries...Math, Stat Finance, EE . . . . . 24.95 ea.
Navigation, Aviation, Survey .39.88 ea.
Blank Cards (40) 14.95

FREE with every SR-52 purchase until 1/15
GAMES Library...20 exciting games of skill, chance and competitive sports ! retail value $29.95

SR-56 . . . . . .89.75

MAIL ORDERS TO:

LECTRO-MEDIA LTD.
P.O. BOX 1770 SA1-7
PHILADELPHIA, PA. 19105
(215) 925-4477

OR VISIT OUR SHOWROOMS -
NEWMARKET, 2nd & Pine (215) 925-9090

CREDIT CARD ORDERS ONLY
Call Toll-Free 1-800-523-2906
PA & Canada Residents: 1-215-925-4477
LECTRO-MEDIA LTD. will match any advertised price..simply clip ad & attach

FREE Digital Themometer with each order.
SHIPPING: 2.50 for 1st unit, (GAMES $3.) .50 for each additional. AIR: add$2./unit
SR-52, PC-100: add$1.50 Pa Res: add 6% tax

easing their work and reducing the dangerous sidewise sweep of waves.

Icebreakers are of increasing value in the polar Arctic, the Baltic and the Great Lakes. They still usually work by sliding a sloping bow up over the ice to break it by vertical force, both during steady progress or by backing and ramming. Such a bow must be strong; the big Russian icebreakers make use of a special steel with high impact strength at low temperatures, in plates more than two inches thick. One Finnish icebreaker of modest size pumps air through holes in the bow; the bubbles reduce friction between the hull and the ice and allow higher speeds. Oceangoing tugs are more needed than ever; the most powerful ply out of Hamburg. They have a range of 20,000 miles for worldwide coverage and have a pull of 150 tons, enough to tow a fully laden 300,000-ton deadweight supertanker at more than seven knots.

Altogether today's ships are full of technological interest and novelty. They sail close to key problems of international economics and ecology. This volume is a good guide for any group involved in shipping, for the many in the news world who need to know more about ships and for libraries, particularly in port towns. It appears to be quite reliable, with very few inconsistencies in the wealth of data offered. Perhaps it is expensive for the general reader, but where else can one read of the modern cargo vessel *Pointe des Colibris,* "for the West Indies run." It carries containers, automobiles and refrigerated cargo, and it also has stainless-steel tanks designed to hold nearly 300 tons of a particular export fluid. Yo-ho-ho and a tanker of rum!

TIME AND THE CALENDARS, by W. M. O'Neil. Sydney University Press, International Scholarly Book Services, Inc., Portland, Ore. ($14.50). NORTH STAR TO SOUTHERN CROSS, AND TWELVE SKY MAPS FROM NORTH STAR TO SOUTHERN CROSS, by Will Kyselka and Ray Lanterman. The University Press of Hawaii, Honolulu (book $8.95, maps in folder $10). ALMANAC FOR COMPUTERS, by George H. Kaplan, LeRoy E. Doggett and P. Kenneth Seidelmann. Circular No. 155, United States Naval Observatory, Washington, D.C. 20390. (Single copies at no charge from the Director of the Nautical Almanac Office.) There is no more appropriate way to begin the new arbitrary cut in the smooth orbital round that we label A.D. 1977 than to review how they managed such cuts long ago and far away, old decisions that still mark our days. A Sydney psychologist, an amateur of analysis with a flair for personal exposition and a droll pen, has made the task attractive. His brief book is not easy, but it is an unhurried and explicit guide to the calendars of mankind, in Egypt, Rome and Babylon and in the Jewish, Islamic, Indian, Chinese and Mesoamerican forms. It abounds with the names of months, with day names in Latinate, Sanskrit, Teutonic, Slavic and other tongues. The account is fuller and more critical than the standard encyclopedias can be, although the author claims little originality. He has worked over considerable monographic material (limited mainly to English, as he disarmingly explains) and puts it forward informed by his own comments on the doubts and uncertainties that still dwell in this part of the historical forest.

Our January 1 (Janus was of course the two-faced Roman god of gates and doorways) began the year long before the famous Julian reform of 46 B.C., when the Roman year had grown "markedly out of phase with the seasons." Thus did Pope Gregory act in the same city 16 centuries later. Both rulers had good astronomical advice (Sosigenes and Clavius). Julius Caesar was tough enough to stretch the calendar year by two unprecedented intercalary months, plus a sprinkling of recurring changes in month-lengths, to add 10.25 days to the old Roman year. Just why he added those 67 extra days in 46 B.C. "escapes me," says the author after considerable discussion. The old story of the jiggling by Augustus to make his month no shorter than July is "fiction . . . enshrined" in standard references going back to perhaps the 13th century.

Islam alone has a purely lunar calendar; it ignores the solar year. It is the insoluble problem of a fit between the lunar month and the solar year, however, that dominates the book, as it does most calendars. Professor O'Neil has a good deal to say about the problem of the first visibility of the young crescent moon, an issue that plays a big role in Islam, in the Jewish calendar and particularly in China. He suggests, after semi-empirical studies of first visibility in Peking, that the scores of Chinese calendar shifts, made to balance lunar months and the solar year precariously by "a great deal of minor tinkering" over two millenniums, were related to actual crescent visibility, although not in a simple way.

This is a book rich in detail, placing arithmetical and astronomical demands on the reader but bound to deepen and rectify superficial views. The literature cited is fascinating; the entire work is likely to incite a number of projects and hobbyists. The ingenious theory of the Roman historian Dio Cassius, which accounts for the order of the planet names in our week by an astrological count, is here somewhat confused by a missing line. O'Neil's philological appendix is interesting. The sun's day becomes the Lord's day in the Romance languages; the Eastern Church avoids the planetary names (made official by Constantine),

© 1976 SCIENTIFIC AMERICAN, INC

Why people who don't join book clubs are joining the Quality Paperback Book Club

QPB is the *only* book club that offers the best of the large-size, well-printed, durable paperbacks being published today. QPB also gives a 20% discount on most Selections and even greater discounts on many other books.

151. $8.95

274. $4.95

118. $5.95

362. WATERSHIP DOWN by Richard Adams. $4.95

221. THE PILL VERSUS THE SPRINGHILL MINE DISASTER, IN WATERMELON SUGAR and TROUT FISHING IN AMERICA by Richard Brautigan. (3 Vols.) Pub prices total **$6.45**

235. THE CATALOG OF KITS by Jeffrey Feinman. Photos. **$6.95**

374. $7.95

372. $5.95

364. $3.95

317. $3.95

215. $4.95

371. THE RAYMOND CHANDLER OMNIBUS. Foreword by Lawrence Clark Powell. **$6.95**

352. AGAINST OUR WILL: Men, Women and Rape by Susan Brownmiller. **$4.95**

368. THE ANATOMY OF HUMAN DESTRUCTIVENESS by Erich Fromm. **$6.95**

119. $8.85

366. Dante's INFERNO. Illustrated by Gustave Doré. **$6.95**

132. CAT'S CRADLE, SLAUGHTERHOUSE-FIVE and BREAKFAST OF CHAMPIONS by Kurt Vonnegut. (3 Vols.) Pub prices total **$8.35**

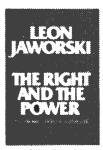
337. FRIENDLY FIRE by C D B Bryan. **$4.95**

310. THE FORMULA BOOK by Norman Stark **$5.95**

114. THE COMPLETE GRIMM'S FAIRY TALES. Introduction by Padriac Colum. Illustrated. **$5.95**

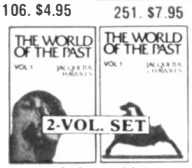
198. $8.95

356. $7.95



106. $4.95

251. $7.95



210. $11.90

353. THE BOOK OF ABIGAIL AND JOHN: Selected Letters of the Adams Family, 1762-1784. Edited by L. H Butterfield, Marc Friedlaender and Mary-Jo Kline **$4.95**

361. PORTRAIT OF PASTA by Anna Del Conte. Photographs **$5.95**

357. SUNFLOWER SPLENDOR Three Thousand Years of Chinese Poetry. Co-edited by Wu-chi Liu and Irving Yucheng Lo. **$6.95**

279. ZEN AND THE ART OF MOTORCYCLE MAINTENANCE: An Inquiry into Values by Robert M. Pirsig. **$4.95**

296. HOW TO FIX ALMOST EVERYTHING by Stanley Schuler New Revised Edition. Illustrated **$5.95**

260. $7.95

307 $7.95

228. $11.90

158. $3.95

128. $5.95

333. SIMPLE JUSTICE: The History of Brown v. Board of Education and Black America's Struggle for Equality by Richard Kluger. **$6.95**

171. AMERIKA, THE TRIAL and THE CASTLE by Franz Kafka (3 Vols.) Pub prices total **$7.35**

355. SAVING THE QUEEN by William F. Buckley, Jr. **$4.50**

231. TROPIC OF CANCER and TROPIC OF CAPRICORN by Henry Miller. (2 Vols.) Pub prices total **$9.90**

373. $3.95

204. $5.95

345. $3.95

146. $3.95

Let's try each other for six months...

CHOOSE ANY 3 BOOKS OR SETS FOR ONLY $1 EACH...

with no obligation to buy another book... and as a member you'll immediately qualify for our Free Book Plan

239. PSYCHOTHERAPY EAST & WEST, NATURE, MAN, AND WOMAN and THE WAY OF ZEN by Alan Watts. (3 Vols.) Pub prices total **$5.85**

360. APARTMENT GREENERY Growing Plants in Unpromising Places by Jean F. Blashfield Illustrated. **$4.95**

270. THE JAZZ BOOK From New Orleans to Rock and Free Jazz by Joachim Berendt. New Revised Edition. **$5.95**

133. THE DECLINE AND FALL OF THE ROMAN EMPIRE by Edward Gibbon. Abridgment by D.M. Low (3 Vols., Boxed) **$4.95**

253. THE ADVENTURES OF SHERLOCK HOLMES and THE MEMOIRS OF SHERLOCK HOLMES by A. Conan Doyle. (2 Vols.) Illus Pub prices total **$7.90**

Join the Paperback Revolution

qpb®

QUALITY PAPERBACK BOOK CLUB, INC.

The Book Club of the '70s

Middletown, Pennsylvania 17057

QUALITY PAPERBACK BOOK CLUB, INC.
Middletown, Pennsylvania 17057

7-0B170-1

Please enroll me as a member of the Quality Paperback Book Club and send me the three books or sets whose numbers I've indicated in the boxes below. Bill me only $3 for all three, plus shipping. As a member, I am not required to buy another book but for every Main Selection I do buy for which there is an established publisher's list price, I will receive a 20% discount of such list price. You will send me the *QPB Review* for at least six months with the understanding that you may cancel my membership at any time thereafter if I have not made—and paid for—at least one purchase in any six-month period. A shipping charge is added to all shipments.

INDICATE BY NUMBER YOUR THREE BOOKS OR SETS [ ] [ ] [ ]

16

Mr.
Ms. (Please print plainly)

Address _____ Apt. _____

City &
State _____ Zip _____

0-14

HOW MEMBERSHIP WORKS

1. You will receive the *QPB Review* fifteen times a year—about every 3½ weeks. This informative catalog describes the Main Selection plus approximately one hundred Alternates. **On any Main Selection you decide to take, for which there is an established publisher's list price, you will receive a discount of 20% of such list price.**

2. If you want the Main Selection, do nothing. It will be shipped to you automatically. If you want one or more of the Alternates—or no book at all—just indicate your decision on the reply form always enclosed with the *Review* and return it by the date specified.

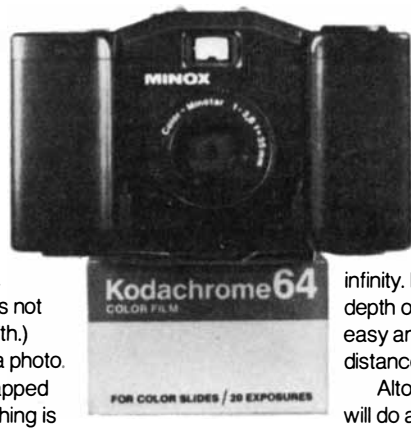
3. Free books. For every book or set you buy (exclusive of your three introductory choices for $1 each), you earn at least one Bonus Point, and the Bonus Points you accumulate entitle you to free books and sets. (You pay only shipping charges.)

4. Return Privilege. If your *Review* is delayed in the mail and therefore you receive the Main Selection without having had ten days to notify us if you did not want it, you may return that Selection at our expense.

5. Cancellation privilege. You may cancel your membership whenever you wish simply by notifying us. We will send you the *QPB Review* for at least six months with the understanding that we may cancel your membership at any time thereafter if you have not made—and paid for—at least one purchase in any six-month period.

© 1976 SCIENTIFIC AMERICAN, INC

Introducing: The world's smallest full-frame 35mm camera.



This is the one 35mm camera you can always carry with you everywhere.

The Minox® 35 EL weighs only seven ounces. It slips into your pocket as easily as a pack of cigarettes.

(And photography is not dangerous to your health.)

After you've taken a photo, the camera can be snapped shut like a turtle. Everything is securely protected inside the fiberglass-reinforced Makrolon shell.

The Minox 35 EL has fully automatic exposure control. You set the aperture. The camera sets the shutter speed. Shutter speeds range from 1/500 to 30 seconds. (That's right, automatic exposures for ASA 25 film can be as long as half a minute.)

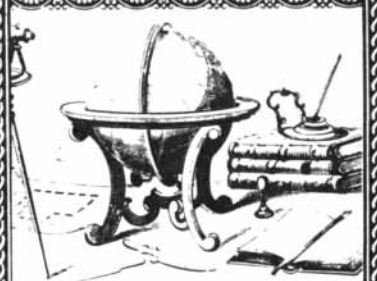
Accurate exposure is half the story. The precision optical system is ideal for almost any shot. The f/2.8 lens is of 35mm focal length, a moderate wide angle with great depth of field. At f/11 with the lens set at 15 feet, pictures are sharp from 7 feet to infinity. Because of the great depth of field, focusing is easy and is not critical at most distances and apertures.

Altogether, this camera will do almost anything an SLR with a normal lens will do. Plus one more thing that can make all the difference in the world. The Minox is small enough to be in your pocket when that once-in-a-lifetime photographic opportunity occurs.

And isn't that what photography is all about anyway?

For information contact Minox U.S.A., division of E. Leitz, Inc., Rockleigh, N.J. 07647. Dept. JJ1.

4512



Write for a free catalog of technical and science books.

Please send me a free copy of the technical and science book catalog.

name

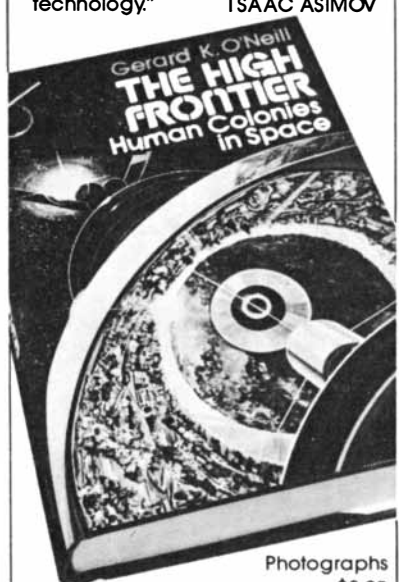
street address

city state zip

mail to: ASUC Store
Mail order dept.
University of California
Berkeley, Ca. 94720

By the year 2150 more people may be living in space than on earth –

The Princeton University physicist who devised the idea discusses all aspects of space colonization: how and when it will be done. "The establishment of such space colonies is possible and even practical in terms of present-day technology." ISAAC ASIMOV



Photographs
$8.95

WILLIAM MORROW
105 Madison Avenue, New York, N.Y. 10016

and so does modern Greek; the Dutch and the Germans mark Mars's day as a day of service ("To whom I am not sure"); the Sanskrit names were translated in India from Constantine's list, and in Sweden every week *lordag,* "more literally soap-day," precedes *sondag.*

The southernmost of our states is Hawaii. Two stargazers there, Kyselka, a planetarium lecturer, and Lanterman, an artist and a devoted amateur astronomer, have together produced the related publications we list. The first is a small, attractive book, a kind of field guide to the stars, with a simple text whose core is a series of maps of the night sky, one for each month along the meridian, from the North Star literally to the Southern Cross. (New Zealand sky watchers are not well served, since the southernmost polar cap is not treated. To be sure, no more of the sky can be seen from any of our states.) The maps are well made, with white star images on a deep blue ground and the conventional marks, such as the simplified constellation outlines and star names, done in a lighter blue. The treatment of the zodiac and the precession of the equinoxes is particularly careful, with the two polar caps being mapped in this context with the path of precession indicated. The text takes up astronomy old and new, with brief stories of the constellations and a set of selected short subjects, from the Crab Nebula to black holes and gravitational pulses. The book is not quite the first book about the stars one might read, but it could well be the second. Coordinates are little used, equations not at all.

In the accompanying folder the deep blue maps that grace the book are enlarged to full newspaper-page size on strong paper stock, more legible, fit for study and admirably decorative. They present stars to the fifth magnitude quite completely, and they also locate the brighter half of the objects of M. Messier. The meridian high these January evenings offers the best show in the sky, in bright stars, clusters, bright galaxies and a few easy variables. Again, no stars are mapped south of the Large Cloud of Magellan.

A few words in favor of Pythagoras. Number rules some worlds, in particular those of navigation and precision geodesy. Practitioners of these arts are sure to own the current volume of the annual *Nautical Almanac* (or its variations for airmen and landlocked surveyors). That worthy book, whose ancestors mark more than two centuries of service, is a thick volume of some 500 close-packed tabular pages for locating and timing the sun, the moon, the stars, the planets and even the moons of Jupiter and Saturn. The much smaller *Almanac for Computers* is perhaps a sign of a profound change to come. Its purpose is not to put into the reader's hands the tabular positions computed to high pre-

© 1976 SCIENTIFIC AMERICAN, INC

cision by the official groups in the three nations that annually publish them. Rather, here are the essential numerical data with which a meticulous user of a pocket calculator (arithmetic or scientific) can reach the same end. The aim is practical; we do not see the heroic trigonometric series the *Nautical Almanac* experts use, with the perturbations neatly laid out. The formulas are carefully smoothed matches of the functions of the *Almanac* (which remains the standard) fitted for each purpose—sun position, planet position and so on—over a particular time period. The numbers given are mainly coefficients of a polynomial expansion due to Chebyshev, but the user need know nothing of such mysteries. The lists yield the desired tabular values after a few multiplications and summings, with a simple function of the time as a variable. The expansion was chosen because it lends itself to easy truncation; if less accuracy will do, only a few of some 20 listed coefficients need be used. The decision can be made almost by inspection of the list with one's precision limits in mind. The tables given here will reproduce the main data of the *Almanac* itself to an accuracy of about a second of arc.

We can only dream, however, of navigating for a lifetime with a carefully written three-by-five card and a calculator! The tabular fits work only over a surprisingly limited time; some of them are listed afresh for every few weeks. (The root cause is apparently the earth's rapid nutation, our gyroscope kicked by the tilted orbit of the moon. To be sure, such effects are small, only seconds of arc.) For precision many terms are needed. Overall the *Almanac for Computers* is largely a computer-printout paperback of some 120 pages, saving a factor of four or so over the complete tables. The work is brand-new and frankly an experiment. If you are an agile computer with a serious interest in sky positions, you can obtain a single copy without charge by writing as indicated above; amateurs are welcome. A few years of public response will help to fix the eventual form of the almanac for the age of microelectronic chips. One minimal version would probably be popular: the sun, the moon, the stars and the planets to an accuracy of several minutes, one part in 1,000, just right for telescopic viewfinding or small-boat navigation. A first estimate would put the bulk of such tables at only 30 pages.

THE LONG-RUN AVAILABILITY OF PHOSPHORUS: A CASE STUDY IN MINERAL RESOURCE ANALYSIS, by Frederick J. Wells. Johns Hopkins University Press ($9.50). In one of the cleverest of Bloomsbury novels an eccentric physiologist bored everyone with his antic concern for phosphorus. That element, indispensable to all life (consider the nucleic acids and ATP), steadily drains from our fields to the bottom of the sea. In 50 years such fears have mutated from the amusing to the trendy.

Indeed, five years ago a research organization in the Middle West published a serious estimate of the total phosphorus potential of the planet implying that, do what we would, supplies of the element would run out "before the end of the 21st century." This small, careful, somewhat mathematical study is a reexamination of that forecast, presented as an instance of mineral-resource analysis. The critical task of the author, a senior analyst at another think tank, is made easy by the fact that the first argument patently contained a decimal-place error that put the overall resource base too low by a factor of 1,000! There are few rejoinders to any analysis that can take off from a runway that has been so well cleared.

The new result should reassure us about phosphorus even as it displays the fallibility of modelmaking (not excluding the author's). Several interesting results appear. A Dutch professor of theoretical crop husbandry has set the ultimate carrying capacity of the earth, given maximum yields, vegetarian diets and the extension of agriculture to desert and tundra, at about 1,000 billion people. It seems plain, however, that the true limit is set not by the area of cropland but by living and working space, the need for natural environment and the load of wastes. An adequately fed population of about 75 billion, with parks and living space, would be half again as dense worldwide as the population of Japan is today.

Even that hubbub of a world, with a density almost 20 times the present one, would not be limited by phosphorus. The main lesson of this study is that recycling and the careful control of loss (with conservation of the phosphorus content of sewage, manure, compost and bone, including human bone) can go a long way toward supplying phosphorus needs. Measures of a more extreme kind are mining common rock for its phosphorus (which is not much less concentrated there than the metal in the ores today being successfully exploited for copper) and eliminating the crop cycle in favor of the direct synthesis of food. In any plausible world, frugal and technically advanced, with say 20 billion eating at table, the runout time of phosphorus is surely thousands of years, perhaps millions. Common-rock mining for it is still a long way off; the realistic question is not whether we will run out of the essential atom but how much more it will cost us to extract it. Perhaps the deep sea will yield economically to the bottom miner. Once again it is today's poor we need to worry about, and gross neglect of the contemporary environment. Posterity is likely to manage better than we do, if it has the energy (both literal and figurative).

© 1976 SCIENTIFIC AMERICAN, INC

BIBLIOGRAPHY

Readers interested in further reading on the subjects covered by articles in this issue may find the lists below helpful.

LEGAL ABORTION

JOINT PROGRAM FOR THE STUDY OF ABORTION (JPSA): EARLY MEDICAL COMPLICATIONS OF LEGAL ABORTION. Christopher Tietze and Sarah Lewit in *Studies in Family Planning,* Vol. 3, No. 6, pages 97–122; June, 1972.

THE ABORTION EXPERIENCE: PSYCHOLOGICAL AND MEDICAL IMPACT. Edited by Howard J. Osofsky and Joy D. Osofsky. Harper and Row, Publishers, 1973.

LEGALIZED ABORTION AND THE PUBLIC HEALTH: REPORT OF A STUDY BY A COMMITTEE OF THE INSTITUTE OF MEDICINE, MAY, 1975. Institute of Medicine. National Academy of Sciences, 1975.

ABORTION SURVEILLANCE: ANNUAL SUMMARY 1974. Center for Disease Control. Public Health Service, U.S. Department of Health, Education, and Welfare, 1976.

MORTALITY ASSOCIATED WITH THE CONTROL OF FERTILITY. Christopher Tietze, John Bongaarts and Bruce Schearer in *Family Planning Perspectives,* Vol. 8, No. 1, pages 6–14; January/February, 1976.

AGRICULTURE WITHOUT TILLAGE

CONSERVATION TILLAGE: THE PROCEEDINGS OF A NATIONAL CONFERENCE SPONSORED BY THE SOIL CONSERVATION SOCIETY OF AMERICA IN COOPERATION WITH THE AMERICAN SOCIETY OF AGRONOMY, AMERICAN SOCIETY OF AGRICULTURAL ENGINEERS, AND NATIONAL ASSOCIATION OF CONSERVATION DISTRICTS, MARCH 28–30, 1973, IN DES MOINES, IOWA. Soil Conservation Society of America, 1973.

INFLUENCE OF LONG TERM TILLAGE, CROP ROTATION, AND SOIL TYPE COMBINATIONS ON CORN YIELD. D. M. Van Doren, Jr., G. B. Triplett, Jr., and J. E. Henry in *Soil Science Society of America Journal,* Vol. 40, No. 1, pages 100–105; January–February, 1976.

THE QUANTUM MECHANICS OF BLACK HOLES

THE FOUR LAWS OF BLACK HOLE MECHANICS. J. M. Bardeen, B. Carter and S. W. Hawking in *Communications in Mathematical Physics,* Vol. 31, No. 2, pages 161–170; 1973.

BLACK HOLES AND ENTROPY. Jacob D. Bekenstein in *Physical Review D,* Vol. 7, No. 8, pages 2333–2346; April 15, 1973.

PARTICLE CREATION BY BLACK HOLES. S. W. Hawking in *Communications in Mathematical Physics,* Vol. 43, No. 3, pages 199–220; 1975.

BLACK HOLES AND THERMODYNAMICS. S. W. Hawking in *Physical Review D,* Vol. 13, No. 2, pages 191–197; January 15, 1976.

THE ANTIBODY COMBINING SITE

BINDING OF 2,4-DINITROPHENYL COMPOUNDS AND OTHER SMALL MOLECULES TO A CRYSTALLINE λ-TYPE BENCE-JONES DIMER. A. B. Edmundson, K. R. Ely, R. L. Girling, E. E. Abola, M. Schiffer, F. A. Westholm, M. D. Fausch and H. F. Deutsch in *Biochemistry,* Vol. 13, No. 18, pages 3816–3827; August 27, 1974.

THREE-DIMENSIONAL STRUCTURE OF IMMUNOGLOBULINS. David R. Davies, Eduardo A. Padlan and David M. Segal in *Annual Review of Biochemistry,* Vol. 44, pages 639–667; 1975.

THREE-DIMENSIONAL STRUCTURE, FUNCTION AND GENETIC CONTROL OF IMMUNOGLOBULINS. Roberto J. Poljak in *Nature,* Vol. 256, No. 5516, pages 373–376; July 31, 1975.

ROTATIONAL ALLOMERISM AND DIVERGENT EVOLUTION OF DOMAINS IN IMMUNOGLOBULIN LIGHT CHAINS. A. B. Edmundson, K. R. Ely, E. E. Abola, M. Schiffer and N. Panagiotopoulos in *Biochemistry,* Vol. 14, No. 18, pages 3953–3961; September 9, 1975.

THE PERCEPTION OF MOVING TARGETS

INHIBITION AND DISINHIBITION OF DIRECTION-SPECIFIC MECHANISMS IN HUMAN VISION. Eugene Levinson and Robert Sekuler in *Nature,* Vol. 254, No. 5502, pages 692–694; April 24, 1975.

MOVING VISUAL PHANTOMS: A NEW CONTOUR COMPLETION EFFECT. Paul Tynan and Robert Sekuler in *Science,* Vol. 188, No. 4191, pages 951–952; May 30, 1975.

THE INDEPENDENCE OF CHANNELS IN HUMAN VISION SELECTIVE FOR DIRECTION OF MOVEMENT. Eugene Levinson and Robert Sekuler in *The Journal of Physiology,* Vol. 250, No. 2, pages 347–366; September, 1975.

EXOELECTRONS

DER METALLISCHE ZUSTAND. Johannes Kramer. Vandenhoeck & Ruprecht, 1950.

A SURVEY OF EXO-ELECTRON EMISSION PHENOMENA. L. Grunberg in *British Journal of Applied Physics,* Vol. 9, No. 3, pages 85–93; March, 1958.

REGISTRATION AND MAPPING OF PLASTICALLY DEFORMED METAL SURFACES BY MEANS OF PHOTOELECTRONS. C. C. Veerman in *Materials Science and Engineering,* Vol. 4, No. 6, pages 329–342; September, 1969.

THE DETECTION OF FATIGUE DAMAGE BY EXOELECTRON EMISSION. William J. Baxter in *Journal of Applied Physics,* Vol. 44, No. 2, pages 608–614; February, 1973.

NEUE ENTWICKLUNGSTENDENZEN IN DER ANWENDUNG DER EXOELEKTRONENEMISSION (FORTSCHRITTSBERICHT). W. Lausch in *Kernenergie,* Vol. 17, No. 5, pages 129–136; May, 1974.

CRATERING IN THE SOLAR SYSTEM

GEOLOGY OF THE MOON: A STRATIGRAPHIC VIEW. Thomas A. Mutch. Princeton University Press, 1972.

MOONS AND PLANETS: AN INTRODUCTION TO PLANETARY SCIENCE. William K. Hartmann. Bogden & Quigley, Inc., Publishers, 1972.

PLANETARY GEOLOGY. Nicholas M. Short. Prentice-Hall, Inc., 1975.

MOON MORPHOLOGY: INTERPRETATIONS BASED ON LUNAR ORBITER PHOTOGRAPHY. Peter H. Schultz. University of Texas Press, 1976.

THE SAMARITANS

THE SAMARITANS: THEIR HISTORY, DOCTRINES AND LITERATURE (THE SCHWEICH LECTURES, 1923). Moses Gaster. Oxford University Press, 1925.

SAMARITAN PENTATEUCH and SAMARITANS in *The Interpreter's Dictionary of the Bible: Vol IV.* Abingdon Press, 1962.

THE SAMARITANS, THE EARLIEST JEWISH SECT: THEIR HISTORY, THEOLOGY AND LITERATURE. James A. Montgomery. Ktav Publishing House, Inc., 1968.

MATHEMATICAL GAMES

CENTRAL TESSELLATIONS. Michael Goldberg in *Scripta Mathematica,* Vol. 21, No. 4, pages 253–260; December, 1955.

THE UNDECIDABILITY OF THE DOMINO PROBLEM. Robert Berger in *Memoirs of the American Mathematical Society,* No. 66; 1966.

UNDECIDABILITY AND NONPERIODICITY FOR TILINGS OF THE PLANE. Raphael M. Robinson in *Inventiones Mathematicae,* Vol. 12, Fascicule 3, pages 177–209; 1971.

THE ROLE OF AESTHETICS IN PURE AND APPLIED MATHEMATICAL RESEARCH. Roger Penrose in *Journal of the Institute of Mathematics and Its Applications,* Vol. 10, No. 7–8, pages 266–271; July–August, 1974.

© 1976 SCIENTIFIC AMERICAN, INC

Is there enough fertile land to feed your children?

Yes. The problem is that only a third of the world potential is now under cultivation, and, on much of that, output is hampered by antiquated methods.

FMC works hard at these problems. All over the world, our technology is turning neglected lands into productive regions. Our pesticides are saving crops, our mechanized equipment intensifying food production.

Food preparation and canning systems engineered by FMC are operating in some 50 countries, helping not only to feed the people of those countries, but, in many instances, enabling them to market their surpluses in other parts of the world.

Tell us your needs. FMC serves worldwide markets for food and agricultural machinery, environmental equipment, construction and mining equipment, industrial and agricultural chemicals. FMC Corporation, 200 East Randolph Drive, Chicago 60601.

FMC

© 1976 SCIENTIFIC AMERICAN, INC

You're going to spend $300 for a receiver? And you never heard of PPR?

Stereo Receivers	Sugg. Ret.† Price	Min. RMS Power Per Channel into 8 Ohms	Total Harmonic Distortion at Rated Power (Max.)	FM Sensitivity IHF '58	FM Sensitivity Stereo -50dB*
SA-5760	$799.95	165 watts from 20Hz–20kHz	0.08%	1.8μV	35.7dBf
SA-5560	499.95	85 watts from 20Hz–20kHz	0.1	1.8μV	36.2dBf
SA-5460	399.95	65 watts from 20Hz–20kHz	0.1	1.8μV	36.2dBf
SA-5360	299.95	38 watts from 20Hz–20kHz	0.3	1.9μV	37.2dBf
SA-5160	229.95	25 watts from 30Hz–20kHz	0.5	1.9μV	37.2dBf
SA-5060	169.95	12 watts from 40Hz–20kHz	0.9	2.0μV	38.2dBf

†Technics recommended price, but actual retail price will be set by dealers. *Now IHF '75 standard.

PPR is price performance relationship. And we feel it's a meaningful way of judging a receiver because it can tell you how much power, technology and performance you're getting for your money.

And when you look at our price performance relationship it's easy to see why your next receiver should be a Technics.

Of course, we want you to listen to our receivers. Especially since all six have the reserve power to float through complex musical passages with a minimum of distortion and clipping. And they all have rugged transformers. Bridged rectifiers. As well as high-capacitance filtering.

Play a record. You'll hear it the way it was recorded. Quietly and with greater dynamic range. Because we use an overload-resistant 3-stage IC in the phono equalizer sections.

Tune in an FM station. Even a weak one. In addition to hearing all the music, you'll also get increased stereo separation. Negligible distortion. And a minimum of noise. Thanks to flat group delay filters and Phase Locked Loop IC's in the tuner sections.

So if you'd like to know a simple way to discover just some of the reasons why a Technics receiver is so good . . . it's as easy as PPR.

All cabinets simulated wood.

Technics
by Panasonic

© 1976 SCIENTIFIC AMERICAN, INC